PREDICT BEHAVIOURAL SCORES IN SLEEP APNEA PATIENTS FROM RESTING-STATE NEAR-INFRARED SPECTROSCOPY (fNIRS)

by

AMNAH ABDELRAHMAN

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfilment

of the Requirements
for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

THE UNIVERSITY OF TEXAS AT ARLINGTON

MAY 2021

To my world

ACKNOWLEDGEMENTS

Thanks to my supervisor, Dr. Yingying Zhu who guided me under her supervision and for introducing me to this research area. My thanks go to Dr. Amal Isaiah who has been supportive, generous with his time and knowledge, and he trusted us throughout this work.

Words may not be able to express my thanks to my beloved mother Fatima AL-Amin and my dear husband Dr. Amar Mahgoub for their understanding, encouragement, spiritual and unconditional love and care. Thanks to my entire family and friends.

Thanks goes to the UTA family. I learned a lot through this experience and the experiences yet to come.

May 7, 2021

ABSTRACT

PREDICT BEHAVIOURAL SCORES IN SLEEP APNEA PATIENTS FROM RESTING-STATE NEAR-INFRARED SPECTROSCOPY (fNIRS)

Amnah Abdelrahman, M.S
The University of Texas at Arlington, 2021

Supervising Professor: Yingying Zhu

Sleep disorders are common among adults and children; They has serious consequences on their heath, cognitive development, and quality of life. However, some sleep disorders are challenging to diagnose and more challenging to treat. Practitioners often rely on apnea-hypopnea index (AIH) for obstructive sleep apnea (OSA) patients' classification task, where considering one measurement could raise the risk of oversimplification. Studies show the correlation between sleep disorders, specifically OSA, and mental health. On the other side, there are an increasing number of studies suggesting evidence of a relationship between the dynamic properties of functional brain structure and the behaviors and cognition attributes.

This novel work objective is to assist medical professionals in screening and treating sleep disorders and any related cognitive disorders as a consequence. In this work, we consider the novel problem of detecting behavioral measurements based on the resting-state functional near-infrared spectroscopy (fNIRS) time series data. Taking the full advantages of the fNIRS neuroimaging technique, which is described as an emerging and promising technology for a wide range of applications and as a diagnosing tool. In addition, it is a useful neuroimaging technology for

research on children's cognitive development without the limitation of functional magnetic resonance imaging (fMRI).

The foundation of the study is capturing the spatiotemporal dynamic features of the resting state FC and fully uses the captured dynamic information for the prediction problem. To do so, we applied a 4D tensor method to build the dynamic functional connectivity matrices. Then the tensor statistical model trained on dynamic functional connectivity (dFC) matrices for the classification task. The high classification accuracy achieved was 74%.

Key words: Resting-state functional near-infrared spectroscopy (fNIRS), Sleep Apnea, Apnea-hypopnea index (AHI), Dynamic functional connectivity (dFC), Tensor Analysis.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF ABBREVIATIONS

fNIRS	functional near-infrared spectroscopy
ADHD	attention deficit hyperactivity disorder
AHI	apnea-hypopnea index
ASD	autism spectrum disorder
BOLD	blood oxygenation level dependent
BRIEF	Behavior Rating Inventory of Executive Function
CBF	cerebral blood flow
CNN	convolution neural network
CPM	Connectome-Based Predictive Modeling
CSA	central sleep apnea
dFC	dynamic functional connectivity
EF	executive functioning
FC	functional connectivity
fMRI	functional magnetic resonance imaging
GLM	General Linear Model
HbO2	oxygenation hemoglobin
HHb	deoxyhemoglobin
ICA	independent component analysis
LSTM	Long-term short-term memory
MVPA	Multivariate pattern analysis or multivoxel pattern analysis
OSA	obstructive sleep apnea
PCA	principal component analysis
PFC	prefrontal cortex region
PLS	partial least squares regression
RNN	recurrent neural network
ROI	regions of interest
RSA	Representational Similarity Analysis

RSS	Residuals Sum of Square
SVM	support vector machine
tHb	total hemoglobin

CHAPTER 1

INTRODUCTION

1.1 Introduction and Background

Children's and adolescent's behavior problems are very common, peculiarly the problems associated with sleeping disorders (Kotagal & Pianosi, 2006; Mindell & Meltzer, 2008). Many studies suggest a direct relationship between sleep problems and neurobehavioral functioning development because of its protracted course of maturation. Sleep problems and short sleep duration contribute to tiredness, lower energy, and obstruct children from developing executive functioning (EF) abilities and cognitive development in general. A study conducted by (Turnbull et al., 2013) shows that 15% to 30% of children between 2 and 5 years old have some sort of sleeping difficulties. Furthermore, among school-aged children 6- to 12-years-old, 11 to 15 % experience behavioral sleep problems. Children experiencing sleep disorders are often diagnosed with attention deficit hyperactivity disorder (ADHD) and autistic spectrum disorder (ASD) (Turnbull et al., 2013).

This issue brings attention to the importance of extensive examination of the impact of sleep problems on children's neurobehavioral functioning. However, there is more research done for adults than for children (Turnbull et al., 2013). The research examines the correlation between the sleep problems and the involved fields of cognitive developmental psychology, neuroscience, and pediatric sleep (Pavlova & Latreille, 2019). This would help to better understand the effects of sleep on the brain. Therefore, develop better practices to deal with various conditions.

Cognitive development is associated with the changes in the prefrontal cortex region (PFC) of the brain. Damages in this region might lead to executive dysfunctions including behavioral inhibition, regulation of affect and arousal, ability to analyze and synthesize, and memory" (Cataletto &

Serebrisky, 2019). Thus, it interferes with cognitive abilities and learning. A potential cause of damage in the PFC is sleeping disorder effects.

Sleep disorders classify into over 80 classes. The major classes are "insomnia, circadian rhythm disorders, sleep-disordered breathing or called sleep Apnea, hypersomnia/narcolepsy, parasomnias, and restless legs syndrome/periodic limb movement disorder" (Abad & Guilleminault, 2003).

First, sleep apnea is a potential serious sleep disorder. The event of apneas occurs when breathing repeatedly stops while the patient is asleep. The main types of sleep apnea are obstructive sleep apnea (OSA), central sleep apnea (CSA), and a combination of OSA and CSA known as mixed sleep apnea. The cause of OSA is when the upper airway gets partially or completely blocked during sleep, usually when the throat muscles relax. CSA happens when there is a failure of the respiratory drive when the brain fails to send the signals required for breathing (Mayo Clinic, n.d.). Studies found that OSA is associated with brain impairment because of profound changes in cerebral blood flow (CBF), in short, because of hypoxia.

Untreated sleep apnea may include serious brain damage, deterioration of or threat to life, diabetes, heart failure, and stroke (Biswal et al., 2018). For severe cases of sleep apnea, the treatment options could be using airway pressure devices, such as continuous positive airway pressure (CPAP) and adaptive servo-ventilation (ASV) machines. Other options include surgery, like removing tonsils and adenoids, jaw repositioning, and creating a new air passageway (tracheostomy) (Epstein et al., 2012).

The apnea-hypopnea index (AHI) is an indicator used to assess executive function in patients with OSA. To calculate the AHI, the experts follow the guidelines to count the number of apnea and hypopnea events occurring every 15 to 120 seconds, then divide the total by actual sleep time over 6 to 8 hours (Biswal et al., 2018). According to the American Academy of Sleep Medicine (AASM) (Borsini et al., 2018), the threshold of AHI is 5. If patients' AHI is less than 5, then this

is normal, otherwise, it is considered to be an abnormal case. Besides, it might have severed OSA (Biswal et al., 2018; Pinti et al., 2020).

Depending on AHI as the only measurement to diagnose OSA cultivates a risk of over-simplification, and it does not provide enough information for physicians to decide the treatment plan for their patients (Biswal et al., 2018). Since AHI is entailing many limitations and assumptions; thus, there is a crucial need for multiple parameters and further tools to assess the OSA.

Nonetheless, the massive amount of data collected from patients, plus the presence of advanced neuroimaging and computational technology, can be a powerful tool for diagnosing the OSA and any related cognitive impairments.

In terms of data, frequently, sleeping studies collects a variety of demographic information, clinical data, and clinical scales. One of the clinical scales is known as a Behavior Rating Inventory of Executive Function (BRIEF). BRIEF is a self-administered rating scale of executive function (EF) for children and adolescents, aged 5 to 18, completed by parents and teachers. It was first published in 2000. It is an evaluation tool for a daily behavior related to the functional connectivity (FC) in home or during school; it is used in clinical, psychoeducational, and research sets (Gioia et al., 2018). It reflects the neuropsychological construct of executive functionality. Therefore, it has been widely used in studies associated with neurocognitive impairment (Viola et al., 2017). BRIEF composed of 86 indices with 8 theoretically and empirically derived clinical scales.

On the other hand, a brain imaging technology known as functional near-infrared spectroscopy (fNIRS) offers researchers the flexibility to expand their human brain exploration, an exploration that can be translated into valuable applications (Ferrari & Quaresima, 2012; Irani et al., 2007).

This study presents novel research aiming to propose a prediction model that can be utilized as a diagnosing tool or an auxiliary tool coupled with additional resources. The goal is to predict children's behavior scores using their fNIRS data along with their demographic and clinical data.

Furthermore, this study can be classified as an exploratory analysis in the field of brain structures, activity, and cognitive functions. The potential model might be uses to identify patients with OSA using their fNIRS- behavior data as an additional measurement.

1.2 Thesis Motivation

Identifying a treatment plan for children with OSA is necessary; to avoid consequences on the brain tissue leads to behavior disorders or cognitive dysfunction. Often the judgement on the treatment plan is a subjective decision; usually it is taken based on one measurement, called the AHI. Using resting-state fNIRS time series data for a prediction method may support one's decision about whether to undergo a specific procedure such as Adenotonsillectomy (AT) surgery or not; the prediction method can avoid unnecessary risk and cost.

The main objective of this initial study is to build a model to predict the clinical behavioral BRIEF scores based on the resting-state fNIRS time series data, and further classify patients into normal or abnormal in terms of their behavior, taking into consideration their demographic information such as age and gender. Then we compare the outcome with the AHI results based on the given apnea events marked by an expert.

1.3 Problem Statement

Our goal in this work is to find a potential model that can address the question of how we can predict the behavioral scores for a subject using their fNIRS time series data. To explore this problem, we address possible open questions:

- How can we predict the behavior scores using resting-state fNIRS time series data?
- Is there a correlation between resting-state fNIRS time series data and a patient's behavior scores?
- Are the patients' resting-state fNIRS time series data different or similar according to the behavior scores? And the opposite?

- Can we rely on the resting-state fNIRS data as an indicator for patient classification? And how is the result accuracy compared to AHI?
- Does the demographic data have an impact on the behavior scores and the possible model performance?
- Which one of the three fNIRS parameters (HbO2, HHb, and tHb) has more information power?

To address the questions, we will review group of related studies uses fNIRS and fMRI time series data to predict a specific condition correlated with cognitive impairment.

1.4 Thesis Organization

This thesis is divided into four chapters as follows: Chapter 1 offers the introduction, background, thesis hypotheses, motivation, and goals; Chapter 2 reviews some related methods for analysis and prediction; Chapter 3, describes and analyzes the dataset, the experimental setup, and the thesis methodology; and in Chapter 4, summarizes our results, presents discussion and future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Functional Near-Infrared Spectroscopy (fNIRS)

The fNIRS is a neuroimaging technology coupled with microelectronics and software advancement. It increases the range of applications and a research institution's ability to monitor the human brain for better understanding and more investigations for daily actions (Di Domenico et al., 2019).

Fields of neurodevelopment, cognitive neuroscience, social neuroscience, and ecological science engaged the fNIRS technology for their analyses. Within these fields, Researchers completed studies of typical brain development such as language processing, functional within the visual, auditory, and sensorimotor systems. Furthermore, they studied atypical development, neurodevelopmental disorders such as autism (ASD), attention-deficit/hyperactivity disorder (ADHD), schizophrenia, depression, and Alzheimer's disease (Di Domenico et al., 2019; Fan et al., 2020; Hu et al., 2020; Hutchison et al., 2013; Sui et al., 2020; von Lühmann et al., 2021).

Some of the advantages of the fNIRS over other biomedical imaging methods such as functional magnetic resonance imaging (fMRI), computerized tomography (CT) scan, X-ray, ultrasound, and electroencephalography (EEG) are as follows: it is highly portable, lightweight, easy to wear, resistant to motion artifact, quiet, safe, and inexpensive. These attributes of fNIRS have made it particularly a great tool for studies with children and infants when they are awake or asleep. In general, the fNIRS provides an opportunity to study almost normal everyday life (Di Domenico et al., 2019; Emberson et al., 2017; Naseer & Hong, 2015; Pinti et al., 2020). For example, Figure 2.1. shows a type of fNIRS device used to measure infant's brain activity, it proves how convenient the device is.

Figure 2.1. A type of fNIRS machine, shows the appropriateness for infant study (Artinis medical SYSTEMS: Fnirs and NIRS devices-portalite mini, n.d.).

Furthermore, the portability of fNIRS makes it suitable for monitoring cortical hemodynamics response during motor tasks or resting. This is in contrast to MRI technology restriction where the subjects have to be in a lab where magnetic fields are needed. However, since the fNIRS is quite recent technology, it lacks the analogy to the MRI sophisticated methods for conducting multivariate pattern analyses (MVPA).

In details, fNIRS is based on the neurovascular coupling, it provides images of blood concentration and oxygen saturation in tissue (Naseer & Hong, 2015). It can reveal the hemodynamic response signal, analogous to the blood oxygen level dependent (BOLD) response perceived by fMRI. The fNIRS signal can measure the changes of the oxy-hemoglobin (HbO2) and deoxyhemoglobin (HHb) concentrations in the brain tissue by accurately measuring the amount of light that is transmitted through the brain tissue. The fNIRS machine has multiple near-infrared light sensors that are sources and detectors inserted into a cap to capture the bulk of photon movement, in order to identify the cortical activation. The source sends light in a banana-like shape through the skin and the skull until it reaches the outlier of the brain for about 3 cm in depth; often the depth depends on the types of fNIRS instrumentation. Then the unabsorbed light travels back to the surface of

the head where the detector measures it (Naseer & Hong, 2015). Therefore, large datasets are created from these targeted regions. Figure 2.2. shows the path of the light between the source (the emitter) and the detector through the surface of the brain.

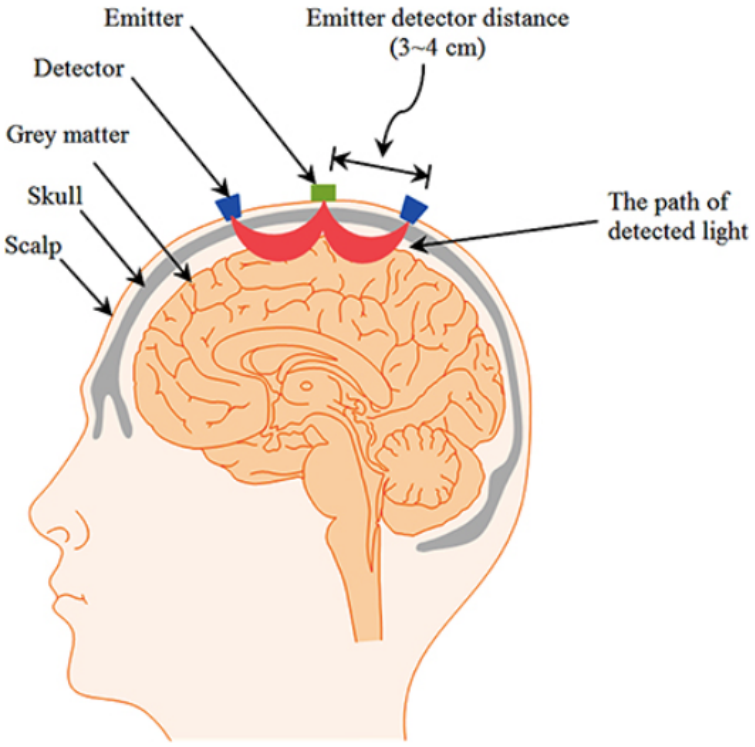


Figure 2.2. fNIRS-based brain-computer interfaces, shows the banana shape photon path (Naseer & Hong, 2015).

The brain neurons communicated with each other and transmitted massages to and from body tissues. The neurons produce a neuroanatomist's chemical to relay their signals, therefore the production process needs the energy that comes from glucose obtained from food and the oxygen carried by the Hemoglobin in the blood. Hence, the active areas in the brain need more energy, so the blood flow with oxygenation hemoglobin (HbO2), as well as the action known as regional CBF, increases in these regions; this process is called functional hyperemia. Figure 2.3. outlines the sequence of the brain hemodynamic. Therefore, there is much light absorbed in these regions.

Based on the difference between the amount of light sent by the source and received by the detector, the machine can calculate the amount of oxygen being consumed (Ayaz et al., 2019; Pinti et al., 2020).

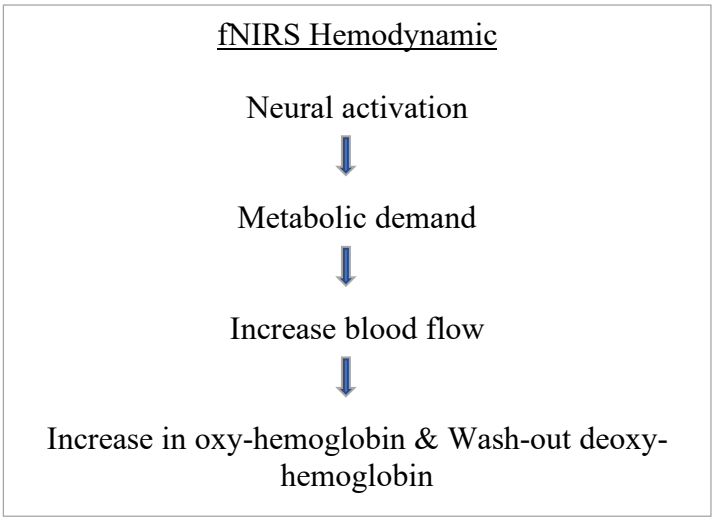


Figure 2.3. shows a simplified sequence of the mechanism of the brain activity (FNIRS & nirx: Fnirs systems: NIRS DEVICES: NIRX, n.d.).

Figure 2.4. from Pinti et al., (2020) illustrates the interface between the sensors and a computer base, showing the hemodynamic responses during a designed memory task. Part (A) is a recording of concentration changes between HbO2 and HbR (they also noted as O2Hb and HHb, respectively) from one channel over the PFC during the stimulation period shown by the gray areas. Part (B) presents the hemodynamic response, computed by averaging the HbO2 and HHb signals within a gray block. Part (C) shows the 16 channels over the PFC, where circled in magenta is the source of the signal interchanges presented in part (A) and part (B).

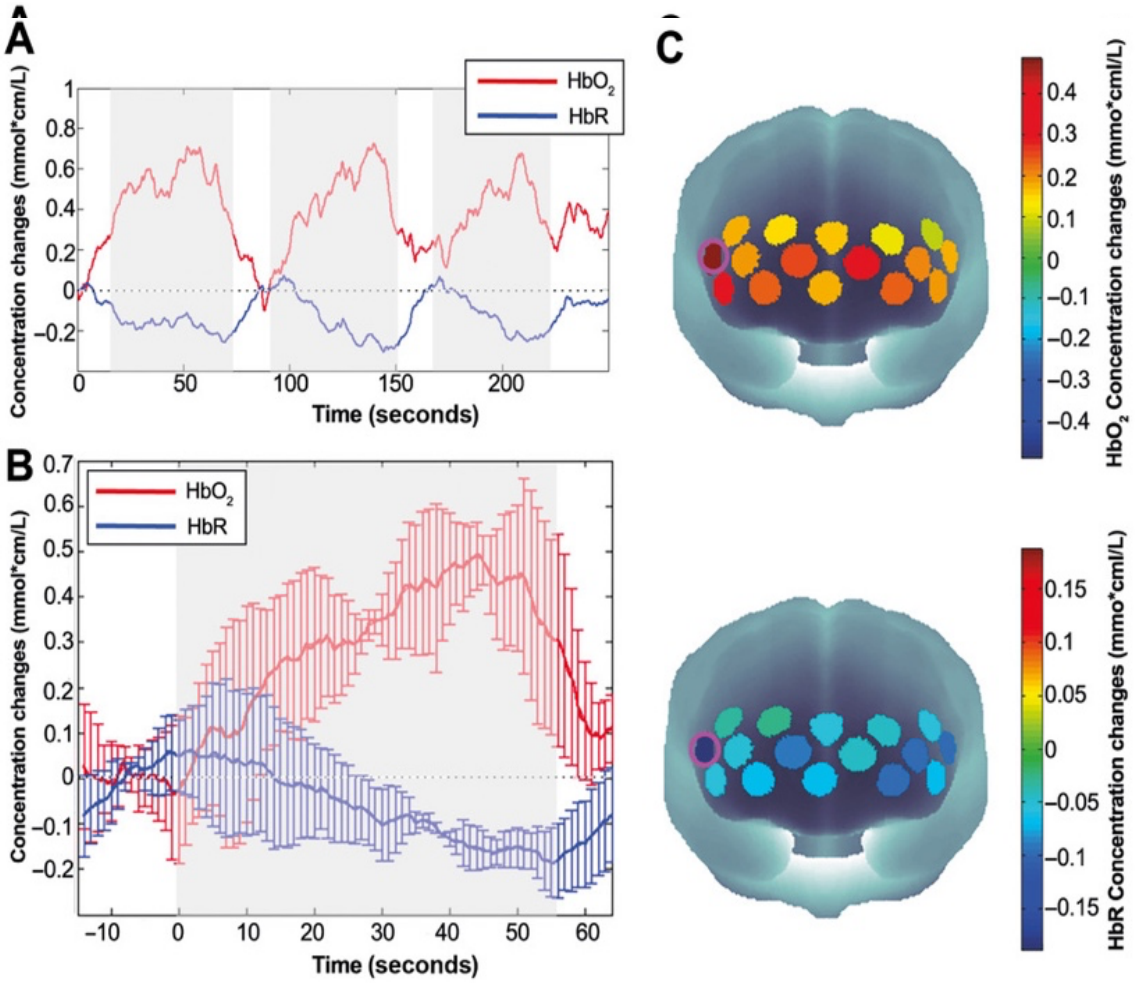


Figure 2.4. "Example of HbO2 (red) and HHb (blue) signals from one channel (circled in magenta in panel C) of a single subject over the PFC using the fNIRS, during memory task. The gray areas refer to the stimulation period. The panels show the block-averaged hemodynamic response, computed by averaging the HbO2 and HHb signals" (Pinti et al., 2020).

2.2 Assessing Dynamic Brain Connectivity

The statistical relationship between particular physiological or brain signals in space and time is known as functional connectivity (FC). For years, the fMRI, electroencephalography (EEG), positron emission tomography (PET), diffusion tensor imaging (DTI), and electroencephalography

(MEG) machines have been used to assess this relationship between brain regions. The FC shows a significant impact of detecting a variety of neurological diseases, cognitive distortions, and effective in the social and behavior research (Hutchison et al., 2013; Hu et al., 2020; Shen et al., 2017; Zhang et al., 2010)

Moreover, FC in the brain is analyzed by multiple demonstrated methods. The methods are derived from computational and statistical backgrounds; examples are (Rogers et al., 2007):

- Seed region correlation maps analysis.
- Inverse covariance methods.
- Multivariate decomposition methods:
  - Principal component analysis (PCA).
  - independent component analysis (ICA).
  - Partial Least Squares.
- Graph theory.

Researchers mostly track the brain activity during specific tasks and collect evidence points to non-stationarity FC (Leonardi et al., 2013). However, recent studies have found that tracking the brain activity during sleep or resting-state offers new perspectives and great opportunities to learn more about the brain's functional structure, since during resting-state brain signals are more robust, particularly in the PFC region. Resting-state dynamic functional connectivity (dFC) was discovered by B. Biswal el., 1995 (Chen et al., 2020).

In summary, the importance of understanding the nature of the data created by fNIRS and the FC is to be able to correctly analyze the data with the right tool to take advantages of modeling the brain signals. In Chapter 3, we provide more details of multivariate decomposition methods used for FC analysis.

CHAPTER 3

CONCEPTS AND DESIGN

This chapter highlights some related studies on resting-state fNIRS and resting-state fMRI time series data for various applications and analysis. We briefly explain the statistical and computational methods that have appeared frequently in those studies.

3.1 Multivariate Pattern Analysis

Multiple methods applied to analysis and study the brain signals recorded by neuroimaging technology, and link between the signals and behaviors or cognitive measurements. In this section we present some of the common methods from statistical and computational backgrounds. Reviewing the methods help to evaluate the appropriate approaches for our problem.

The approaches to study behaviors are the behavior-first approach and the brain-first approach. The behavior-first monitors participants as they complete a prospective memory task or behavior then link it to the occurrence of brain activity. The second approach, the brain-first, does the opposite. Both cases require some sort of FC analysis (Pinti et al., 2020).

The objective of calculating the function connectivity is finding a deep value of the normally distributed time series data from two or more separate brain regions for either a subject or group of subjects. It can be attained in many ways. The most obvious statistical method used to analyze the fNIRS or fMRI data is averaging. The averaging approach is a static metric; it does not take advantage of the temporal dynamics in brain signals (Pinti et al., 2019).

A more powerful methodology has been developed by researchers and includes transforming time series data into a correlation coefficient matrix first, then performing the matrices group analysis, or what is known as multivariate pattern analysis or multivoxel pattern analysis (MVPA). MVPA is a statistical method. It is a collection of tools used to study the joint behavior of multiple random variables, understand the complex spatial disease affecting the brain, and most importantly, extract

information from the pattern of brain activity. Examples of multivariate techniques are correlation analysis, principal components analysis, factor analysis, and cluster analysis (Linn et al., 2016; Emberson et al., 2017).

3.1.1 Time-Series and Structures Data

Our work requires combining the fNIRS data and the expert labeled Apnea events, along with the targeted labels and the demographic information for the prediction task. One approach of combining time-series and structure data for classification is by either merging or aggregating the structured data into the time-series data (Du, 2019; Chablani, 2020).

Specifically, in the case of aggregating the time series data into the structured data, three methods can be used: (1) Window-based summary: for each window prediction generate new features for the next time points. (2) Time-series encoding: treat the data as multimodal deep learning using methods such as recurrent neural network (RNN), convolution neural network (CNN) or seq2seq. (3) Time-series heuristics: the most known methods for this case are ARIMA, TimeNet, and CNN-Long-term short-term memory (LSTM), and the most naïve model is the moving average. Moreover, in the case of merging, the structured data can be added to the time series data as new features.

3.1.2 Representational Similarity Analysis (RSA)

Representational Similarity Analysis (RSA) is a multivariate technique; it is derived from the same method as classification-based MVPA. The similarity-based decoding attempts to classify new observations based on the similarity with other known observations for conditions such as ADHD/autism. It combines data from different sources by using a common representational space, hereby denoted as the geometric representation of information. Then it measures the similarity between the brain activity in the regions of interest (ROIs) and measurements such as behavioral data in a higher order representation. A modest version of an RSA approach is simply using the Pearson correlation (Emberson et al., 2017; Kriegeskorte, 2008).

3.1.3 Pearson Correlation

The common term used to describe the relationship between two variables is either covariance or correlation. The Pearson Correlation is a statistical method; it measures the strength and the direction of linear correlation between two variables. Usually, the notation r is used for the Pearson Correlation result; the value of r is between -1 and 1. The coefficient r -value represents the strength of the correlation, while the sign is for the direction of the correlation, either positive or negative, where zero value means no correlation is found. In a scientific context, describing the strength of the correlation depends, and it has to be a specific term between weak, moderate, strong, and so on, accompanied by a confidence interval (Mukaka, 2012; Akoglu, 2018).
The Pearson Correlation formula is as follows:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Where *r* is the correlation coefficients, x_i is the values of the x-variable , y_i is the values of the y-variable, $\bar{x}$ and $\bar{y}$ is the mean values of the x-variable and y-variable respectively.

Other common values of the Pearson Correlation are R-squared and P-value. The R-squared shows the amount of variation between the variables shared, and the P-value with a confidence level is used to test the Null hypothesis; the lower the P-value is, the greater the statistical significance and the Null hypothesis rejected in favor of the alterative hypothesis. The P-value shows the probability that this strength may occur by chance.

In order to use the Pearson Correlation, there are some underlying assumptions taken. The assumptions are that the variables are independent, continuous, normally distributed, no outliers

exist, and linear association exist. The correlation between two variables is called bivariant, and multivariant for more than two variables.

The Pearson Correlation is used to construct the brain connectivity matrix for both dynamic and static brain FC (Mukaka, 2012).

3.1.4 Linear and Polynomial Regression

Linear regression is a simple and powerful method for modelling and estimating the relationship between the input variable X and output variable Y. The output variable is also called the dependent variable, while the input variable X is also called the independent variable or the predictor variable. When there is only one input variable, the model is simply a linear regression, whereas if there are many input variables, the model is a multiple linear regression. The regressor is, in this case, a first order polynomial.
The simplest linear regression model consists of the mean function coupled with the variance function Var as follows:

$$E(Y|X = x) = \theta_0 + \theta_1 x\,,$$

$$Var\,(Y|X = x) = \sigma^2$$

where, $X = X$ is the predictor, $h\theta(x) = \theta_0 + \theta_1\, x$ is the regression function, θ_0 and θ_1 are an unknown parameter named the intercept (bias) and the slope respectively. $Y = \theta_0 + \theta_1 x + \varepsilon$ is the general linear regression model or general multivariate regression, for short the GLM, where $Y = \theta_0 + \theta_1 x$ is the model and ε is the error term.

Usually, we need to estimate the model parameters in order to come up with an optimal model or fitted line $h^{\theta}(x) = \theta_0 + \theta_1 x$. In the case of several independent variables $x1, x2, \dots, x_n$ the model is defined as follows:

$$Y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n + \varepsilon.$$

Hence the regression function is given by: $h\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n = \theta^T x$.

A more advanced version of regression methods is the partial least squares regression (PLS). It is an effective covariance-based regression technique. It finds a linear regression model by projecting the predicted and the observable variables to a new space. It is recommended when there are a high number of explanatory variables, that are likely to be correlated (Meidenbauer et al. , 2021).

3.1.5 Support Vector Machine (SVM)

The support vector machine (SVM) gained popularity among the machine learning and neuroimaging community (Linn et al., 2016); Due to an easy accessibility of implementation tools, SVM has been widely employed for behavior predictions (Sui et al., 2020; Emberson et al., 2017). SVM **c**an produce significant accuracy for a wide range of problems with less computational power. It is used for classification and regression tasks; the regression method is SVR. The SVM models learn from a subset of data points called support vectors; these points are hard to classify. They are places on the opposite sides of a linear or non-linear decision boundary. In addition, the algorithm can control the complexity of the model, mapping the higher order dimension to a lower one using the kernel function (Sui et al., 2020). Moreover, SVM does not suffer the overfitting problem when the model fits the training data but not the testing data, which ensures a good generalization performance.

The decision boundary is known as the margin of separating hyperplane. The universal equation of the hyperplane is:

$$w^T x + b = 0,$$

where x is the data point and (w, b) represents the parameter of the hyperplane. Figure 3.1. illustrates an example of a simple explanation of the linear SVM classification, using soft margins where the red and blue circles are based on two predictors.

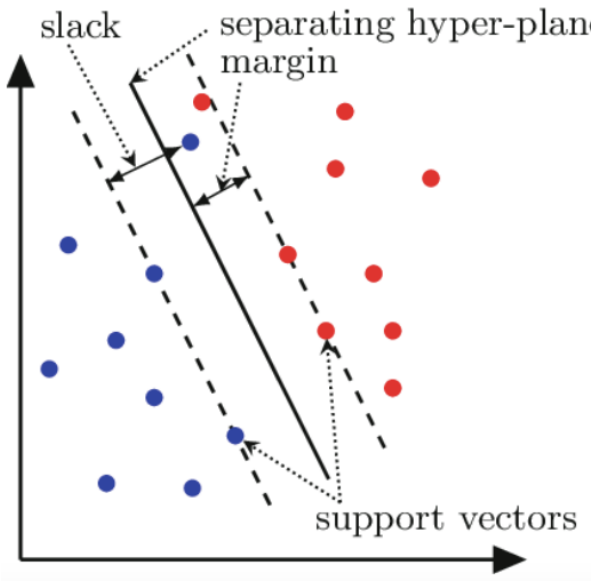


Figure 3.1. shows the separating (classification) hyperplane, margin, support vectors, and slack for a two-dimensional feature space (Caputo et al., 2015).

3.1.6 Principal component analysis (PCA)

The principal component analysis (PCA) algorithm is an exploratory tool for data analysis. It is used for dimensionality reduction purpose, mapping the data from high dimensions to low dimensions. As a result, it increases the interpretability of the dataset while minimizing information loss and keeping as much variation as possible. The algorithm creates new uncorrelated variables that maximize the variance (Jolliffe & Cadima, 2016; Leonardi et al., 2013).

The pseudocode of the principal component's algorithm shows the steps for data set X and number of components required denote k,

Procedure for Principal Component Analysis (PCA): (X, k)

1: Standardize the dataset.

2: Calculate the covariance matrix for the features in the dataset.

3: Calculate the eigenvalues and eigenvectors for the covariance matrix.

4: Sort eigenvalues and their corresponding eigenvectors.

5: Select k eigenvectors, called principal components, that have the k largest eigenvalues.

6: Pick k eigenvalues and form a matrix of eigenvectors (these are the top k PC) (the function return).

7: Transform the original matrix (Only in case we interested on the new projection).

8: End procedure

Algorithm 1: Principal Component Analysis (PCA)

3.1.7 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a computational method. It is a generalization of PCA; it considers the high-order dependencies among variables. Again, it decomposes a complex dataset or multivariate signal into simpler components. For instance, it takes the spatial-temporal time series data forms multiple voxel/cortices and breaks it down into spatial and temporal components (McKeown, 2003). Moreover, ICA is capable of separate noise and artifacts from resting-state fNIRS data (Zhang et al., 2010).

3.2 Sliding Window Analysis

The sliding window technique is considered one of the main tools to explore time series data. It is a commonly used method to examine dFC dealing with the high dimensional time series data (Hutchison et al., 2013).

Based on the defined time window length, the window the data points to calculates a specific statistical measure, such as the correlation coefficient. It gives a perception on the changing patterns within the data points through time.

The limitations of the sliding window technique are that the length of the window affects the result analysis, and it lacks sensitivity to transient changes in network connectivity (Hutchison et al., 2013). To overcome these issues, the sliding window can be implemented with time–frequency analysis (Mokhtari et al., 2019).

3.3 Cross Validation

Cross validation is a model validation technique. It is also a re-sampling technique used for model selection, and most importantly, evaluates model generalizability (Shen et al., 2017). Prediction methods may use cross validation to select the most appropriate model after evaluating the latter on the validation set in order to limit overfitting issues, for instance. In our use of the method, to find a balance between the bias and variance, we need to test the hypothesis on the validation set separately from the training set. It would obviously be a mistake to evaluate the model on the same dataset used to train the model as the model would have a perfect score (but would then fail to generalize with unseen data). It is therefore important to test the model on unseen samples called validation or evaluation sets and then measure the errors for train and test sets. The earliest form of estimation and inference in linear regression was the method of least squares. The purpose is to minimize a quantity called Residuals Sum of Square (RSS) or cost function. The cost function is the square root of the difference between the actual (y_i) and predicted values ($h_\theta(x_i)$). It can also be used as one of the model selection measurements. The mathematical form of RSS for a sample of size n is given by:

$$J_{train}(\theta) = \frac{1}{2n}\sum_{i=1}^{n}(h_\theta(x_i) - y_i)^2 ,$$

$$J_{cv}(\theta) = \frac{1}{2n}\sum_{i=1}^{n}\left(h_\theta\left(x_{i_{cv}}\right) - y_{i_{cv}}\right)^2 ,$$

$$J_{test}(\theta) = \frac{1}{2n}\sum_{i=1}^{n}\left(h_\theta\left(x_{i_{test}}\right) - y_{i_{test}}\right)^2$$

The model complexity increases with the number of samples or the number of parameters to estimate. Adding more training samples reduces the estimation error though increasing the bias. In contrast, increasing the number of unseen samples for the prediction function on the validation set increases the estimation error. In an underfitting case we have $J_{train} \approx J_{cv}$, meaning that the model has high bias (low variance). On the other hand, in the case of overfitting, $J_{train} \ll J_{cv}$ meaning that the model has high variance (low bias).

The validation set is meant to tune the model parameters and the testing set is for evaluating the model accuracy. Another important point of cross validation is to make sure that the sampling is fair and crosses over (Refaeilzadeh et al., 2018). The cross validation means that every sample from the original dataset has the same chance of appearing in the training and validation set.

Different cross validation methods can be used to split datasets. One of the themes is random selection, in which a fixed number of samples from the train set are randomly selected without replacement. The final estimation true error is the average for k estimators:

$$E = \frac{1}{k}\sum_{i=1}^{k} E_i$$

The K-fold method is a popular cross validation method. With this approach, the samples (x_i, y_i) are divided into K groups called folds. The K-1 fold is then used for learning while the other folds are meant for testing. The 10-fold cross validation is the most commonly used in machine learning by the data mining community.

A special case of the K-fold method is the shuffle split, where K=N, N being the dataset size. It generates N group of samples; N-1 are used for training and one for testing. It is known as the leave-one-out cross-validation. If the data ordering is not arbitrary, shuffling may be essential to get a meaningful cross-validation result.

3.4 Cognitive Impairment prediction using fNIRS/fMRI

This section summarizes a selected group of compound methods in processing fNIRS data and using it for either a classification of a cognitive impairment or a regression of measurements.

3.4.1 Detection of Cognitive Impairment Using Recurrent Neural Network (RNN)

Zhu et al.'s (2019) work used the resting-state fNIRS to detect cannabis-associated cognitive impairment. They examined two machine learning algorithms: the SVM and the RNN model. Both of the models trained on the dynamic FC matrices. Zhu et al. coupled their model with data augmentation strategy apposite for the data to boost the predicative model performance. In their work, 20 channels over the PFC were recorded during resting-state.

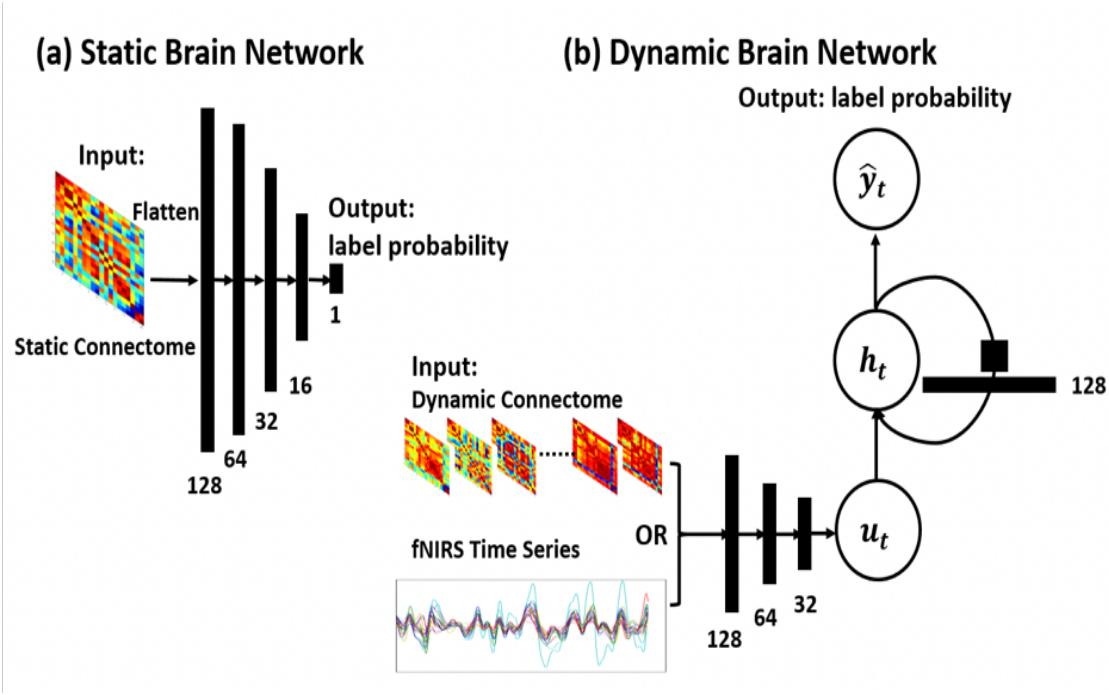


Figure 3.2. illustration of classification model using (a) Fully connected neural network on a static connectivity matrix. (b) RNN on a dynamic connectivity matrices or input fNIRS timeseries (Zhu et al., 2019).

The architecture of the RNN prediction model (Zhu et al., 2019) applied is shown in Figure 3.2. Here, part (b) shows a fully connected network with four hidden layers, containing 128, 64, 32, and 16 nodes, respectively. They implemented a cross-entropy loss function and a *tanh* nonlinearities activation function to compute the probability of cognitive impairment status as output. They feed the model two types of inputs for a comparative analysis, one with raw signals for the 20 channels and the second with the FC matrix.

3.4.2 Detection of Cognitive Impairment Using convolutional neural network (CNN)

fan et al. (2020), employed a model using deep learning algorithms. They combined the CNN and LSTM network. The model is proposed to capture the spatial temporal structures of the dynamic FC based on resting-state fMRI data. The application focused on predicting individual's cognition performance and gender classification. They took advantage of the LSTM model architecture rather than the RNN, where LSTM has a long-term and short-term memory, which improves the learning progression. Moreover, RNN has sufficient ability to solve long-term dependencies. Fan et al. (2020), used LSTM to detect temporal dynamics and the CNN designed for spatial feature extraction.

Additional research takes the advantage of the deep learning algorithms computational power for classification of ASD using resting-state fNIRS and resting-state fMRI time-series data done by Dvornek et al., 2017; and Thomas et al., 2020.

3.4.3 Connectome-Based Predictive Modeling (CPM)

Connectome-Based Predictive Modeling (CPM) is a simple linear regression model. CPM combines simple linear regression and feature selection together to predict individual differences in traits and behavior from connectivity data. It was proposed by Hutchison et al. (2013) for evaluating the brain-behavior relationship. CPM is trained on fMRI connectivity data using cross-

validation and generates predictions of behavioral measures. The main characteristic of CPM is that it is a regression model that can predict continuous variables, while the other methods we reviewed are classification models. It is worth noting that working with a regression-based problem can be more challenging. However, CPM is relatively easy to implement and understand for the neuroimaging community, where it considered as more straight forward and simple method (Sui et al., 2020).

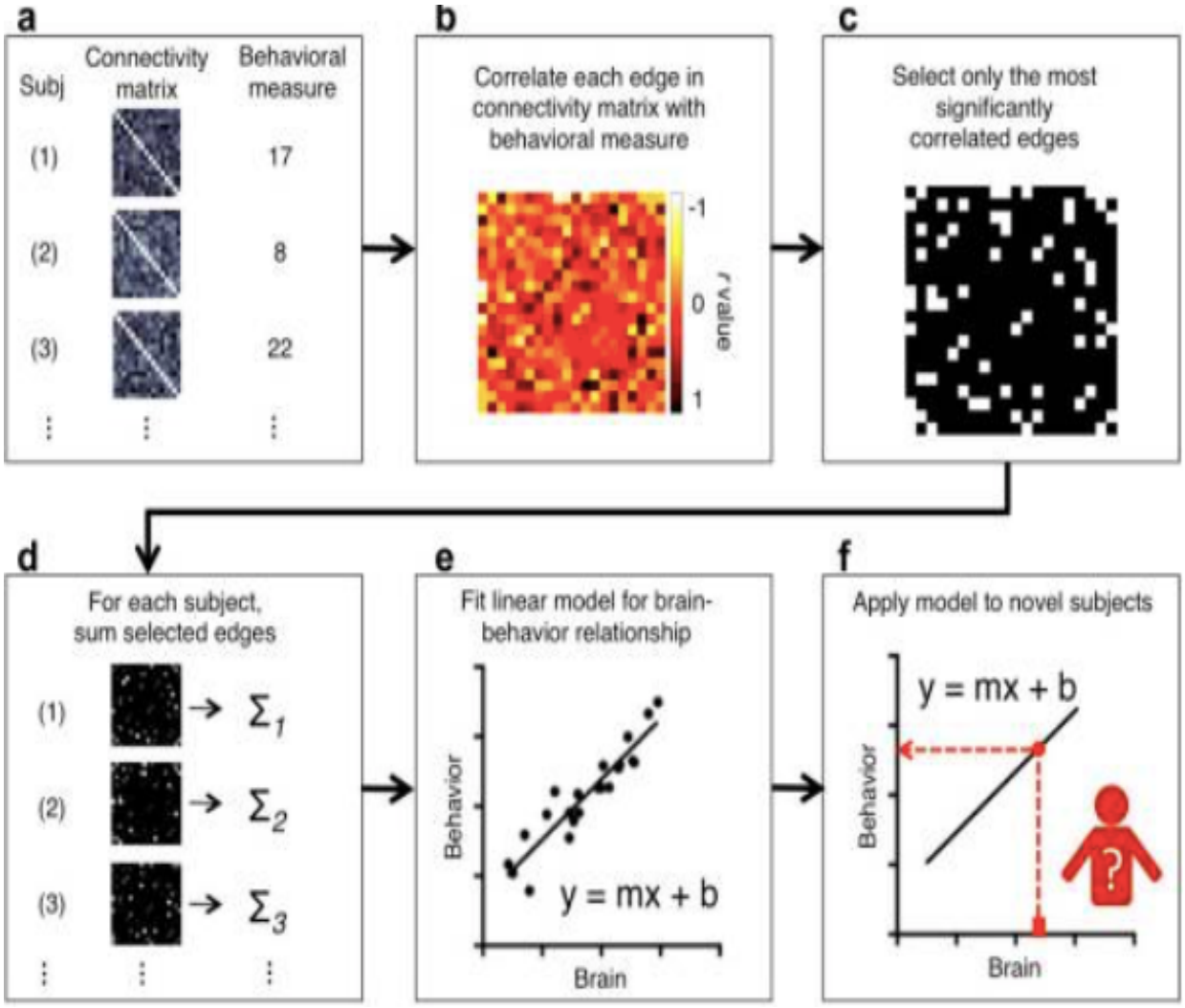


Figure 3.3. shows the schematic of CPM (Shen et al., 2017).

Figure 3.3. presents the schematic of the CPM model. It starts in part (a) by forming the connectivity matrix and behavioral measures for each subject and dividing the data into a training set and a testing set. Second, part (b) forms a connectivity matrix for all subjects where each edge in the connectivity matrices is associated to the behavioral measures using a form of linear regression. Third, in part (c), the most principal selected edges are selected based on a pre-defined threshold. Part (d) summarizes the selected edges into a single subject value. Part (e) builds a predictive model, assuming a linear relationship between the single-subject summary value of connectivity and the behavioral variables. Lastly, part (f) calculates the summary values for each subject in the testing set and uses it as an input for the model to predict the behavioral measures.

Sui et al. (2020), summarized regression-based approaches for single-task and multi-task problems. The single-task predicts a single behavioral measurement or cognitive score, while the multi-task is designed to predict multiple behavioral measures assuming the correlation between these variables. They mentioned the success of the CPM model for the prediction of multiple human behaviors. Their work is conducted on a new large data set known as Adolescent Brain Cognitive Development (ABCD). Figure 3.4. lists these regression approaches.

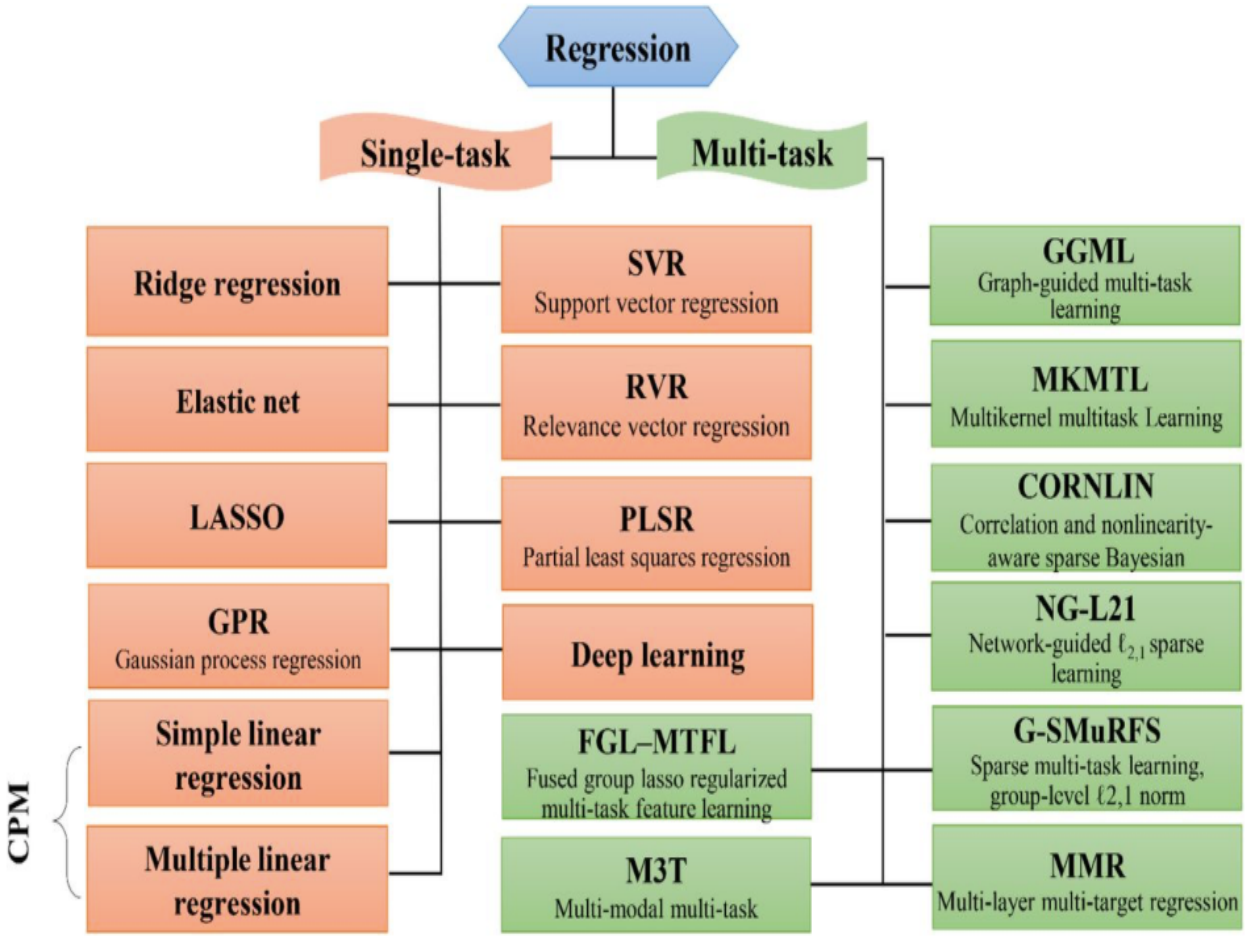


Figure 3.4. summarize the regression approaches by (Sui et al., 2020).

CHAPTER 4

METHODOLOGY

4.1 Methodology

The general sequence to build a prediction model starts by understanding the problem and describing the dataset by using visualization tools as an example. Then, the researcher selects the features and prepares the dataset to fit the proposed predictive models. After that, the researcher applies the model evaluation and select the most promising model. Finally, the researcher optimizes and measures the final model performance before presenting the last result.

For the rest of this chapter, we will follow part of the general framework above since this is a novel problem and it has two main challenges. The first is combining the multimodal data for the analysis, and the second tis he original problem, being a multivariate regression problem, that requires a large quantity of data. However, our dataset is limited at this stage.

4.2 Data Description

The data comes from the University of Maryland School of Medicine. The data is for 50 patients, of whom 33 were included for analysis. The patients' ages are between 4 to 12 years old, and their data was recorded in a sleeping lab. The data was separated into two files. The first file consists of 274,370 records for all patients. It contains the fNIRS time series data for three channels, measuring the prefrontal cortex (PFC) blood flow, and recording neuronal activation during resting-state for between 5 to 7 hours. The fNIRS captures the concentration changes of oxygenated hemoglobin (HbO2) and deoxygenated hemoglobin (HHb) in the of brain tissue. The measurements were recorded every five seconds. The fNIRS data merged with other discrete events called apnea, which is identified by a certified sleep technician. The second file contains 13 clinical behavioral scores, some statistical (demographics) data, apnea hypopnea index (AHI),

and other measurements. The mean age of the 33 children was 7.24 years, and the gender distribution among them was 18 to 15 males to females.

That biometrics for sleep apnea are: OBS apnea, central apnea, mixed apnea, hypopnea, desaturation, arousal, spontaneous arousal, apnea arousal, arousal hypopnea, and arousal desaturation. Each of these events recorded as a discrete value, either 0 or 1, when it occurred during the sleep time.

From BRIEF, nine clinical scales as well as four composite scales are included for the prediction task: Inhibit, Emotional Control, Behavioral Regulation Index (BRI), Shift, Emotional Control, Emotion Regulation Index (ERI), Initiate, Working Memory, Plan/Organize, Task Monitoring, Organization of Materials, Cognitive Regulation Index (CRI), and Global Executive Composite (GEC) (Jacobson et al, 2016).

4.3 Data Preparation

The fNIRS data for some records needed cleaning from spike artifacts; most of the records of HbO2 and HHb fell into range between -25 and 25. To correct the fNIRS recorded signals, we used a median filter, with a sliding window size of 100. The records for patients 1, 2, 44, and 50 were missing some values, therefore, we replaced the missing values by the average for the continuous variables and the mode for the discrete variables.

Then to prepare the data for the classification algorithm, we classified patients into normal and abnormal groups. First, we utilized the Inhibit clinical score (inhibit_pct), the control impulsive, to distinguish the cases with threshold equal to 65. Patients with inhibit_pct score above 65 are considered abnormal. In this case, the number of patients with a score above or equal to 65 were 20 out of 33. Then we fed the data to the tensor connectome model to be shaped into 4D Tensor.

Next, we calculated the average score based on the 13 behavior scores of BRIEF to classify the patients. The number of abnormal patients was 21 and normal patients was 12. To compare the

model results we also classified the patients using AHI. As a result, the number of abnormal patients was 19 and normal patients was 14.

In our effort to study the impact of the fNIRS data based on the location of the three sensors [Tx1, Tx2, Tx3] for each of the HbO2, HHb or tHb, we created separate data files for each measurement to check how the model behaved using different time series, along with the classified patients based on the inhibit_pct and the average score and compared the results with the AHI.

The preprocessed data file was fed to the tensor statistical model to learn the statistics of connectome patterns. The training data was shaped onto a 4D tensor, where the four dimensions were the stacked FC matrix within each sliding window, the time dimension, and all the training subjects. The 4D tensor present the independent variable and the class target is our dependent variable we want to predict (Zhu et al., 2017).

4.4 Model Implementation

The model we deployed in this work is called Tensor Connectome, developed by Zhu et al. (2017). The tensor statistical model is implemented for quantifying dFC in the brain. It has been applied to classify patients diagnosed with Autism Spectrum Disorder (ASD) based on the learned dFC patterns. For the rest of this section, we summarize the model for the implantation intent.

The high dimensionality of the dFC makes it challenging to quantify. Common work used the PCA or clustering coefficients to reduce the dimensions, but the approaches mechanism dose not preserve the temporal dynamics and changes along time. However, the learning-based tensor model is originally designed to drive connectome biomarkers from resting-state fMRI images for each individual. Nonetheless, this model works well for data from multiple regions of the brain, easing the issues of high dimensionality data.

In general, the dFC for each individual subject is a matrix calculated by the Pearson's Correlation between the ROIs to measure the strength of FC. Then, this learning-based approach is integrated

with the sliding-window technique to expose the dynamics of FC. The dFC matrix for an individual is stacked within each sliding window to form the 3D tensor. Next, the 4D tensor is formed based on the 3D, along with all subjects. It is important to emphasize that this model combines the practices of evaluating FC for individuals and extracting connectome features for the population.

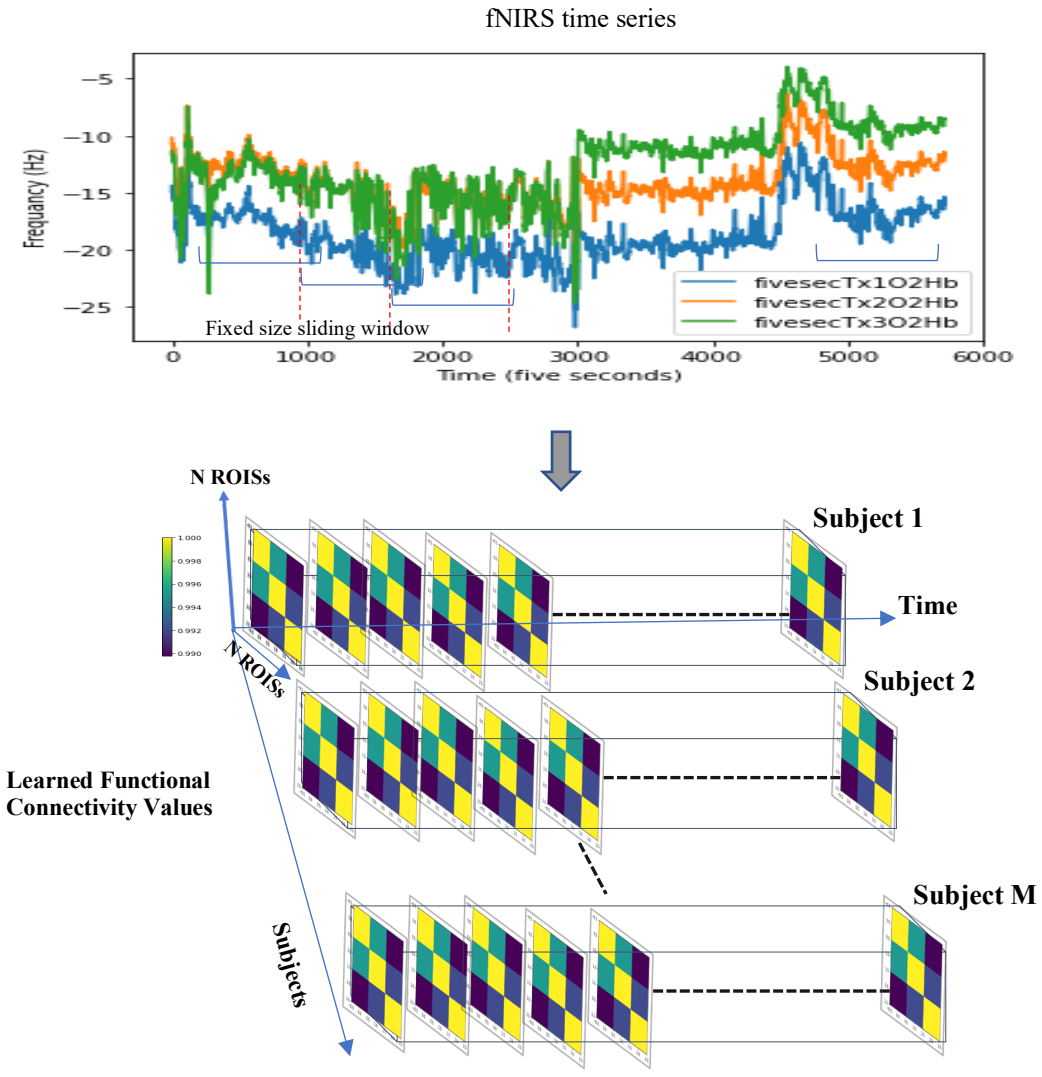


Figure 4.1. is illustration of the learning-based 4D tensor model.

In Figure 4.1. presents the tensor model where the tensor connectome model is trained on a four dimensions tensor constructed from each individual's connectivity matrix.

The tensor connectome model is a module-wise network structure; it is enhanced modeling the FC by reflecting on the region-to-region signal correlations and the similarity between high level module-to-module connection patterns, contrasting to the convectional models. It suggested a dynamic sliding windows technique. Furthermore, based on the learned pattern and the intrinsic feature, the model optimizes the dFC level on individuals and the population.

The main steps for running the tensor statistical model are: (1) Construct the 3D tensor of Pearson's correlation for each individual subject, (2) Implement the proposed regularization model to develop the dFC, (3) Construct the 4D tensor based on the learned model features, (4) Implement the regularization model including all subjects, and (5) Decompose the 4D tensor into the 3D tensor to conduct the temporal dFC nature and obtain the compact connectome feature representation. Following that, we will go through each phase in depth, reflecting our own resting-state fNIRS time series data.

Let $x1$, $x2$, and $x3$ be our three channels from fNIRS, for each of the HbO2, HHb, and tHb channels. Each of $xi \in \mathbb{R}^{W\times 1}$ where W is the length of time points within the sliding window; in our solution the initial sliding window size is 40, later overlapped multiple length of sliding windows is adopted for the whole-time course for each subject. Doing that would make uniform the length of fNIRS data for all subjects. Based on the values of $x1, x2$ and $x3$ the Pearson's Correlation matrix size is $NxNx40$ where $N = 3$ in our case, and $3x3$. The result is what is called the connectivity matrix denoted by tensor $\mathcal{S} \in \mathbb{R}^{3x3xT}$. Since the data channels we have are within the same region, they are strongly correlated and there is no need for a threshold to remove the weak connections or the spurious connection, therefore, our matrix would not be sparse.

The tensor statistical model applies the FC optimization method to avoid quantifying the dFC based on signal correlation only. The optimization equation for the 3D tensor model is as follows:

$$\arg\min_{\mathcal{S}} \|\mathcal{S} - \mathcal{C}\|_F^2 + \alpha\|\mathcal{S}\|_* + \gamma\|\mathcal{S}\|_1 \quad (1)$$

Where, $\mathcal{S} \in \mathbb{R}^{3x3xT}$ is the stacked FC matrix within the sliding window length t, for a single subject. $\mathcal{C} \in \mathbb{R}^{NxNx\mathrm{T}}$ presents all the arranged Pearson's correlation values, and the information from the module-to-module connection to guide the measurement of low-level region-to-region connectivity strength and represent sparsity. The scaler α and γ defined to balance the strength of the low rank constraint with the $L1$ norm, since the $L1$ norm is effective to remove the redundant connections in the FC matrix. Moreover, low rank constraint $\| \|_*$ was applied on $\mathcal{S}$ to control the changes in the temporal domain. The equation outcome is denoted as $\mathcal{S}^*$.

As extension for the above framework into the 4D tensor for the whole populations, the equation becomes:

$$\arg\min_{\mathbb{S}} \|\mathbb{S} - \mathbb{C}\|_F^2 + \alpha\|\mathbb{S}\|_* + \gamma\|\mathbb{S}\|_1 \tag{2}$$

Where $\mathbb{S} \in \mathbb{R}^{\mathrm{N\times N\times T\times M}}$ and $\mathbb{C} \in \mathbb{R}^{\mathrm{N\times N\times T\times M}}$, M is the number of samples, T is the number of sliding windows.

$$\arg\min_{\mathbf{U}_{(c)},\mathbf{U}_{(t)},\mathcal{F}^{\#},\mathbb{S}} \|\mathbb{S} - \mathbb{C}\|_F^2 + \alpha\|\mathbb{S}\|_* + \gamma\|\mathbb{S}\|_1$$
$$\mathcal{S}^{\#} = \mathcal{F}^{\#} \times_1 \mathbf{U}_{(c)} \times_2 \mathbf{U}_{(t)}, \tag{3}$$

Then $\mathcal{S}^{\#} \in \mathbb{R}^{\mathrm{N2\times T\times M}}$ is the transformed matrix form the 4D tensor to the 3D tensor in order to learn the principal connectome space. The result of the decomposing of the tensor $\mathcal{S}^{\#}$ is:

$\mathrm{U}_{(\mathbf{c})} \in \mathbb{R}^{\mathrm{R\times N^2}}$ is the brain low dimensional connectome space, $\mathbf{U}_{(t)} \in \mathbb{R}^{R\times T}$ is the low dimensional temporal dynamic space, $\mathcal{F}^{\#} \in \mathbb{R}^{R\times R\times M}$ is a coefficients tensor, $\times$ 1 and $\times$ 2 represents the tensor mode multiplication. Finally, the classifier equation developed as:

$$\mathbf{F}^* = \mathbf{U}_{(\mathbf{c})}^{\mathrm{T}} \; \mathbf{S}^{*\#} \; \mathbf{U}_{(\mathbf{t})} \tag{4}$$

Where $\mathbf{F}^* \in \mathbb{R}^{R \times R}$ is a compact connectome feature representation as a result of the projection of the learned connectivity spaces $\mathbf{U}_{(c)}$ and $\mathbf{U}_{(t)}$ on the training dataset. $\mathbf{S}^{*\#} \in \mathbb{R}^{N2 \times T}$ is a reshaped matrix from the matrix $\mathcal{S}^*$ from equation (1). The classifier $\mathbf{F}^*$ is trained on the learned dynamic FC feature representation to predict the class label for each testing subject.

The model training was conducted using the 10-fold validation technique. One subject holds for testing and the remaining subjects are used for training. Note that the learned dynamic brain connectivity features can be used to train the classic SVM or other classifiers.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Results

We applied the tensor statistical model we presented in Chapter 4 mixed with the cross-validation technique and SVM from Chapter 3. The tensor statistical model employed the sliding window, the Pearson correlation, and the optimization technique. We separate the 33 patients' fNIRS time series data into different files, classified as normal and abnormal. The files were set up into two levels; level one is based on the three variants O2Hb/HbO2, and HHb and tHb exhibits oxygenation-dependent light absorbing characteristics of hemoglobin. The second level includes the measures used to classify the patients' records, either Inhibit_pct, average score of all the given clinical scores, or AHI with threshold equals to 65. Nonetheless, the AHI factor is used for a comparison purpose.

The quantitative measurement we are including are the accuracy, precision, and recall evaluating the classification performance. The final results are the average accuracy score of 10-folds. The following tables show the final accuracy and the standard deviation (std) for each subgroup of the three channels classified.

In Table 5.1. we set up the sliding window size to 40 data points, and the tensor model optimization parameters α and γ in Equation 3. to be equal to one.

Sliding window size = 40 , $\alpha = 1$, $\gamma = 1$						
fNIRS channel	**Behavior score**	**Patient Classes**	**Mean Accuracy (based on 10-Folds)**	**Std (based on 10-Folds)**	**Recall (based on 10-Folds)**	**Precision (based on 10-Folds)**

oxyhemoglobin (HbO2)	Inhibit_pct (control impulsive)	Abnormal = 20 Normal = 13	0.64	0.143	0.529	0.470
	Average score	Abnormal = 21 Normal = 12	0.74	0.150	0.629	0.524
	AHI	Abnormal = 19 Normal = 14	0.53	0.156	0.470	0.525
deoxyhemoglobin (HHb)	Inhibit_pct (control impulsive)		0.56	0.222	0.417	0.368
	Average score		0.67	0.149	0.629	0.423
	AHI		0.53	0.105	0.376	0.423
total hemoglobin (tHb)	Inhibit_pct (control impulsive)		0.64	0.096	0.62	0.42
	Average score		0.67	0.125	0.62	0.57
	AHI		0.56	0.149	0.525	0.425

Table 5.1. lists the prediction accuracy of the tensor statistical model trained on fNIRS data Sliding window size = 40 , $\alpha = 1$, $\gamma = 1$.

We evaluate the effects of the α and γ model parameters by the accuracy for lower values. In Table 5.2. we set $\alpha = 0.1,\ \gamma = 0.1$, and $\alpha = 0.01,\ \gamma = 0.01$ with window size =40.

fNIRS channel	**Behavior score**	**Patient Classes**	**$\alpha = 1,\ \gamma = 1$**		**$\alpha = 0.1,\ \gamma = 0.1$**		**$\alpha = 0.01,\ \gamma = 0.01$**	
			Mean Accuracy (based on 10-Folds)	**Std (based on 10-Folds)**	**Mean Accuracy (based on 10-Folds)**	**Std (based on 10-Folds)**	**Mean Accuracy (based on 10-Folds)**	**Std (based on 10-Folds)**
oxyhemoglobin (O2Hb)	Inhibit_pct (control impulsive)	Abnormal = 20 Normal = 13	0.62	0.162	0.54	0.0966	0.56	0.1265
	Average score	Abnormal = 21 Normal = 12	0.74	0.151	0.62	0.1229	0.66	0.1350
	AHI	Abnormal = 19 Normal = 14	0.62	0.079	0.44	0.2119	0.57	0.1160
deoxyhemoglobin (HHb)	Inhibit_pct (control impulsive)		0.61	0.144	0.53	0.1889	0.61	0.0738
	Average score		0.71	0.137	0.67	0.0823	0.62	0.1135
	AHI		0.59	0.119	0.51	0.1370	0.48	0.1398
total hemoglobin (tHb)	Inhibit_pct (control impulsive)		0.64	0.096	0.53	0.1947	0.55	0.1269
	Average score		0.67	0.125	0.63	0.0949	0.62	0.1549

	AHI		0.56	0.149	0.58	0.1687	0.44	0.2459

Table 5.2. lists the prediction accuracy of the tensor statistical model trained on fNIRS data Sliding window size = 40 , $\alpha = 0.1$, $\gamma = 0.1$, and $\alpha = 0.01$, $\gamma = 0.01$.

Moreover, we investigate the effects of the sliding windows size on the model performance. The results are presented in Table 5.3.

fNIRS channel	**Behavior score**	**Patient Classes**	Sliding window size =**15**		Sliding window size =**100**		Sliding window size =**200**	
			Mean Accuracy (based on 10-Folds)	**Std (based on 10-Folds)**	**Mean Accuracy (based on 10-Folds)**	**Std (based on 10-Folds)**	**Mean Accuracy (based on 10-Folds)**	**Std (based on 10-Folds)**
oxyhemoglobin (O2Hb)	Inhibit_pct (control impulsive)	Abnormal = 20 Normal = 13	0.48	0.147	0.59	0.152	0.60	0.133
	Average score	Abnormal = 21 Normal = 12	0.65	0.097	0.71	0.099	0.71	0.152
	AHI	Abnormal = 19 Normal = 14	0.56	0.222	0.57	0.149	0.58	0.091
deoxyhemoglobin (HHb)	Inhibit_pct (control impulsive)		0.52	0.168	0.61	0.087	0.51	0.144
	Average score		0.63	0.094	0.65	0.108	0.67	0.067
	AHI		0.52	0.139	0.50	0.163	0.55	0.085
total hemoglobin (tHb)	Inhibit_pct (control impulsive)		0.57	0.176	0.56	0.189	0.55	0.165

	Average score		0.59	0.137	0.65	0.135	0.58	0.113
	AHI		0.47	0.170	0.57	0.105	0.57	0.105

Table 5.3. lists the prediction accuracy of the tensor statistical model trained on fNIRS data Sliding window size = [15, 100, 200] , α = 1, γ = 1.

Further supportive results of the sliding window effect can be found in the appendix Table.1 considering a step size of 35, for each window in a range of [15, 190].

5.2 Discussion

From our experiments, we intended to evaluate the current model performance based on four factors: the oxygen-dependent absorption variant, the classification factor, the size of the sliding window, and the value of the optimization parameters α and γ.

First, there is no evidence of variant results under HbO2 or HHb. However, the performance based on HbO2 and HHb is slightly better than tHb. Second, all the results accompanying the average score as a classification factor produce a steadily better performance than Inhibit_pct and AHI. AHI gives the lowest score in all cases. Nonetheless, it is worth noting that currently AHI is the primary factor used by clinical experts. Third, we claim that the choice of the sliding window size directly inferences the model learning performance. From Table 5.4, we can see that the accuracy was much lower with a sliding window of size 15, where the window size of 100 and 200 presents close results. Therefore, based on our experiments, the best results occurred, it seems, when the sliding window size is 40. Lastly, when α=1 and $\gamma = 1$ the accuracy is much higher than when (α=0.1, $\gamma = 0.1$) and (α=0.01, $\gamma = 0.01$). Additional factor affects the overall model performance is the size of the training data. In general, the model could improve if we feed more data for training stage.

5.3 Conclusion

The domains of neuroscience, the fNIRS technology, the FC, and children's sleeping disorders are recognized as important fields, but a great deal still remains unknown. The fact that fNIRS is a child-friendly method to examine cognitive development moderates the limitation of understanding the neural changes. However, there is a lack of computation and statistical techniques recommended that can extract the full advantages of it, compared to the fMRI.

The presented work offers a cognitive impairment classification problem. The targeted measurements for prediction are a set of clinical behavioral measurements derived from BREIF scores. The data sample represents children aged between 4 to 12, with a possible OSA condition. This data is a unique data set; the data associated with their resting-state fNIRS time series, AHI scores, and BREIF scores, and additional apnea events, along with demographic information. The overall goal of this work is to build a regression model that can learn the dFC in order to predict the set of continuous values from BREIF scores. The outcome of this model can be used as a tool for screening children with OSA in order to treat them accordingly.

Up to this point, we overviewed the potential of the fNIRS for quantifying the FC. Also, we highlighted selected related works and models using fNIRS or fMRI to detect a cognitive impairment.

To conclude, we applied one technique in this work, because it appeared promising for tackling the problem. It did answer part of our research questions. However, further study is required since this is a novel problem.

5.4 Future work

The current work is based on a limited data set with 33 patients and fNIRS time series data from three channels on PFC. Yet, by July we expect more data from patients with additional fNIRS channels for the analysis, which would improve the current tensor statistical model robustness and allow us to approach our goal of building a regression model. More analysis on the clinical and

demographic information is required on how we can integrate the given information into the learning model. Another possibility is to focus on regression methods combined with the neural network approaches.

The research on the behavioral problems using fNIRS technology is a developing field. Where many efforts were presented to understand the FC and build useful models based on fMRI than fNIRS were exists. Thus, additional work and research in fNIRS is still required. Moreover, since fNIRS has many advantages, it is also can combined with other neuroimaging technologies such as fMRI and EGG for more gains. Moreover, the researchers might need to explain more of their methodologies and their implementation details. Hence, many papers lack the details and focus on the final results. The efforts in this field can help the researcher community to distribute the benefits of previous works and improve their techniques. The implication of this area of study, is beneficial for the clinical professionals, physicians, neuroscientists, and social and sociological workers. It will help to enhance the practices and our understanding of the human brain.

REFRENCES

[1] Abad, V. C., & Guilleminault, C. (2003). Diagnosis and treatment of sleep disorders: A brief review for clinicians. *Chronobiology and Mood Disorders, 5*(4), 371-388. doi:10.31887/dcns.2003.5.4/vabad

[2] Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine, 18*(3), 91-93. doi:10.1016/j.tjem.2018.08.001

[3] Artinis medical SYSTEMS: Fnirs and NIRS devices-portalite mini. (n.d.). Retrieved April 19, 2021, from https://www.artinis.com/portalite-mini

[4] Ayaz, H., Izzetoglu, M., Izzetoglu, K., & Onaral, B. (2019). The use of Functional Near-infrared spectroscopy in Neuroergonomics. *Neuroergonomics,* 17-25. doi:10.1016/b978-0-12-811926-6.00003-8

[5] Biswal, S., Sun, H., Goparaju, B., Westover, M. B., Sun, J., & Bianchi, M. T. (2018). Expert-level sleep scoring with deep neural networks. *Journal of the American Medical Informatics Association, 25*(12), 1643-1650. doi:10.1093/jamia/ocy131

[6] Borsini, E., Nogueira, F., & Nigro, C. (2018). Apnea-hypopnea index in sleep studies and the risk of over-simplification. *Sleep Science, 11*(1), 45-48. doi:10.5935/1984-0063.20180010

[7] Caputo, M., Denker, K., Franz, M. O., Laube, P., & Umlauf, G. (2015). Support vector machines for classification of geometric primitives in point clouds. *Curves and Surfaces,* 80-95. doi:10.1007/978-3-319-22804-4_7

[8] Cataletto, M. E., & Serebrisky, D. (2019, February 13). Childhood sleep apnea. Retrieved April 19, 2021, from https://emedicine.medscape.com/article/1004104-print

[9] Chablani, M. (2020, May 28). Deep learning on a combination of time series and tabular data. Retrieved April 20, 2021, from https://towardsdatascience.com/deep-learning-on-a-combination-of-time-series-and-tabular-data-b8c062ff1907

[10] Chen, K., Azeez, A., Chen, D. Y., & Biswal, B. B. (2020). Resting-State functional Connectivity: SIGNAL origins and analytic methods. *Neuroimaging Clinics of North America, 30*(1), 15-23. doi:10.1016/j.nic.2019.09.012

[11] Di Domenico, S. I., Rodrigo, A. H., Dong, M., Fournier, M. A., Ayaz, H., Ryan, R. M., & Ruocco, A. C. (2019). Functional near-infrared spectroscopy. *Neuroergonomics,* 169-173. doi:10.1016/b978-0-12-811926-6.00028-2

[12] Du, K. (2019, December 29). Combining time-series and tabular data for prediction. Retrieved April 20, 2021, from https://dxiaochuan.medium.com/combining-time-series-and-tabular-data-for-prediction-9815a5a17cd

[13] Dvornek, N. C., Ventola, P., Pelphrey, K. A., & Duncan, J. S. (2017). Identifying autism from resting-state fmri using long short-term memory networks. *Machine Learning in Medical Imaging,* 362-370. doi:10.1007/978-3-319-67389-9_42

[14] Emberson, L. L., Zinszer, B. D., Raizada, R. D., & Aslin, R. N. (2017). Decoding the INFANT mind: Multivariate pattern analysis (MVPA) Using fnirs. *PLOS ONE, 12*(4). doi:10.1371/journal.pone.0172500

[15] Epstein, L. J., Weinstein, M. D., Weaver, E. M., Schwab, R. J., Rogers, R., Ramar, K., Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine (2009). (2012). Clinical guidelines for the evaluation of adults with obstructive sleep apnea. *Sleep Loss and Obesity,* 191-202. doi:10.1007/978-1-4614-3492-4_13

[16] Epstein, L. J., Kristo, D., Strollo, P. J., Jr, Friedman, N., Malhotra, A., Patil, S. P., Ramar, K., Rogers, R., Schwab, R. J., Weaver, E. M., Weinstein, M. D., & Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine (2009). Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, 5(3), 263–276.

[17] Fan, L., Su, J., Qin, J., Hu, D., & Shen, H. (2020). A deep network model on dynamic functional connectivity with applications to gender classification and intelligence prediction. *Frontiers in Neuroscience, 14*. doi:10.3389/fnins.2020.00881

[18] 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. doi:10.1016/j.neuroimage.2012.03.049

[19] FNIRS & nirx: Fnirs systems: NIRS DEVICES: NIRX. (n.d.). Retrieved April 19, 2021, from,https://nirx.net/fnirs-and-nirx?gclid=Cj0KCQiAvbiBBhD-

ARIsAGM48bzGFoq_B6Fmu4j-VjHTAcTNUAd252RLWwA3F3z9lAJ9MM8cmxkABZ0aAkzYEALw_wcB

[20] Gioia, G. A., Isquith, P. K., & Roth, R. M. (2018). Behavior rating inventory for executive function. *Encyclopedia of Clinical Neuropsychology,* 532-538. doi:10.1007/978-3-319-57111-9_1881

[21] Hu, Z., Liu, G., Dong, Q., & Niu, H. (2020). Applications of resting-state fnirs in the developing brain: A review from the connectome perspective. *Frontiers in Neuroscience, 14*. doi:10.3389/fnins.2020.00476

[22] Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., . . . Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage, 80*, 360-378. doi:10.1016/j.neuroimage.2013.05.079

[23] Irani, F., Platek, S. M., Bunce, S., Ruocco, A. C., & Chute, D. (2007). Functional near infrared Spectroscopy (FNIRS): An Emerging Neuroimaging technology with important applications for the study of brain disorders. *The Clinical Neuropsychologist, 21*(1), 9-37. doi:10.1080/13854040600910018

[24] Jacobson, L. A., Pritchard, A. E., Koriakin, T. A., Jones, K. E., & Mahone, E. M. (2016). Initial examination of the brief2 in clinically referred children with and without adhd symptoms. *Journal of Attention Disorders, 24*(12), 1775-1784. doi:10.1177/1087054716663632

[25] Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374*(2065), 20150202. doi:10.1098/rsta.2015.0202

[26] K. Pavlova, M., & Latreille, V. (2019). Sleep disorders. *The American Journal of Medicine, 132*(3), 292-299. doi:10.1016/j.amjmed.2018.09.021

[27] Kotagal, S., & Pianosi, P. (2006). Sleep disorders in children and adolescents. *BMJ, 332*(7545), 828-832. doi:10.1136/bmj.332.7545.828

[28] Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*. doi:10.3389/neuro.06.004.2008

[29] Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J., Schluep, M., . . . Van De Ville, D. (2013). Principal components of functional connectivity: A new approach to study

dynamic brain connectivity during rest. *NeuroImage, 83*, 937-950. doi:10.1016/j.neuroimage.2013.07.019

[30] Linn, K. A., Gaonkar, B., Doshi, J., Davatzikos, C., & Shinohara, R. T. (2016). Addressing confounding in predictive models with an application to neuroimaging. *The International Journal of Biostatistics, 12*(1), 31-44. doi:10.1515/ijb-2015-0030

[31] Mayo Clinic. (2020, July 28). Sleep apnea. Retrieved April 19, 2021, from https://www.mayoclinic.org/diseases-conditions/sleep-apnea/diagnosis-treatment/drc-20377636

[32] McKeown, M. (2003). Independent component analysis of functional mri: What is signal and what is noise? *Current Opinion in Neurobiology, 13*(5), 620-629. doi:10.1016/j.conb.2003.09.012

[33] Meidenbauer, K. L., Choe, K. W., Cardenas-Iniguez, C., Huppert, T. J., & Berman, M. G. (2021). Load-dependent relationships BETWEEN FRONTAL fNIRS activity and performance: A DATA-DRIVEN PLS approach. *NeuroImage, 230*, 117795. doi:10.1016/j.neuroimage.2021.117795

[34] Meltzer, L. J., & Mindell, J. A. (2008). Behavioral sleep disorders in children and adolescents. *Sleep Medicine Clinics,3*(2), 269-279. doi:10.1016/j.jsmc.2008.01.004

[35] Mokhtari, F., Akhlaghi, M. I., Simpson, S. L., Wu, G., & Laurienti, P. J. (2019). Sliding window correlation analysis: Modulating window shape for dynamic brain connectivity in resting state. *NeuroImage, 189*, 655-666. doi:10.1016/j.neuroimage.2019.02.001

[36] Mukaka, M. (2012, September). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Retrieved April 20, 2021, from https://pubmed.ncbi.nlm.nih.gov/23638278/

[37] Naseer, N., & Hong, K. (2015). Fnirs-Based brain-computer interfaces: A review. *Frontiers in Human Neuroscience, 9*. doi:10.3389/fnhum.2015.00003

[38] Pinti, P., Scholkmann, F., Hamilton, A., Burgess, P., & Tachtsidis, I. (2019). 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*. doi:10.3389/fnhum.2018.00505

[39] Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2020). The present and future use of functional near-infrared spectroscopy (fnirs) for

cognitive neuroscience. *Annals of the New York Academy of Sciences, 1464*(1), 5-29. doi:10.1111/nyas.13948

[40] Refaeilzadeh, P., Tang, L., & Liu, H. (2018). Cross-Validation. *Encyclopedia of Database Systems,* 677-684. doi:10.1007/978-1-4614-8265-9_565

[41] Rogers, B. P., Morgan, V. L., Newton, A. T., & Gore, J. C. (2007). Assessing functional connectivity in the human brain by fmri. *Magnetic Resonance Imaging, 25*(10), 1347-1357. doi:10.1016/j.mri.2007.03.007

[42] Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols,12*(3), 506-518. doi:10.1038/nprot.2016.178

[43] Sui, J., Jiang, R., Bustillo, J., & Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and HEALTH: Methods and promises. doi:10.1101/2020.02.22.961136

[44] Thomas, R. M., Gallo, S., Cerliani, L., Zhutovsky, P., El-Gazzar, A., & Van Wingen, G. (2020). Classifying autism spectrum disorder using the Temporal statistics Of RESTING-STATE Functional MRI data with 3D convolutional neural networks. *Frontiers in Psychiatry, 11*. doi:10.3389/fpsyt.2020.00440

[45] Turnbull, K., Reid, G. J., & Morton, J. B. (2013). Behavioral sleep problems and their potential impact on developing executive function in children. *Sleep, 36*(7), 1077-1084. doi:10.5665/sleep.2814

[46] Viola, A., Balsamo, L., Neglia, J. P., Brouwers, P., Ma, X., & Kadan-Lottick, N. S. (2017). The behavior Rating inventory of executive FUNCTION (BRIEF) to IDENTIFY Pediatric acute lymphoblastic Leukemia (ALL) survivors at risk FOR neurocognitive impairment. *Journal of Pediatric Hematology/Oncology, 39*(3), 174-178. doi:10.1097/mph.0000000000000761

[47] Von Lühmann, A., Zheng, Y., Ortega-Martinez, A., Kiran, S., Somers, D. C., Cronin-Golomb, A., . . . Yücel, M. A. (2021). Toward neuroscience of the everyday world (new) using functional near-infrared spectroscopy. *Current Opinion in Biomedical Engineering, 18*, 100272. doi:10.1016/j.cobme.2021.100272

[48] Zhang, H., Zhang, Y., Lu, C., Ma, S., Zang, Y., & Zhu, C. (2010). Functional connectivity as revealed by independent component analysis of resting-state fnirs measurements. *NeuroImage, 51*(3), 1150-1161. doi:10.1016/j.neuroimage.2010.02.080

[49] Zhu, Y., Gilman, J., Evins, A. E., & Sabuncu, M. (2019). Detecting cannabis-associated cognitive impairment using resting-state fnirs. *Lecture Notes in Computer Science,* 146-154. doi:10.1007/978-3-030-32254-0_17

[50] Zhu, Y., Zhu, X., Kim, M., Yan, J., & Wu, G. (2017). A tensor statistical model for quantifying dynamic functional connectivity. *Lecture Notes in Computer Science,* 398-410. doi:10.1007/978-3-319-59050-9_32

APPENDIX

Experiment on raw fNIRS time series data

The figures in this section illustrate the dynamic of the recorded channels based on the HbO2 (O2Hb), HHb and tHB. It reflects the hemodynamics changes for one patient as an example. It exposes the interchanges and the similarities between the activities in the three channels.

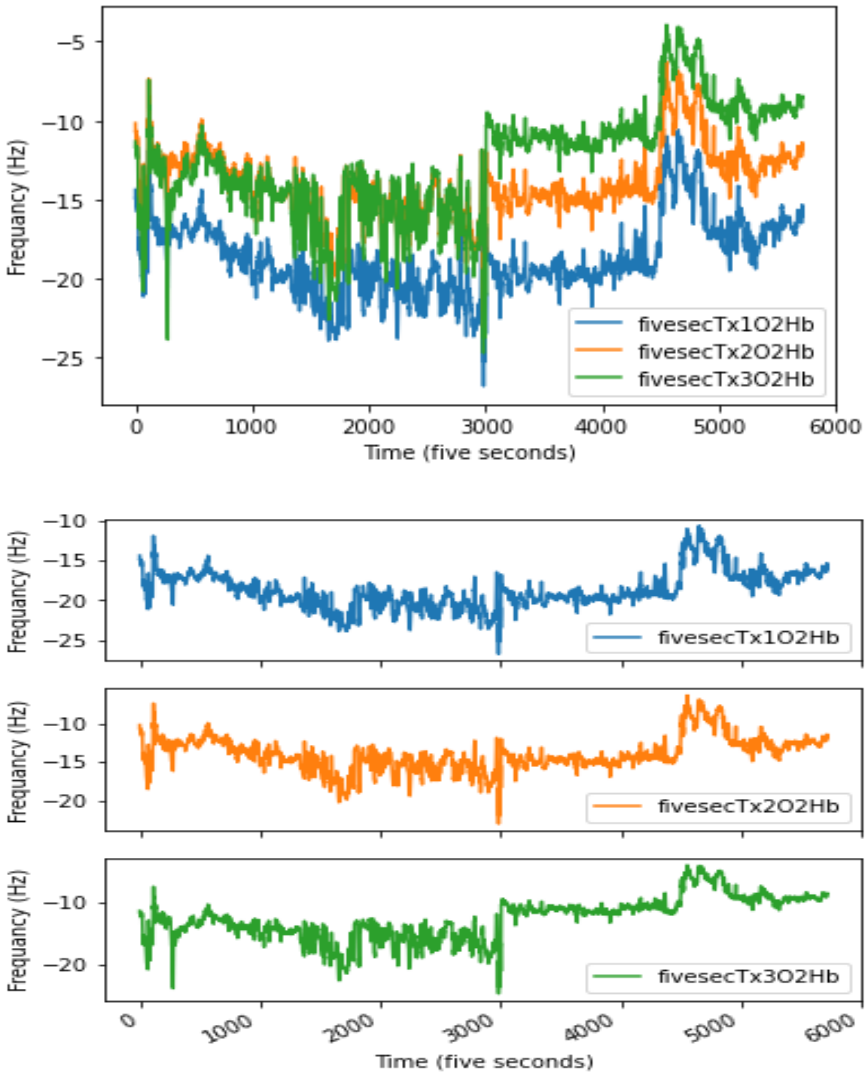


Figure 1. The hemodynamics based on O2Hb of three channels.

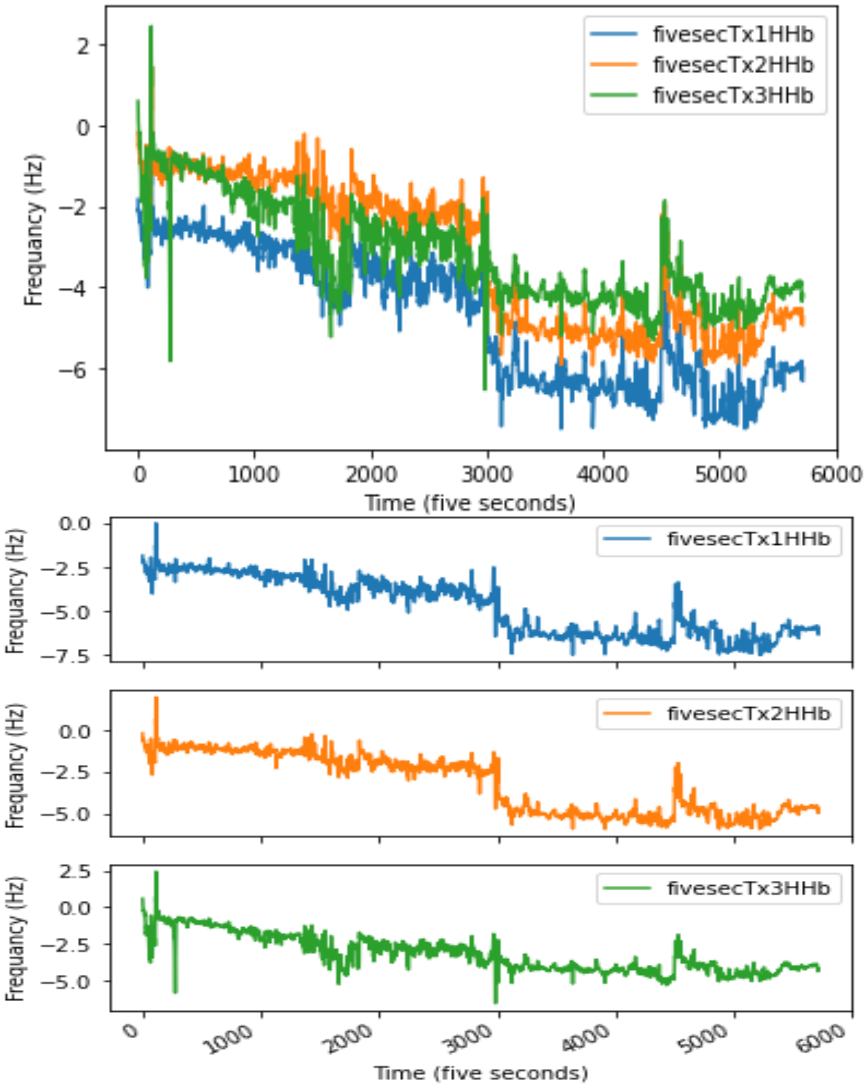


Figure 2. The hemodynamics based on HHb of three channels.

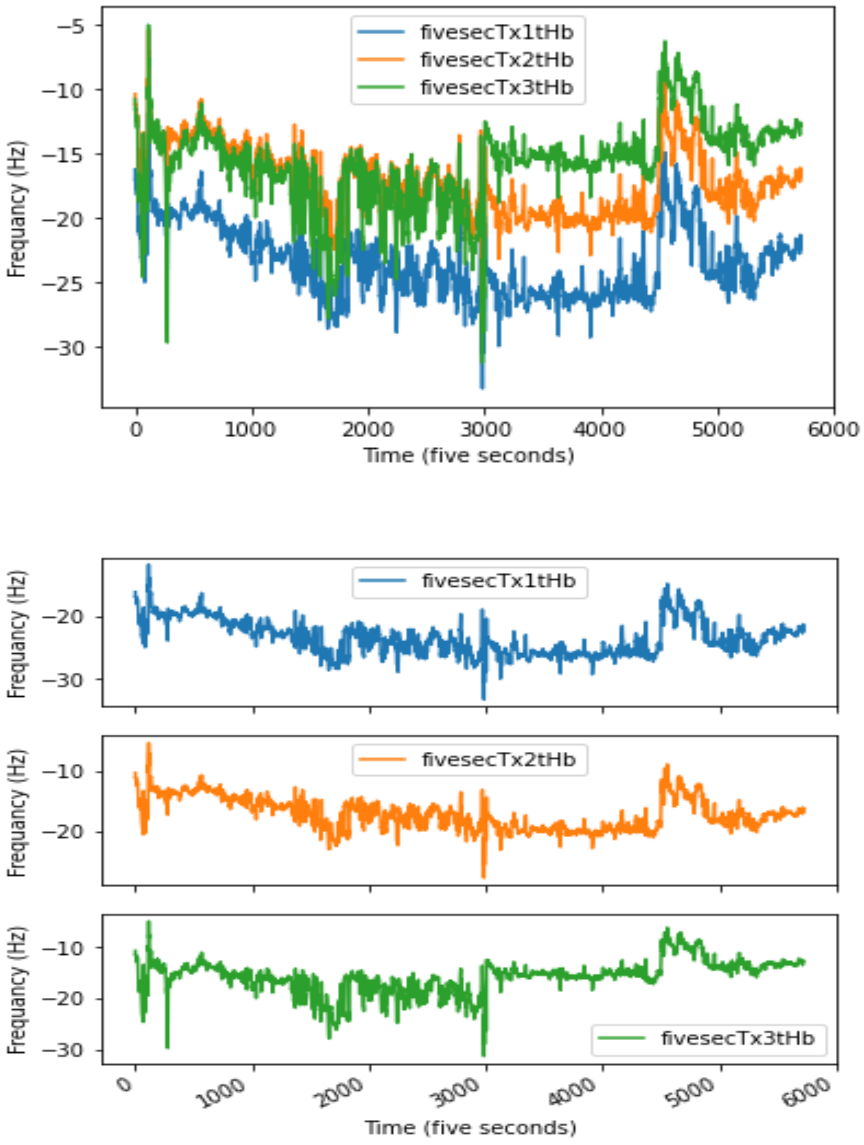


Figure 3. The hemodynamics based on tHb = 02Hb +HHb of three channels.

Next, Figure 4. shows the difference in the dynamic of the O2Hb and HHB for measurements.

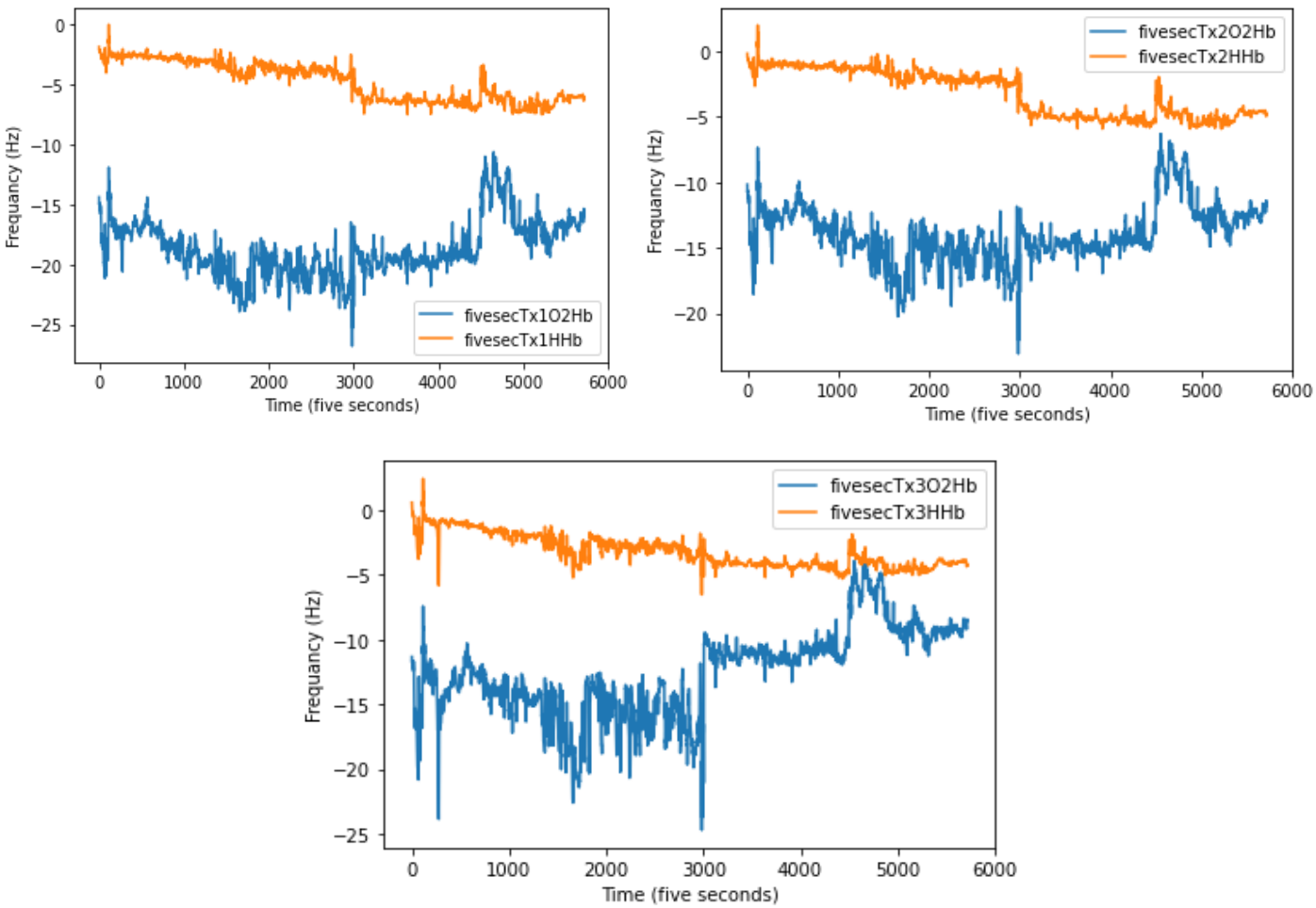


Figure 4. presents the three channels based O2Hb and HHb activity recorded by fNIRS.

Median Filter

The following figures shows one case with signal artifacts. We denoised the time series data using median filter. The median filter calculates the medina for data points within the sliding window we defined. The example shows the filter effects on the first channel based on O2Hb and HHb.

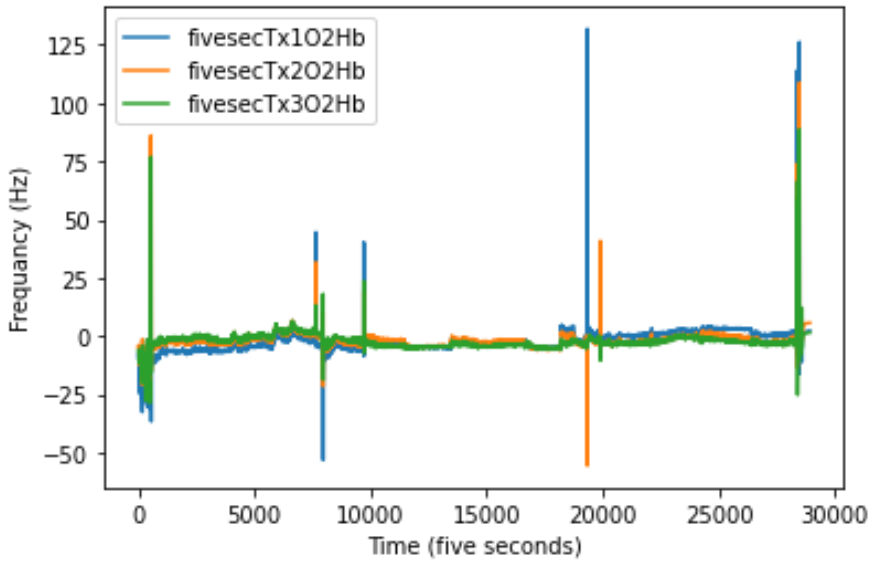
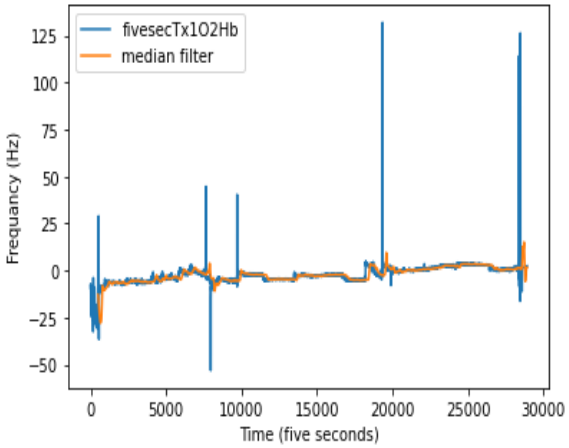
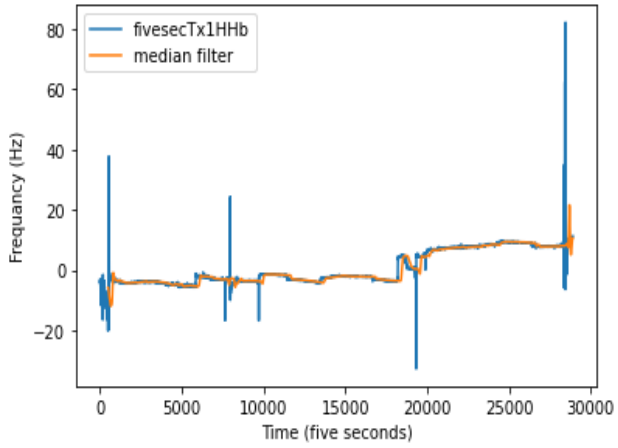
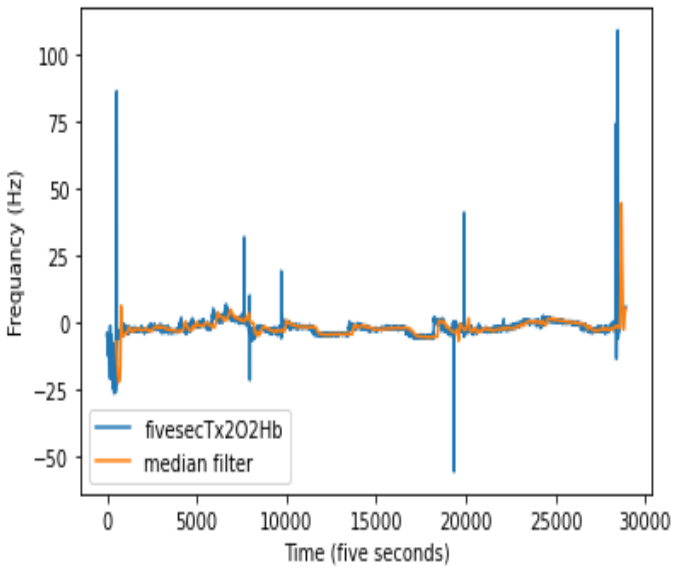


Figure 5. Raw fNIRS data based on O2Hb of three channels with artifacts.

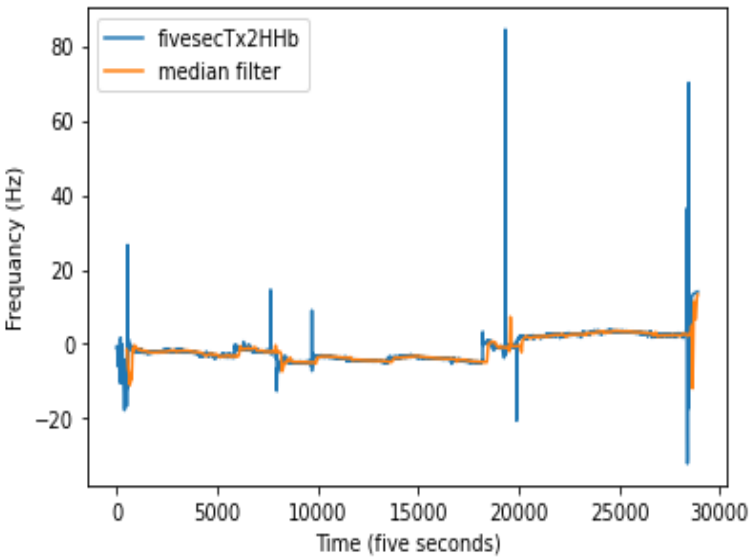
(a) (b)



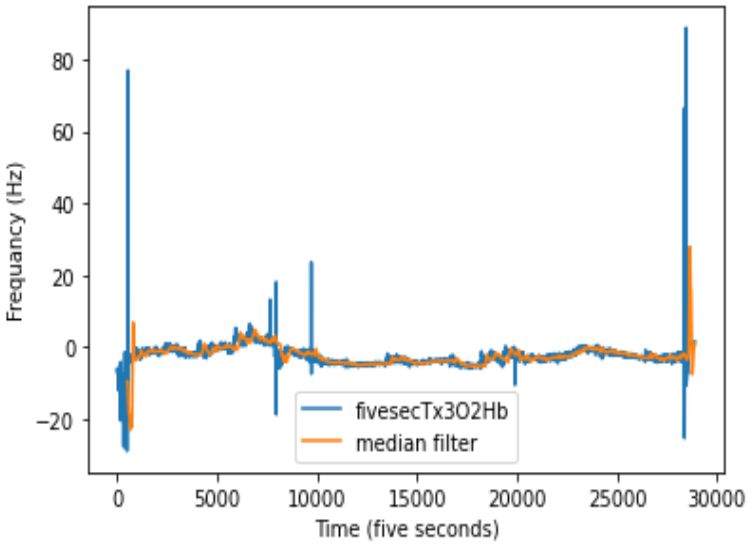




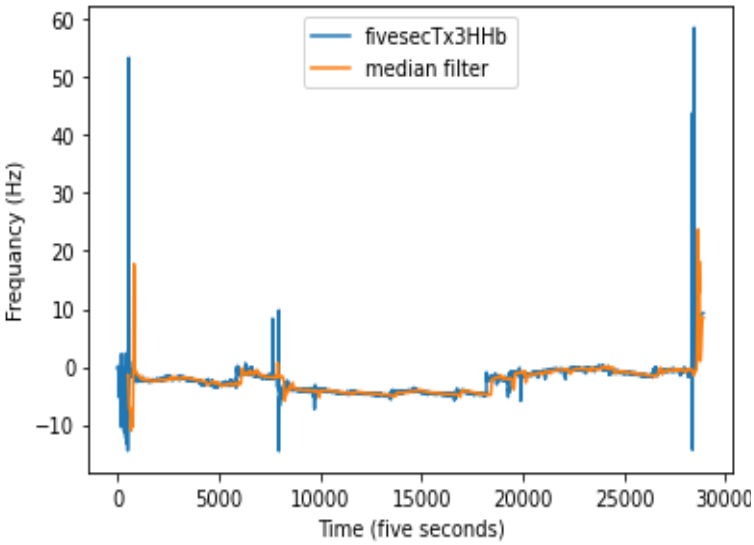
(c)



(d)



(e)



(f)

Figure 6. Shows the effect on medina filter on all the three channels disjointedly based on O2Hb and HHb.

Experiments on sliding window size for model performance

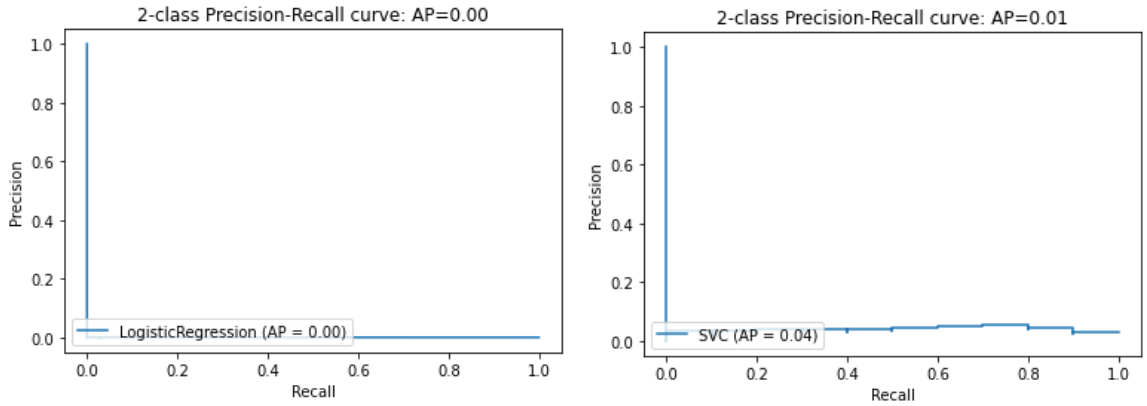
As we mentioned on Section 5.1, Table 5.3 the effects of the size if the sliding window to calculate the Pearson correlation on the final model accuracy. Table.1 is an extension of Table5.1.

fNIRS channel	**Behavior score**	**Patient Classes**	Sliding window size =15 **Mean Accuracy**	Sliding window size =50 **Mean Accuracy**	Sliding window size =85 **Mean Accuracy**	Sliding window size =120 **Mean Accuracy**	Sliding window size =155 **Mean Accuracy**	Sliding window size =190 **Mean Accuracy**
oxyhemoglobin (O2Hb)	Inhibit_pct (control impulsive)	Abnormal = 20 Normal = 13	0.48	0.60	0.61	0.55	0.63	0.67
	Average score	Abnormal = 21 Normal = 12	0.65	0.60	0.68	0.64	0.62	0.63
	AHI	Abnormal = 19 Normal = 14	0.56	0.50	0.49	0.49	0.56	0.58
deoxyhemoglobin (HHb)	Inhibit_pct (control impulsive)		0.52	0.56	0.64	0.49	0.66	0.52
	Average score		0.63	0.72	0.71	0.67	0.67	0.71
	AHI		0.52	0.48	0.53	0.51	0.51	0.50
total hemoglobin (tHb)	Inhibit_pct (control impulsive)		0.57	0.53	0.66	0.59	0.64	0.59
	Average score		0.59	0.62	0.63	0.67	0.63	0.63
	AHI		0.47	0.53	0.51	0.58	0.51	0.57

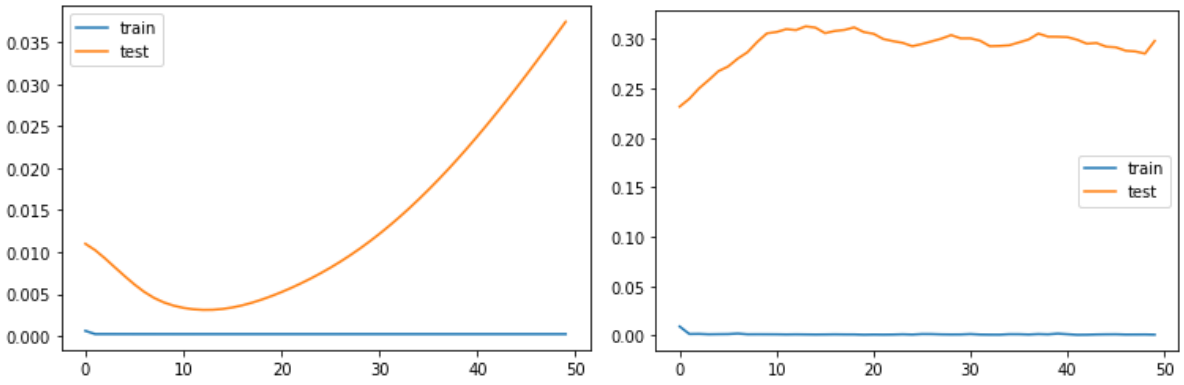
Table 1. List the prediction accuracy of the tensor statistical model trained on fNIRS data Sliding window size = [15, 50, 85, 120, 155, 190] , $\alpha = 1$, $\gamma = 1$.

Experiments on predicting OSA events from fNIRS time series data

We used the time series model and the linear regression (LR), support vector classifier (SVC), RNN, and LSTM machine learning models to predict the apnea events from the resting state fNIRS time series data points for a single subject. However, we focused on the OSA events as show case. Table 2. show high accuracy but lower precision and recall for all models, where the reason is model overfitting problem. Since the training data set is limited and the OSA events are unbalanced, meaning is rarely detecting during the 4-7 hours for multiple subjects, the model can suffer from the overfitting problem and not being able to generalize on new data.



(a) Precision-Recall curves for Logistic Regression and SVC models.



(b) Learning curves for RNN and LSTM models.

Figure 7. shows performance evaluation LR, SVC, RNN and LSTM models used to predict OSA.