Unraveling the Cognitive Secrets of Chess Experts: Investigating Dynamic Functional Brain Connectivity through rs-fMRI Analysis

Malisha Kapugamage, Dr. Rasika Rajapaksha
University of Kelaniya, Sri Lanka

***Abstract*—This study investigates the dynamic functional connectivity (DFC) in resting-state fMRI (rs-fMRI) data of chess players using a Vector Auto-Regression (VAR) model. The VAR model was constructed using the Group Lasso and Sliding Window technique. The study included 116 brain regions, and their correlation was examined in the context of their dynamic connection. Statistical feature selection techniques were used to determine which dynamic connections of brain areas were crucial in discriminating chess masters from novice players. After identifying key DFCs related to these brain regions, a classification model was built to classify chess experts and normal control individuals.**

Our classification model achieved an accuracy of 96.33% under a 10-fold cross-validation framework. This performance represents a substantial improvement over previous studies utilizing only rs-fMRI data, which reported a maximum accuracy of 85.45%, indicating a 10.88% enhancement in accuracy. Moreover, our model outperformed methods that combined rs-fMRI with T1-weighted MRI data, which achieved an accuracy of 88%, yielding an additional 8.33% improvement. These results demonstrate that our approach, relying solely on rs-fMRI data, offers a notable advancement in the classification of chess expertise.

***Index Terms*—Dynamic Functional Connectivity, Vector Auto-Regression, resting-state fMRI, Chess Masters and Novices, Classification**

I. Introduction

Brain network analysis is an exciting subject in human connectome projects whose goal is to create a network map that solves the anatomical and functional connectivity within the healthy human brain along with producing a body of data that clears the way for researching brain disorders [1]. The most often used imaging modalities to investigate functional and structural connection patterns of brain areas are fMRI (functional magnetic resonance imaging) and DTI (diffusion tensor imaging). Diffusion imaging aids in the identification of structural connections between brain regions, while functional magnetic resonance imaging (fMRI) involves functional interactions between brain regions at macroscopic scales. The fMRI infers brain activity by indirectly monitoring variations in blood flow. Resting-state fMRI (rs-fMRI), a sub-modality of fMRI, is a valuable technique for analyzing regional interactions that occur within the brain while a subject is not doing an explicit task. In both rs-fMRI and task-based fMRI, the output depends on past and current conditions, representing a causal system. The Granger Causality model was used to model the impact of previous tasks on distinct brain areas [2], [3]. The application of rs-fMRI in the research and clinical setting has been growing for the past 2 decades. The resting-state functional connectivity technique has been applied for different mental disorders such as the detection of Autism, Schizophrenia, and Alzheimer's disease, etc. An accurate rs-fMRI classification model will help to provide diagnostic and prognostic information for the above applications.

A. Summary of Our Contributions

- A comprehensive literature review of the most recent techniques for analyzing rs-fMRI data was conducted.
- The application of vector auto-regression for detecting dynamic functional connectivity (DFC) in experimental rs-fMRI data was investigated, and various correlation metrics were compared.
- A classification model was developed to differentiate rs-fMRI data based on their DFC patterns.
- The effectiveness of the proposed method was evaluated by classifying rs-fMRI scans of chess experts and control individuals.

B. Background and motivation

Resting-state fMRI (rs-fMRI), a sub-modality of fMRI, is a valuable technique for analyzing regional interactions that occur within the brain while a subject is not doing an explicit task. Since there is no task, rs-fMRI is going to be attractive to patients who may struggle with task instructions, such as those with neurologic, neurosurgical, or psychiatric disorders, as well as children [4]. Thus, during the past two decades, the use of rs-fMRI in research and clinical settings has increased. The resting-state functional connectivity technique has been applied for different mental disorders such as the detection of Autism, Schizophrenia, and Alzheimer's disease, etc.

Brain scan rs-fMRI results can also be interpreted as causal signals. When no external stimulus (task) is applied during the resting state, activation of brain areas is thought to be steady [5]. The existence and dominance of several types of mental activity during the resting state were identified

Correspondence: Malisha Kapugamage (E-mail: malishagk15@gmail.com)
Received: 16-06-2024 **Revised:** 12-08-2024 **Accepted:** 09-09-2024
Malisha Kapugamage and Rasika Rajapaksha are from University of Kelaniya (malishagk15@gmail.com, rasikar@kln.ac.lk)
DOI: https://doi.org/10.4038/icter.v18i2.7294
The 2025 Special Issue contains the full papers of the abstracts published at the 24th ICTer International Conference.

in [6]. Resting-state patterns of correlated brain activity have been demonstrated in numerous studies to be constant across subjects and diagnostic of disease states when abnormal [7]. rsfMRI may be thought of as a dynamic imaging modality that monitors brain activity across time. However, the frequently used methods for analyzing rs-fMRI signals are based on the examination of complete time series [8]. This method, on the other hand, fails to observe inter-region connections and might miss subtle patterns that might be useful in distinguishing functionally different networks. DFC (dynamic functional connectivity) describes the functional interaction between brain areas and the changes that occur over time [9]. One of the most commonly used methods for detecting DFC is the sliding window technique, which divides the signal into overlapping subsamples [10].

However, the challenge of identifying change points in the signal in order to establish an optimal window size has proven to be difficult. The majority of approaches take an experimental approach, in which window size is chosen based on the best result [11], [12]. Identifying change points in the signal to produce windows of similar width from the rs-fMRI data has been a challenge. After identifying change points in the signal, a novel approach for learning, discriminating features (connections), and generating functional connectivity networks is also required. The identified features with dynamic functional connectivity can then be used to create a classification model for classifying rs-fMRI data.

II. Related work

Dynamic functional connectivity (DFC) or its network analog, dynamic functional network connectivity (DFNC), refers to the connectivity between pairs of brain regions (or networks) within sub-intervals of time series, as opposed to conventional static functional connectivity (SFC), which is derived from the correlation within an entire time series [13]. Indeed, DFC research shows that cognitive impairments and clinical symptoms associated with many mental illnesses are dependent not just on the degree of connectivity between brain areas, but also on how that connectivity varies over time [13], [14]. In [15], a comprehensive comparison of several correlation metrics was carried out. They demonstrated that Mutual Information (MI) and Variation of Information (VI) metrics produced the most consistent findings by attaining high reliability with regard to DFC estimations for various window widths using test-retest datasets. On the basis of resting-state functional MRI data, a variational inference method to a Bayesian multi-subject vector autoregressive (VAR) model for inference on effective brain connectivity was presented in [16], [17].

An experiment revealed a transitory and selective structural alteration in brain areas related to the processing and storage of complex visual motion [18]. In response to external demands, their findings suggest that the structure of the mature human brain changes.

Long-term learning skills and activity repetition can result in significant anatomical and functional changes in the brain [19]. To identify the neurological basis of long-term professional chess training, the changed functional couplings of Visual motion area (MT) subregions in rs-fMRI data of chess players were analyzed using resting-state functional connectivity (RSFC) and Granger causality analysis (GCA). They discovered that long-term professional chess training might change the functioning of the MT region [20].

Resting-state functional connectivity and graph-theoretical studies were used to analyze the overall architecture of brain functional networks in chess players. The results show an increase in functional connectivity between the hippocampus, thalamus, basal ganglia, and many parietal and temporal regions in chess masters. These results reveals a link between brain network functional architecture and the subjective nature of cognitive expertise, which might give more insight into the processes underlying expert behavior [21].

Spatial independent component analysis (sICA) has been used to examine dynamic connectivity parameters in rs-fMRI data from chess players [22]. Four meta-state dynamic fluidity indices were investigated using sliding window correlation and meta-state approaches. According to the findings, Professional chess players demonstrated increased dynamic fluidity and dynamic range, as well as increased travel between successive meta-states.

The voxel-wise changes of functional connectivity patterns have been identified in rs-fMRI data from chess players [20]. The alteration of corresponding functional couplings was analyzed using Seed-based functional connectivity analysis. According to the findings, there were fewer functional connections between the precentral gyrus and the primary visual cortex (V1) in chess players. Additionally, a classification model based on support vector machines (SVM) distinguished between experienced players and beginners with an accuracy of 85.45%.

In order to classify professional and amateur chess players, a novel method for combining metrics from anatomical and functional brain images was presented [23]. Chess players' functional connectivity (FC) and morphometric connectivity (MC) were estimated using rs-fMRI data and T1-weighted MRI data, respectively. As a result, the machine learning-based model achieved a classification accuracy of 88%, demonstrating an improvement over the conventional functional connectivity-based approach where the accuracy was 76.33%.

III. Method

A. Dataset

The Chess Masters and Novices (CM&N) dataset was used to evaluate the effectiveness of our proposed technique. Previous studies have shown that long-term learning and repeated practice can lead to substantial anatomical and functional changes in the brain. Specifically, they have found that extended professional chess training may enhance functional connectivity across multiple brain regions. Resting-state functional magnetic resonance imaging (rs-fMRI) provides a valuable tool for assessing changes in intrinsic functional connections and interactions within the brains of chess players. The dataset contained a multi-modal rs-fMRI data of 27 professional Chinese chess players, the majority of whom

are chess grand masters and masters (CM/Ms), as well as 27 chess beginners whose age and gender well-matched. Our goal is to classify Chess Masters and Novices (CM&N) testing individuals as either professional or novice chess players. The rs-fMRI row scans were transformed into time series data, which indicated the changes in 116 brain regions over time (200s).

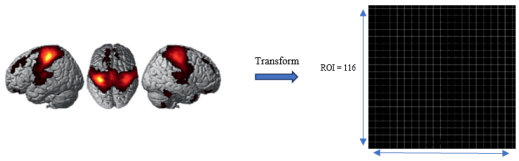

Fig. 1: rs-fMRI Row Data Transformation

- Region of interest (ROI) = 116
- n time points = 200s
- All number of subjects = 54
- Chess masters: 27
- Chess novices: 27

B. Stride and Window Length Selection with Vector Auto-Regression (VAR)

Using the VAR approach, the window size k and stride s are selected. The order of the VAR processes is determined by the window size. Ten participants were randomly chosen from the Chess Masters and Novices dataset while guaranteeing representation from both master and novice classes in order to define the order of the VAR processes. The regularization parameter λ was set to 0.1, and errors were computed for various window lengths and strides to predict the regression order using Lasso, The experiment considered 610 combinations of window length and stride in order to determine the best combination of window length and stride. The window length k ranges from 60 to 120, while the stride s ranges from 1 to 10. The relative prediction errors from all window length and stride combinations are recorded, and the threshold of 0.02 is utilized to determine the best pair. Please refer to the following figure for an explanation of stride and window length selection process.

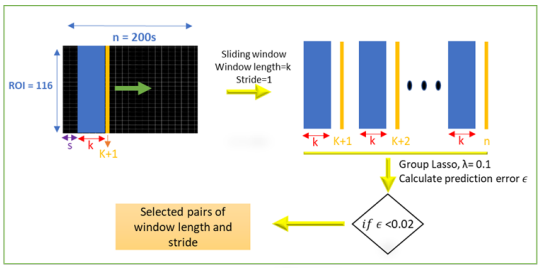

Fig. 2: Stride and Window Length Selection Process

C. Feature Generation and Selection

After determining the optimal combination of window length and stride, a sliding window of the chosen length was applied to the rs-fMRI signal, advancing by the selected stride. The process resulted in r total blocks when the sliding window traversed the entire signal. The value of r is calculated using the following equation,

$$r = \frac{(T-k)}{s} + 1$$

Where:

- $T = 200$ is the total number of time points in the rs-fMRI data.
- k is the selected window length for the sliding window.
- s is the selected stride for the sliding window.

Afterward, a correlation matrix was generated for each block, and the average of all these correlation matrices was computed, resulting in a single correlation matrix for each subject. The precision matrices were then calculated for the averaged correlation matrices of all subjects. From each precision matrix, the upper triangular entries were extracted and vectorized into a vector containing 6,676 entries, giving each subject a unique vector. The final dataset was constructed with subjects as rows and the 6,676 vector entries as columns. This presented a high-dimensionality challenge, as the number of features exceeded the number of subjects. To mitigate this issue, feature selection and dimensionality reduction techniques were employed.

In this study, dimensionality reduction was performed using the Wilcoxon Rank Sum (WRS) test, Mutual Information (MI), and Principal Component Analysis (PCA). By adjusting parameters (such as the p-value in the WRS test), altering algorithm structures, and combining different approaches, these three methods yielded nine distinct feature selection strategies. These strategies effectively reduced the original 6,676 features to a much smaller subset. For a detailed explanation of the feature generation and selection process, please refer to the following figure.

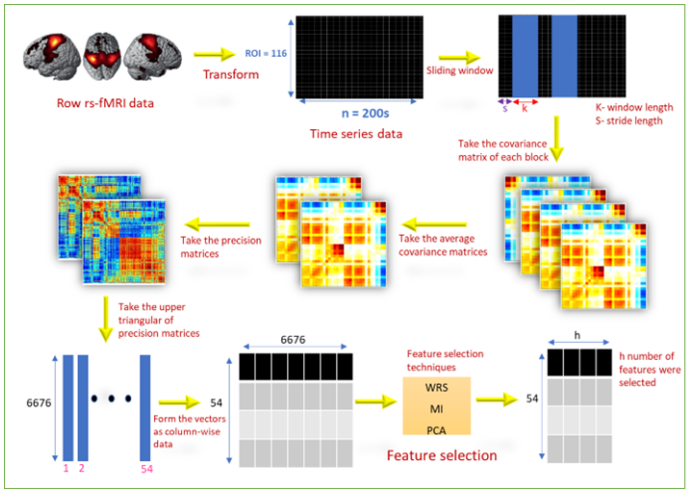

Fig. 3: Feature Generation and Selection Process

D. Classification

The objective of this classification task was to distinguish between chess masters and novices using the CM&N dataset. In total, eight machine learning algorithms and one deep learning algorithm were tested. Hyperparameter tuning was conducted alongside model training to identify the optimal parameters. Ultimately, 11 learning models were developed, consisting of nine individual models and two ensemble models, where the latter combined the best-performing models that achieved over 90% accuracy. Following the feature selection process, nine distinct feature sets were available. Each of these feature sets was used to train the nine individual learning models. The top-performing models and feature selections were then combined to form the ensemble models. Detailed descriptions of all 11 learning models are provided in the following table.

TABLE I: Learning Models Description

Model	Model Description
Model 1	Support Vector Machine (SVM)
Model 2	Extreme Gradient Boosting (XGBoost)
Model 3	Logistics Regression (LR)
Model 4	Light Gradient Boosting Machine (LightGBM)
Model 5	Multilayer Layer Perceptron (MLP)
Model 6	Extremely Randomized Trees (ExtraTrees)
Model 7	Random Forest (RF)
Model 8	Adaptive Boosting (AdaBoost)
Model 9	Category Boosting (CatBoost)
Model 10	The ensemble model of Model 1 + Model 3
Model 11	The ensemble model of Model 3 + Model 9

E. Evaluation

Before applying a classification model to real-world data, it is essential to evaluate its performance. Performance metrics such as accuracy, precision, recall, and F1-score are commonly used to assess how well the models perform. In this study, model evaluation was conducted using 10-fold cross-validation. For each combination of learning model and feature selection, the average values of accuracy, precision, recall, and F1-score were recorded, along with their standard deviations (SD).

IV. Results

A. Stride and Window Length Selection

The heatmap of prediction errors for various window lengths and strides is displayed in the figure below. The window length varies from 60 to 120, and the stride ranges from 1 to 10. The range of errors that each block of the heatmap follows can be determined using the bottom color bar. If the block is darker, it indicates that the error is minor. If the block is lighter, it indicates that the error is considerable. In addition, there were some special hues in the color bar that allowed us to identify specific error points. The objective, however, is to determine which block has the least error. According to the heatmap, the blocks with stride 1 appear to be darker than the other blocks. Then stride 1 might be considered the best stride.

The number of pairs under various error thresholds is shown in the table below. There are three thresholds: 0.1, 0.05, and 0.02. If 0.02 is chosen as the threshold, there will be 61 potential window length and stride pairs.

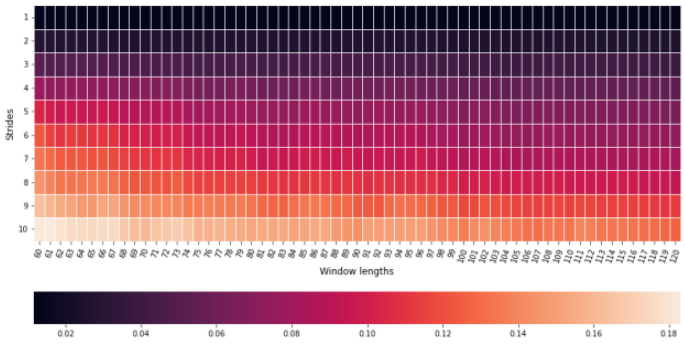

Fig. 4: Heatmap of Prediction Errors

TABLE II: Number of Pairs under Error Thresholds

Threshold	Number of pairs
0.1	410
0.05	182
0.02	61

Upon thorough examination, all 61 potential pairs were found within stride 1. As a result, stride 1 is confirmed to be the optimal stride. The following line plot represents the variation of the prediction error under different window lengths with the selected optimal stride length 1. Using the line plot, it is simple to see that the window length with the least error is 64. The precise error of window length 64 with stride 1 was 0.011556.

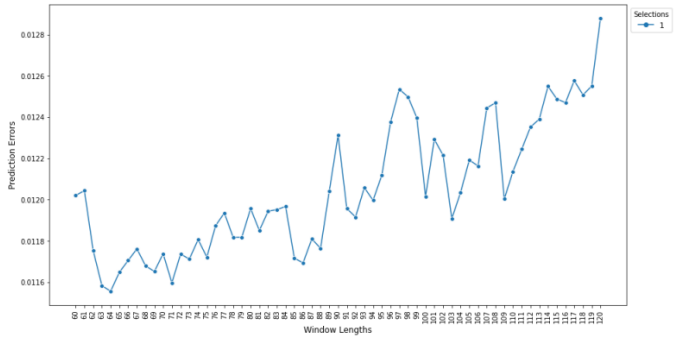

Fig. 5: Deviation of Errors under Window Lengths with Stride 1

Using the preceding results, 1 and 64 can be selected as the optimal stride and window respectively.

B. Feature Generation and Selection

The following table shows the number of features for each feature selection approach that were used to select discriminating features from a dataset of 6676 features. The WRS test was used to generate 8 out of 10 feature selections. The number of features in each feature selection is less than 100.

TABLE III: Number of Features under each Feature Selections

Feature Selections	Number of features	MI	PCA	WRS test
Selection 1	51	✓	✗	✗
Selection 2	55	✗	✓	✗
Selection 3	71	✗	✗	✓
Selection 4	20	✗	✗	✓
Selection 5	41	✗	✗	✓
Selection 6	21	✗	✗	✓
Selection 7	43	✓	✗	✓
Selection 8	86	✓	✗	✓
Selection 9	18	✓	✗	✓

C. Classification and Model Evaluation

This study employed a range of learning algorithms to experimentally determine the best classification model from among the most commonly used models. A total of nine individual learning models were tested. The following multi-line plot (Figure 6) illustrates the accuracy deviations for each combination of model and feature selection, allowing for a comparison of model performance across different feature sets. Each line represents a feature selection, with the X-axis indicating the models and the Y-axis displaying accuracy based on 10-fold cross-validation. The red horizontal line at 90% accuracy marks a threshold to highlight feature selection and model combinations that achieved greater than 90% accuracy. From this plot, nine combinations that exceed this threshold can be identified. Notably, the green line corresponding to Selection 3 shows that Model 1 and Model 3 achieved the highest accuracy with this feature selection. Therefore, Model 1 with Selection 3, and Model 3 with Selection 3 can be regarded as the best combinations of learning models and feature selections based on accuracy.

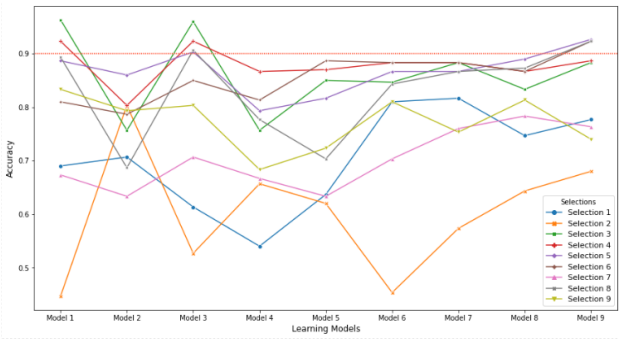

Fig. 6: The Change of Accuracy with Different Learning Models and Feature Selections

Since accuracy was determined using 10-fold cross-validation, it represents the average accuracy across 10 subsamples of the dataset. Therefore, it is crucial to examine the standard deviation (SD) of accuracy to check for overfitting. The following multi-line plot (Figure 7) depicts the SD for each model and feature combination. The feature selections are represented by lines, while the models are represented by the X-axis. The Y-axis represents the standard deviation of the accuracy measured using a 10-fold cross-validation. The red horizontal line at an SD of 0.10 serves as a threshold value, with lower SDs being preferable. An SD greater than this threshold suggests that the model may have overfitted to some subsamples during cross-validation. The plot highlights 7 model and feature selection combinations with SDs below this threshold. Among the two top-performing combinations identified in Figure 6 based on accuracy, only Model 1 with Selection 3 has an SD below the threshold value. Consequently, this combination could be selected as the most effective combination of individual learning model and feature selection.

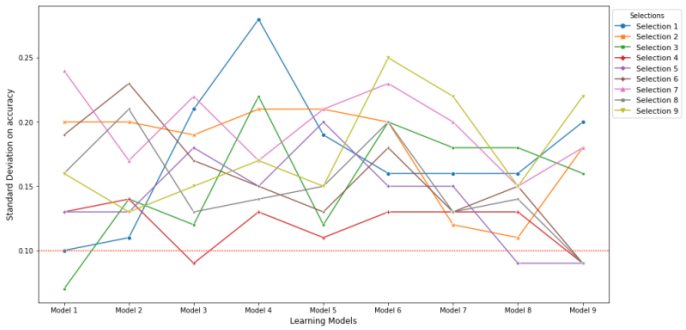

Fig. 7: The Change of SD of Accuracy with Different Learning Models and Feature Selections

The table below presents 9 combinations of models and feature selections with accuracy greater than 90%, along with their precise accuracy and standard deviation (SD). Additionally, the table highlights the two combinations identified from Figure 6, with their corresponding accuracy and SD.

TABLE IV: Models and Feature Selections Combinations with 90 Percent Plus Accuracy

Number	Feature Selections	Model	Accuracy: Mean	SD
S3M1	**Selection 3**	**Model 1**	**0.9633**	**0.07**
S4M1	Selection 4	Model 1	0.9233	0.13
S3M3	**Selection 3**	**Model 3**	**0.9600**	**0.12**
S4M3	Selection 4	Model 3	0.9233	0.09
S5M3	Selection 5	Model 3	0.9033	0.18
S5M9	Selection 5	Model 9	0.9267	0.09
S6M9	Selection 6	Model 9	0.9233	0.09
S8M3	Selection 8	Model 3	0.9067	0.13
S8M9	Selection 8	Model 9	0.9233	0.09

Additionally, the most accurate models were combined using a voting classifier, which utilizes majority voting based on predicted class labels. This ensemble learning approach was applied to Models 10 and 11. Consequently, a new model and feature selection combination, S1M10, was identified, achieving over 95% accuracy, as shown in Table V.

TABLE V: Ensembled Models and Feature Selections Combinations with Accuracy

Number	Feature Selections	Model	Accuracy: Mean	SD
S3M10	**Selection 3**	**Model 10**	**0.9633**	**0.07**
S4M10	Selection 4	Model 10	0.9433	0.09
S5M11	Selection 5	Model 11	0.9067	0.13
S8M11	Selection 8	Model 11	0.9067	0.13

Overall, three combinations of models and feature selection methods—**S3M1**, **S3M3**, and **S3M10**—have been identified with an accuracy of over 95%. The performance of these

models can be evaluated using metrics such as Accuracy, Precision, Recall, and F1-score.

TABLE VI: Final Feature Selections and Models Combination with Performance Metrics

Number	Feature Selections	Model	Metric	Value
S3M1	Selection 3	Model 1	Accuracy: Mean	0.9633
			SD	0.07
			Precision: Mean	0.9750
			SD	0.07
			Recall: Mean	0.9667
			SD	0.10
			F1-score: Mean	0.9657
			SD	0.07
S3M3	Selection 3	Model 3	Accuracy: Mean	0.9600
			SD	0.12
			Precision: Mean	0.9500
			SD	0.15
			Recall: Mean	0.9500
			SD	0.15
			F1-score: Mean	0.9500
			SD	0.15
S3M10	Selection 3	Model 10	Accuracy: Mean	0.9633
			SD	0.07
			Precision: Mean	1.0000
			SD	0.00
			Recall: Mean	0.9167
			SD	0.17
			F1-score: Mean	0.9467
			SD	0.11

Because the feature selection is the same for all combinations, Selection 3 can be chosen as the best feature selection with discriminating features for classifying rs-fMRI data from the CM&N dataset. The accuracy and SD of S3M1 and S3M10 are the same. S3M10 has the highest precision (100%), which indicates that all positively predicted data is indeed positive. However, its recall is lower than S3M1, implying that the probability of S3M10 which can predict some positive data as negative is higher than the probability of S3M1 which can predict some positive data as negative. However, the F1-score of S3M1 was higher than the F1-score of S3M10. As the harmonic mean of precision and recall, the F1-score for S3M1 is high, indicating that it is theoretically the best-performing model.

V. Discussion

A. Stride and Window Length Selection

The selection window and length and stride were a crucial part of the implementation of the proposed approach since the selected window length and stride are key to estimating DFCs in rs-fMRI data. In the process, a wide range of window lengths and strides were evaluated. As shown in Figure 4, as the size of the stride increases, so does the prediction error. When the stride size is set to 10, large window lengths generate lower error rates than small window lengths, whereas when the stride size is set to 1, small window lengths generate lower error rates than larger window lengths (Figure 5).

As presented in Table II and Figure 4, The majority of the 410 pairs fall between strides 1 and 5 at the error threshold of 0.1. The majority of the 182 pairs fall between strides 1 and 3 when the error threshold is set to 0.05. All 61 pairs have stride 1 at the final threshold of 0.02. Using that insight, stride 1 can be considered as the best window length.

Figure 5 represented the deviation of the prediction error under different window lengths with the selected stride length 1. Figure 5 showed a significant positive relationship between prediction errors and window length. That is, as the window length increases, so does the prediction error.

By carefully analyzing the graph, it is easy to identify the window length with the lowest error. The window length with the minimum error of 0.011556 is 64, with the error calculated to the fourth decimal place. However, evaluating the error change to this precision might not be necessary. When the threshold is set to 0.02, there are 61 window lengths that could be considered as potential candidates for the optimal window length, as their error differences are minimal enough to be considered negligible.

B. Feature Generation and Selection

3 distinct feature selection methods were utilized in this study. By combining and modifying the parameters of these algorithms, it was possible to produce 9 feature selections. These 9 feature selections vary in number of features. Nevertheless, no feature selection contains more than 100 features. This implies that 6676 features were able to be reduced to a significantly lower amount. In terms of the methods utilized to develop these feature selections, the WRS test stands out because it has contributed to 8 of the 10 feature selections.

C. Classification and model evaluation

Figure 6 illustrates the accuracy changes of all individual learning models with all feature selections. There have only been 9 times where any line has surpassed the red threshold line. That indicates that 9 different combinations of individual learning models and feature selections had an accuracy that was more than 90%. Table IV elaborates on these 9 combinations. Nevertheless, there were only two different combinations that had an accuracy of more than 95%. Table V demonstrates another model and feature selection combination with an accuracy greater than 95%. This model was created using combining previous models which has the accuracy over 95%. These three combinations are then summarized in Table VI along with the performance metrics to evaluate their performance. The first insight is that the feature selection process is the same for all combinations. Selection 3 consistently appears across all combinations, indicating that it is the most effective feature selection approach for our model. Therefore, Selection 3 identifies the most significant discriminative features of rs-fMRI brain scans, enabling the classification of Chess Masters and Novice players based on their dynamic functional connectivity. By analyzing these key features, we can trace them back to specific brain regions, providing insights into which areas are most influential in mastering chess.

In Table VI , we compare final three models with accuracy over 95%: S3M1, S3M3, and S3M10. S3M1 and S3M10 both exhibit the highest accuracies, with identical standard deviations, making them strong candidates for the best model.

However, when examining additional metrics—precision, recall, and F1-score—the differences between these models become more significant.

1) Understanding Precision and Recall in the Chess Expert Classification Context: High precision indicates that when the model predicts a subject as a novice, there is a high likelihood that the subject is indeed a novice.

High recall indicates that the model successfully identifies most novices, but it may include some false positives (chess masters misclassified as novices).

Given that novices are labeled as positive and chess masters as negative, precision and recall must be interpreted with care. S3M10, with its perfect precision (100%), ensures that every subject predicted as a novice is truly a novice. However, its lower recall suggests that some novices might be overlooked and incorrectly classified as chess masters. This scenario is less concerning if the goal is to ensure that every novice identified is accurate, but it may miss true novices.

Conversely, S3M1, with a higher recall, is more effective at identifying all novices, but its slightly lower precision means it may misclassify some chess masters as novices. This is crucial if the priority is to ensure that no novice is misclassified, even if it means some chess masters are incorrectly identified as novices.

2) Proposed Criteria for Model Selection: The selection of the best model depends on the specific costs associated with false positives (novices misclassified as chess masters) and false negatives (chess masters misclassified as novices) in the given scenario. Below are two possible scenarios:

a) Scenario 1: Prioritizing the Avoidance of False Positives (Novices Misclassified as Chess Masters): In situations where it is critical to ensure that only true chess masters are identified (e.g., for a high-stakes chess competition or an exclusive chess master certification), S3M10 might be preferable. Its perfect precision ensures that no novices are mistakenly classified as chess masters, even if it means some true novices are missed.

b) Scenario 2: Prioritizing the Avoidance of False Negatives (Chess Masters Misclassified as Novices): In cases where missing out on identifying a true chess master is more detrimental (e.g., for selecting a candidate with a high level of chess expertise), S3M10 might be the better choice. Its higher recall ensures that most chess masters are correctly identified, even if it means accepting some novices being misclassified as chess masters.

3) Conclusion and Final Recommendation: While S3M1 and S3M10 are both strong candidates, the decision on which model is the best should be based on the real-world application and the relative costs of false positives and false negatives.

To further refine model selection, a cost-sensitive learning approach is recommended, where the misclassification costs are explicitly incorporated into the model's decision-making process. This could involve adjusting the decision threshold or employing a weighted F1-score that reflects the relative importance of precision and recall for the specific application.

In summary, while the selection of the best model can be subjective, it is essential to align the model choice with the specific needs and risks associated with the real-world scenario in which the model will be deployed.

VI. Conclusion

A classification model was created in this study to classify rs-fMRI data of chess players. The objective of the classification model was to separate chess players into distinct categories of Chess Masters and Novices. In 10-fold cross-validation, our developed classification model achieved an accuracy of 96.33%. When compared to prior research, which obtained the highest accuracy of 85.45% utilizing solely rs-fMRI data in chess players, this represents a significant improvement of 10.88%. Additionally, our model demonstrates an 8.33% improvement over the combined approach of rs-fMRI and T1-weighted MRI data, which reached an accuracy of 88%. This accuracy enhancement was achieved through the exclusive use of rs-fMRI data.

The improvement in accuracy can be attributed to several factors. In the data pre-processing phase, a range of window length and stride combinations was analyzed, with the optimal combination selected based on the lowest prediction error of 0.011556. This optimal combination proved effective in estimating dynamic functional connectivities (DFCs) in rs-fMRI. Moreover, after testing various feature selection techniques, Selection 3 emerged as the most effective approach, consistently appearing across all combinations. This selection identifies the most significant discriminative features of rs-fMRI brain scans, facilitating the classification of Chess Masters and Novice players based on their dynamic functional connectivity. In addition, several types of learning algorithms, such as machine learning, deep learning, and ensemble learning, were tested. In total, 11 different learning models and 9 different feature selection strategies were tested. In each combination of learning model and feature selection, hyper-parameter tuning was used to fine-tune the model with the best parameters. Following a comprehensive review of learning models and feature selection methods, the best learning model and feature selection technique were selected.

The next step in this research will involve identifying the specific brain regions associated with key features identified in Selection 3. This will not only provide deeper insights into the neural correlates of chess mastery but also hold potential for making medically significant discoveries. Understanding these brain regions could offer valuable information for cognitive neuroscience and clinical applications, potentially informing interventions and strategies for enhancing cognitive skills and understanding brain function in various contexts.

Data Availability Statement

Publicly available datasets were analyzed in this study. This data can be found here: http://fcon_1000.projects.nitrc.org/indi/pro/wchsu_li_index.html

Ethics Statement

The studies involving human participants were reviewed and approved by The Research Ethics Committee of West China Hospital of Sichuan University. Written informed consent to

participate in this study was provided by the participants' legal guardian/next of kin.

Author Contributions

All authors significantly contributed to the design, analysis, and evaluation of the proposed method and results. All authors reviewed the manuscript and agreed on the final version.

References

[1] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, and K. Ugurbil, "The wu-minn human connectome project: An overview," *NeuroImage*, vol. 80, pp. 62–79, oct 2013.

[2] R. Goebel, A. Roebroeck, D.-S. Kim, and E. Formisano, "Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping," *Magn. Reson. Imaging*, vol. 21, no. 10, pp. 1251–1261, dec 2003.

[3] X. Wen, G. Rangarajan, and M. Ding, "Is granger causality a viable technique for analyzing fmri data?" *PLoS ONE*, vol. 8, no. 7, jul 2013.

[4] H. Lv *et al.*, "Resting-state functional mri: Everything that nonexperts have always wanted to know," *Am. J. Neuroradiol.*, jan 2018.

[5] E. C. A. Hansen, D. Battaglia, A. Spiegler, G. Deco, and V. K. Jirsa, "Functional connectivity dynamics: Modeling the switching behavior of the resting state," *NeuroImage*, vol. 105, pp. 525–535, jan 2015.

[6] P. Delamillieure *et al.*, "The resting state questionnaire: An introspective questionnaire for evaluation of inner experience during the conscious resting state," *Brain Res. Bull.*, vol. 81, no. 6, apr 2010.

[7] D. Mastrovito, "Interactions between resting-state and task-evoked brain activity suggest a different approach to fmri analysis," *J. Neurosci.*, vol. 33, no. 32, aug 2013.

[8] C.-Y. Wee, P.-T. Yap, and D. Shen, "Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks," *CNS Neurosci. Ther.*, vol. 22, no. 3, mar 2016.

[9] R. Hindriks *et al.*, "Can sliding-window correlations reveal dynamic functional connectivity in resting-state fmri?" *NeuroImage*, vol. 127, pp. 242–256, feb 2016.

[10] Sakoğlu, G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun, "A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia," *Magn. Reson. Mater. Phys. Biol. Med.*, vol. 23, no. 5–6, dec 2010.

[11] N. Leonardi *et al.*, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, dec 2013.

[12] Y. Zhou, L. Qiao, W. Li, L. Zhang, and D. Shen, "Simultaneous estimation of low- and high-order functional connectivity for identifying mild cognitive impairment," *Front. Neuroinformatics*, vol. 12, feb 2018.

[13] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The chronnectome: Time-varying connectivity networks as the next frontier in fmri data discovery," *Neuron*, vol. 84, no. 2, oct 2014.

[14] E. Damaraju *et al.*, "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia," *NeuroImage Clin.*, vol. 5, pp. 298–308, 2014.

[15] A. D. Savva, G. D. Mitsis, and G. K. Matsopoulos, "Assessment of dynamic functional connectivity in resting-state fmri using the sliding window technique," *Brain Behav.*, vol. 9, no. 4, apr 2019.

[16] S. Chiang *et al.*, "Bayesian vector autoregressive model for multi-subject effective connectivity inference using multi-modal neuroimaging data: Bayesian multi-modal var model," *Hum. Brain Mapp.*, vol. 38, no. 3, mar 2017.

[17] J. H. Kook, K. A. Vaughn, D. M. DeMaster, L. Ewing-Cobbs, and M. Vannucci, "Bvar-connect: A variational bayes approach to multi-subject vector autoregressive models for inference on brain connectivity networks," *Neuroinformatics*, vol. 19, no. 1, jan 2021.

[18] B. Draganski, C. Gaser, V. Busch, G. Schuierer, U. Bogdahn, and A. May, "Neuroplasticity: changes in grey matter induced by training," *Nature*, vol. 427, pp. 311–312, jan 2004.

[19] C. Gaser and G. Schlaug, "Brain structures differ between musicians and non-musicians," *J. Neurosci.*, vol. 23, pp. 9240–9245, oct 2003.

[20] L. Song *et al.*, "Professional chess expertise modulates whole brain functional connectivity pattern homogeneity and couplings," *Brain Imaging Behav.*, vol. 16, no. 2, pp. 587–595, apr 2022.

[21] X. Duan *et al.*, "Functional organization of intrinsic connectivity networks in chinese-chess experts," *Brain Res.*, vol. 1558, pp. 33–43, apr 2014.

[22] E. Premi *et al.*, "Enhanced dynamic functional connectivity (whole-brain chronnectome) in chess experts," *Sci. Rep.*, vol. 10, apr 2020.

[23] H. RaviPrakash, S. Anwar, and N. Biassou, "Morphometric and functional brain connectivity differentiates chess masters from amateur players," *Front. Neurosci.*, vol. 15, feb 2021.