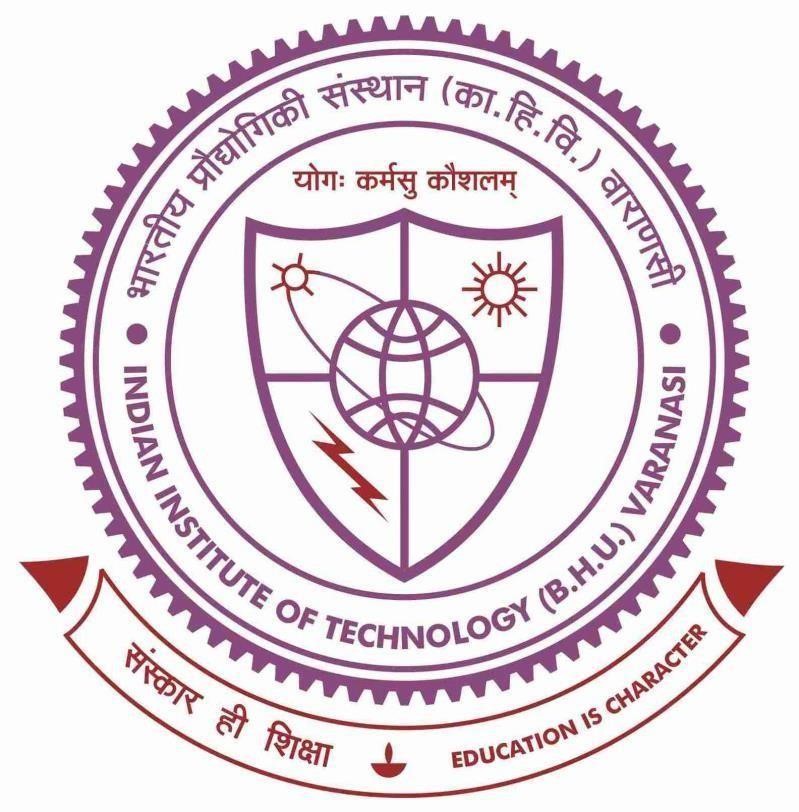
**Dynamic Functional Connectivity Analysis in fMRI for Autism Diagnostic Classification Using Machine Learning**



An Exploratory Project is submitted as a part of the coursework of IDD-3rd semester

BIOMEDICAL ENGINEERING

SUBMITTED BY

**Smarth Sood**

22024018

UNDER THE SUPERVISION OF

**Jac Fredo Agastinose Ronickom**

School of Biomedical Engineering

Indian Institute of Technology (BHU), Varanasi, 221005, India

**DECLARATION**

I, Smarth Sood hereby declare that the following document, titled “Dynamic Functional Connectivity Analysis in fMRI for Autism Diagnostic Classification Using Machine Learning” represents my original work and intellectual contributions. This exploratory project was conducted under my supervision, and I take full responsibility for its content and conclusions.

I affirm that:

All ideas, concepts, and theories presented in this project are the result of my independent research and exploration, unless otherwise cited.

Any external sources of information, including data, images, or ideas, are appropriately cited and referenced in accordance with academic and ethical standards.

The methodology employed in this project is accurately described, and the findings presented are a true representation of the data collected and analysed.

All collaborators and contributors to this project are acknowledged appropriately for their respective roles and contributions.

The project complies with the ethical standards and guidelines set forth by Indian Institute Of Technology BHU

I understand the consequences of academic dishonesty and plagiarism, and I affirm the integrity of my work in adherence to the academic standards of Indian Institute Of Technology BHU.

Date: 22nd November 2023

Signature:

Smarth Sood

22024018

**ACKNOWLEDGEMENTS**

I would like to express my sincere gratitude to all those who have contributed to the success of the exploratory project. The collaborative effort and dedication of the following individuals have been instrumental in achieving our goals:

**Dr. Jac Fredo AGASTINOSE RONICKOM** for providing

encouragement, guidance, and kindness in all respects

regarding work enabling me to work freely and develop an understanding of the subject

I would also like to extend my appreciation to Phd Scholar Chetan Rakshe who, through their collective efforts, played a vital role in the project's progress.

Furthermore, I am grateful to PARAM SHIVAY Super Computing Facility IIT for the technical support, without which this project would not have been possible.

Lastly, I want to thank the entire IIT BHU for fostering an environment that encourages innovation and exploration.

Thank you once again to everyone involved for their hard work, commitment, and enthusiasm. Your contributions have undoubtedly made a significant impact on the success of the exploratory project.

Sincerely,

Smarth Sood

22024018

smarth.sood.bme22@itbhu.ac.in

**INTRODUCTION**

Autism Spectrum Disorder (ASD) affects approximately one in 100 children globally, according to the World Health Organization's 2022 report (https://www.who.int/). ASD is characterized by difficulties in eye contact, social cues, communication, attention, and engaging in repetitive behaviors. Diagnosis involves clinical assessments of interaction, communication, behavior patterns, parental interviews, and growth history. Commonly used diagnostic tools include the Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5) (Allegra J. Johnson et al. 2023). However, the diagnostic process is complicated by diverse symptoms, comorbidities, and the time-consuming nature of these methods (American Psychiatric Association 2013). Consequently, there's an urgent need to develop automated diagnostic methods using brain biomarkers for early intervention to improve social skills in individuals with ASD.

The exploration of brain imaging data through quantitative analysis holds promise in providing valuable biomarkers to enhance the precision of brain disorder diagnoses. Non-invasive techniques like structural magnetic resonance imaging (sMRI), functional magnetic resonance imaging (fMRI), electroencephalogram (EEG), magnetoencephalography, and diffusion tensor imaging contribute to a deeper understanding of the neural circuitry associated with ASD (Asrar G. Alharthi, and Salha M. Alzahrani 2023). Despite the high-resolution capabilities of MRI modalities, challenges such as time-consuming image interpretation, handling multiple slices and images per participant, computational costs, and diverse imaging protocols exist. These challenges may burden clinicians, leading to increased workload and comprehensive examinations, potentially reducing diagnostic capabilities (Farooq et al., 2023, Moridian P et al., 2022). Moreover, the complex structure, non-linear separability, high dimensionality of data, and sequential changes of traceable signals in each voxel pose additional challenges. Nevertheless, resting-state fMRI emerges as a powerful tool for exploring the relationship between brain function and cognitive processes. This technique allows capturing the functional organization of the brain without relying on a specific task or stimuli. Additionally, the analysis of 1D time series data derived from 4D fMRI data, achieved through averaging signal intensity values for selected regions of interest across time and volumes, assists clinicians in overcoming challenges and contributes to the automation of the ASD diagnosis process (Tikaram et al., 2023).

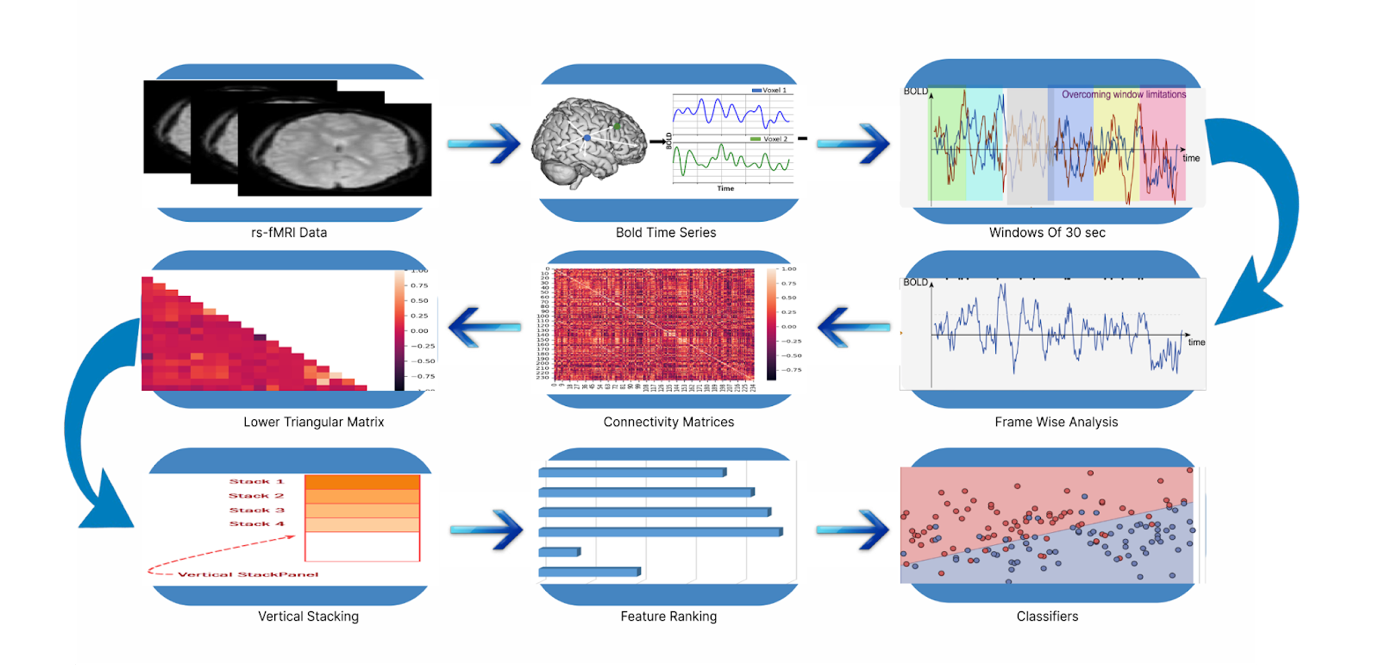
Functional connectivity (FC) is represented as a matrix with rows and columns representing nodes, and each element indicating the edge strength or functional connection between nodes. While the Pearson correlation coefficient is commonly used, it doesn't indicate direct connections. Partial correlation is employed to estimate direct connections by regressing out possible indirect connections through other nodes. Past studies often assumed that FC remains unchanged over the data acquisition time, although recent research challenges this assumption, suggesting non-stationary or dynamic functional connectivity (dFC). Our study segmented fMRI time series data into six equal segments and computed the functional connectivity matrix from each segment.

Various FC modeling approaches have been developed, including Pearson correlation, Pearson partial correlation, Spearman's rank correlation coefficient, and Gaussian covariance. Resting-state fMRI signals exhibit fractal behavior with long-range dependence, self-similarity, and power-law scaling properties in time and frequency domains. Our study leveraged the fractal properties of fMRI signals to distinguish between individuals with ASD and typically developing (TD) participants.

Advancements in artificial intelligence and machine learning have enhanced clinicians' ability to diagnose ASD. We utilized various machine learning classifiers such as XGBOOST, ANOVA F-test, Support Vector Machines (SVM), Multilayer Perceptron (MLP), k-Nearest Neighbors (KNN), Random Forest, and Logistic Regression to identify brain networks associated with these conditions (Sricheta Parui et al., 2023).

**PROCEDURE**

The process pipeline followed in the study is shown in Figure 1, which includes the

**Figure 1 Process pipeline of the study**

This study utilized fMRI data obtained from the publicly accessible Autism Brain Imaging Data Exchange (ABIDE) database. Specifically, we focused on fMRI data from the KKI site, with 47 individuals diagnosed with Autism Spectrum Disorder (ASD) and 125 typically developing (TD) individuals. Demographic information of the study participants can be found in Table 1. The fMRI data underwent standard preprocessing steps, including the creation of a whole-brain mask encompassing voxels containing BOLD signals in 95% of the participants. In our research, we employed 236 regions of interest (ROIs), which comprised 215 cortical ROIs from Gordon's atlas, 14 subcortical ROIs from the Harvard Oxford atlas, and 7 cerebellar ROIs from the Diedrichsen atlases. The blood-oxygen-level-dependent (BOLD) time series were extracted from these selected ROIs. Our investigation focused on assessing the interplay between brain regions through a method called fractal functional connectivity (FC), involving wavelet correlation of a multivariate long memory process. For each participant, we generated a 236x236 FC matrix, resulting in 27,730 diagnostic features derived from the lower or upper triangular portion of the FC matrix. We ranked these features using the XGBoost feature ranking algorithm. To evaluate our analysis, we assessed the performance of the top 100, 200, and so forth, up to 27,700 features using 5-fold cross-validation with a variety of classifiers, including:

**Table 1 Demographic information of the subjects**

|  |  |  |
| --- | --- | --- |
|  | **ASD** | **TD** |
| **Age** | 10.45±1.28 | 10.45±1.28 |
| **Gender** | 34 (M), 13 (F) | 88 (M), 37 (F) |
| **PIQ/FIQ** | 115±10 | 99±21 |

**M-Male, F-Female, PIQ-Performance intelligence quotient, FIQ-Full scale intelligence quotient.**

**RESULT:**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (a) | (b) | (c) |
|  |  |  |
| (d) | (e) | (f) |

**F**

**Figure 2 Dynamic functional connectivity of a TD subject at different window of time series (a) 0-30 seconds, (b) 30-60 seconds, (c) 60-90 seconds, (d) 90-120 seconds, (e) 120-150 seconds, (f) 150-180 seconds.**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (a) | (b) | (c) |
|  |  |  |
| (d) | (e) | (f) |

**Figure 2 Dynamic functional connectivity of an ASD subject at different window of time series (a) 0-30 seconds, (b) 30-60 seconds, (c) 60-90 seconds, (d) 90-120 seconds, (e) 120-150 seconds, (f) 150-180 seconds.**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **(a)** | **(b)** | **(c)** |
|  |  |  |
| **(d)** | **(e)** | **(f)** |
|  |  |  |
| **(g)** | **(h)** |  |

**Figure 4 Classification accuracy on different ML models (a) SVM Linear, (b) SVM RBF, (c) SVM Sigmoid, (d) Random Forest, (e) Logistic Regression, (f) KNN, (g) XGBOOST and (h) MLP.**

**Table 2 Summary of classification results**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Classifier** | **Number of features** | **Accuracy (%)** | **Sensitivity**  **(%)** | **Specificity**  **(%)** | **Precision**  **(%)** | **F1-score**  **(%)** |
| SVM Linear | 2100 | 96.28 | 95.44 | 97.25 | 96.71 | 96.03 |
| SVM RBF | 500 | 97.77 | 97.39 | 98.20 | 98.03 | 97.70 |
| SVM Sigmoid | 2400 | 95.35 | 95.61 | 95.31 | 94.99 | 95.26 |
| Logistic Regression | 1900 | 96.65 | 96.95 | 96.61 | 95.97 | 96.41 |
| KNN | 500 | 96.27 | 94.03 | 98.26 | 97.88 | 95.91 |
| Random Forest | 400 | 86.78 | 79.01 | 94.49 | 92.65 | 85.05 |
| MLP | 800 | 96.28 | 97.09 | 95.84 | 95.26 | 96.10 |
| XG BOOST | 400 | 87.53 | 87.73 | 88.24 | 86.79 | 86.91 |

**CONCLUSION**

In summary, our research demonstrates the efficacy of utilizing fractal functional connectivity (FC) as a reliable method for diagnosing Autism Spectrum Disorder (ASD). We assessed the impact of varying numbers of fractal FC features on classifier performance and found that the top 500 features yielded superior classification results. Employing the SVM RBF algorithm, our classification model achieved an impressive accuracy of 97.77%. Notably, classification accuracy was suboptimal with fewer features, and it decreased with an excessive number of features, reaching its peak at the optimal number. This investigation holds potential for future applications in the diagnosis of neurodevelopmental disorders.