Machine learning diagnosis of schizophrenia using structural and functional images of the brain

Jingyu Song

Phillips Exeter Academy, NH, USA

jsong6@exeter.edu

Abstract. This study aims to provide earlier, and faster schizophrenia diagnosis based on neurocognitive, structural, and behavioral measures using machine learning. This is because current ways of diagnosis, while accurate, delay patients' ability to seek medical care before observable, lasting symptoms develop and are prone to error and discrimination. To provide diagnosis, we used Linear Support Vector Machine (linear SVM), Random Forest (RF), Multilayer Perceptron (MLP), and k-Nearest Neighbor (kNN), all trained with neurocognitive and behavioral measures combined with either structural or functional MRI data or both of 99 subjects from the OpenNeuro public dataset. 100 iterations of classification were run, and results showed a higher-than-average accuracy for all classifiers using all combinations of parameters, with a highest accuracy of 0.75 using linear SVM trained with behavioral and neurocognitive measures and fMRI data. We found correlations between structural changes in AAL3 brain regions and n-back working memory task performance, noting that the inferior parietal gyrus, right precuneus, supplementary motor area, and the central lateral thalamic nucleus have the highest feature importance. This means that future studies can select these features for further clinical examination or for machine learning diagnosis. We conclude that linear SVM provides the highest average diagnostic accuracy, and that fMRI data often leads to more accurate algorithmic decisions than sMRI data and thus should be weighed more in future studies.

Keywords: machine learning, schizophrenia, working memory, MRI.

1. Introduction

Schizophrenia is a psychiatric disorder with characteristic symptoms such as delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms as described in DSM-V [1]. The average age of onset is from late teens to early 30s, and tends to be earlier in males (from late teens to early 20s), than in females (from late 20s to early 30s) [2-3]. Up to 2023, there are an estimated 24 million people in the world who are diagnosed with schizophrenia [4-8].

Traditional ways of diagnosis rely on identifying two or more symptoms that last for at least one month and eliminating possibilities of other medical conditions such as substance-induced psychosis or depressive and bipolar disorder. The current diagnostic process poses several issues preventing patients from receiving adequate help as early as possible. When lasting, observable symptoms have already arisen, the disorder is often in an advanced stage, making treatment difficult or ineffective. Additionally, clinical diagnosis is prone to error due to the similarity of many schizophrenia symptoms with those of other disorders, and it is also prone to discrimination, especially that against females, blacks, and

individuals with learning disabilities. Therefore, we aim to use machine learning to perform schizophrenia diagnosis using neurocognitive and behavioral measures involved in traditional diagnosis to provide earlier and faster diagnosis and build the foundation for risk factor assessment. Behaviorally, schizophrenia affects the normal functioning of many cognitive abilities, particularly working memory. Working memory falls under short-term memory and is a psychological and cognitive construct that provides temporary storage for information needed to perform complex cognitive tasks such as reasoning, learning, and language comprehension.

In recent years, magnetic resonance imaging (MRI) techniques have been applied to the diagnosis of schizophrenia, with past research revealing the relevance of total brain volume decrease, ventricular enlargement, reduction in hippocampal and thalamic volumes, enlargement in globus pallidus volume, and changes to the neocortical temporal lobe regions [10]. In addition, abnormalities in cortical thickness, and gray and white matter distribution [10-11], as well as changes in connectivity among brain networks and systems, are associated with schizophrenia [12].

To analyze brain imaging data that reflect brain structural changes associated with disorders, researchers have trained machine learning algorithms to analyze MR images of schizophrenia patients versus controls. Examples of commonly used machine learning algorithms include Random Forest (RF), Support Vector Machine (SVM), Ridge, Lasso, Gradient boosting, and logistic regression. Other studies have also used deep learning and convolutional neural networks. This study uses an open dataset and applies linear SVM, RF, multilayer perceptron (MLP), and k-Nearest Neighbor (kNN) to classify participants as schizophrenic or healthy based on demographic, cognitive, brain anatomical, and functional MRI data, and find important features in schizophrenia pathology. Our model expands from preexisting models in that it uses demographic and clinical data of patients that would be incorporated in clinical diagnosis currently but have not been in previous classifiers. We also aim to compare the performance of four classifiers when trained with different combinations of data. This will tell us whether structural or functional data contribute more to accurate diagnosis for each classifier.

2. Materials and Methods

2.1. Participants

We used an open dataset named "Working Memory in Healthy and Schizophrenic Individuals" on openneuro.com [14]. Participants in the study were recruited at Washington University School of Medicine in St. Louis. The open dataset consists of 99 participants (40 female, 59 male), including 23 participants diagnosed with Schizophrenia based on semi-structured interview and Structured Clinical Interview for DSM-IV Axis I Disorders [15]. One participant (sub-03) was removed from the analyses due to erroneous and incomplete measures. Demographic information on study participants can be found in **Table S1**, and more details on participation eligibility criteria can be found in [16].

2.2. Clinical and Cognitive Measures

A series of clinical and cognitive measures were collected from each participant. The clinical measures were based on the Scale for the Assessment of Positive Symptoms (SAPS) and Scale for the Assessment of Negative Symptoms (SANS), whose sub-scores were grouped to yield measures on three composite scores: positive symptom domain, negative symptom domain, and disorganization symptom domain. Details on the computation of these composite scores can be found in the Supplemental Materials of Repovš and Barch (2012). Cognitive abilities were assessed using subscales from the Wechsler Adult Intelligence Scale, Wechsler Abbreviated Scale of Intelligence, Wechsler Memory Scale–Third Edition [17], N-back task [18], Continuous Performance Task, California Verbal Learning Test [19], Trails B [20], category and verbal fluency tasks [21], Wisconsin Card Sort [22]. These sub-scores were grouped and combined into four composite scores, reflecting participants' intelligence, working memory, episodic memory, and executive function. Summary statistics of these clinical and cognitive composite scores are shown in **Table S1**.

2.3. N-back Working Memory Task

All participants performed three N-back tasks, in separate runs of functional MRI scanning, with different levels of demand on their working memory load. The order of these tasks was counterbalanced across participants. In the 0-back task, participants responded to whether the letter shown on the screen was the same as a pre-specified letter ("A" and "X"), where there was little working memory load. In the 1-back task with intermediate working memory load, participants responded to whether the current letter was the same as the immediately preceding letter. In the 2-back task, participants responded to whether the current letter was the same as the letter shown 2 trials before, and thus had the highest working memory load. However, in the OpenNeuro dataset, all trials were coded as "non-target" perhaps due to some coding error. We were therefore not able to determine the actual type of individual experimental trials, i.e., whether the participant should have responded or not. As such, both the behavioral and fMRI analyses were conducted at the condition level, e.g., mean accuracy and fMRI activity during each N-back task for each participant.

2.4. MRI Scanning

Magnetic resonance imaging (MRI) is divided into the subtypes structural MRI (sMRI) and functional MRI (fMRI), with sMRI mainly used to record structural brain information while subjects are not completing a specific cognitive task and fMRI mainly used to record brain regional activity while subjects complete cognitive activities [23]. sMRI is based on translating local differences in water content in the brain into a gray scale through which the outline of the patient's brain can be created. Structural magnetic resonance images provide information regarding the shape, size, and some regional information about the brain. From the resulting images, abnormal tissues with changes in tissue density or composition as well as overall and regional brain structural changes can be detected and analyzed, and this information used in disease diagnosis [24].

fMRI works by measuring small changes in the magnetic field strength caused by paramagnetic deoxyhemoglobin in the blood vessels. Changes in the concentration of deoxyhemoglobin reflect regional consumption of oxygen, which is an indication of local metabolic and neural activity. fMRI has been widely applied to clinical trials such as disease detection and diagnosis, surgical planning, monitoring treatment outcomes, and searching for biomarkers in pharmacologic and training programs.

2.4.1. Scanning Metrics. Scanning was completed at Washington University Medical School using a 3T Tim TRIO Scanner. Structural images were acquired at a spatial resolution (voxel size) of 1 mm × 1 mm × 1 mm. Functional images of the working memory task were acquired in 3 runs of 17.1 min of scanning (13.1 min of task performance) with a repetition time (TR) of 2.5 seconds and a spatial resolution (voxel size) of 4 mm × 4 mm × 4 mm. Each participant completed three functional runs of the working memory tasks. More details can be found in the published study [16].

2.5. sMRI Preprocessing and Analysis

2.5.1. Preprocessing. sMRI data preprocessing was run on MATLAB R2023a with the toolboxes SPM12 and CAT12. The preprocessing process included brain extraction, spatial normalization (to a template brain), and tissue segmentation (gray matter, white matter, and cerebrospinal fluid).

2.5.2. Structural Regions of interest (SROIs). We conducted a literature review of schizophrenia-related or -diagnostic features in brain anatomy. Typical brain anatomical changes include ventricle enlargement, gyrus volume reduction, and gray matter volume and density reduction distributed over several regions, including amygdala, thalamus, hippocampus, and insula [11, 16, 25-47]. We used Neuromorphometrics parcellation of the brain to extract relevant volumetric features (see **Table S2** in Supplementary Materials) [48].

2.6. fMRI Preprocessing and Analysis

2.6.1. Preprocessing. fMRI preprocessing was done by the original authors using fMRIPrep 21.0.1 [49]. fMRI preprocessing procedures include realigning functional images to a reference image to correct for head motion across time, co-registering functional images to the participant's own structural MRI image for better spatial alignment, resampling the functional images into a standard template brain space (MNI 152) for ease of comparison, and spatially smoothing the signal with a Gaussian kernel of 6mm full-width half-maximum to improve signal-to-noise ratio. More details on fMRI preprocessing can be found in the fMRIPrep-generated methods description from "Working Memory in Healthy and Schizophrenic Individuals".

2.6.2. First-level analysis. To separate behavior-related signals from noise, we conducted a first-level analysis [50] using the SPM12 toolbox in MATLAB. A general linear model (GLM) with a regressor that indicates the expected hemodynamic response of individual experimental trials, by convolving the task duration with the canonical hemodynamic response function [50], was fitted to the preprocessed BOLD signal time series. Nuisance regressors of head motion and signal in the white matter and cerebrospinal fluid were also included in the GLM to remove their artifacts. The GLMs yield a regression coefficient (beta) that reflects how closely the observed BOLD signal mirrored the expected activity, thereby providing us a measure of neural activity. We obtained the neural activity estimates for each of the 3 working memory tasks for each participant.

2.6.3. Functional Regions of Interest. To identify the brain regions implicated in working memory, we referred to a meta-analysis of 46 existing fMRI studies on working memory in younger adults (age range comparable to our sample) [46]. The meta-analysis identified multiple regions in the dorsal frontal cortex, inferior parietal lobule, and precuneus (see **Figure 2** and **Table 3** of [46]). A term search of "working memory" on Neurosynth, an automated online meta-analysis, yielded very similar activation patterns across the brain [51]. We parcellated the brain using the Automated Anatomical Labeling 3 (AAL3; [43]) atlas and selected relevant parcels (see **Table S3** for functional regions of interest and **Table 1** below for all features used in classification, including behavioral, sMRI, and fMRI data.)

Table 1. Behavioral, structural, and functional features used in classification

Behavioral	Structural	Functional
gender	third ventricle	precentral gyrus (PreCG)
age	lateral ventricles	left superior frontal gyrus, dorsolateral (SFG)
1-back target score	amygdala	middle frontal gyrus (MFG)
2-back target score	hippocampus	inferior frontal gyrus, opercular part (IFGoperc)
	parahippocampal gyrus	supplementary motor area (SMA)
	superior temporal gyrus	left superior frontal gyrus, medial (SFGmedial)
	fusiform gyrus	insula (INS)
	insula	inferior parietal gyrus, excluding supramarginal and angular gyri (IPG)
	right caudate nucleus	right angular gyrus (ANG)
	middle temporal gyrus	precuneus (PCUN)
	supramarginal gyrus	right ventral lateral nucleus of the thalamus (tAV)
	angular gyrus	
	thalamus	

* If the left or right side is unspecified, then the whole structure is considered.

2.7. Classification Setup

Using the scikit-learn package in Python [52], we implemented four machine learning classification methods: linear SVM, RF, MLP, and kNN. All algorithms were trained to perform a binary classification task: Control vs. Schizophrenic. Linear SVM segregates a given dataset into multiple categories with a single straight line. RF is a random composition of various decision trees through which the composition fits various sub-samples of a dataset and uses averaging to improve predictive accuracy and control over-fitting. kNN classifies new data points based on the classification of the k nearest data points to the given data points. We compared the predictive accuracy and feature importance for each classifier.

2.7.1. Feature Preprocessing. Data cleaning was done on all subjects before building the classifiers. Missing data were imputed with median values. Global standardization of features was done across all participants.

2.7.2. Cross-validation. We employed 100-fold cross-validation. All classification results reported below are the averaged performance from 100 iterations.

2.7.3. Train-Test Split. For each iteration, the dataset was randomly split into a training set and a test set, with the requirement that each test set contained 5 CON and 5 SCZ, leaving 70 CON + 18 SCZ for classifier training. All numeric columns in the training set were standardized (mean=0, sd=1), and the same transformation was applied to the test set.

2.7.4. Data Augmentation. Data augmentation was performed to match the number of SCZ and CON subjects in the training set to resolve class imbalance, which could lead to classifier bias. We made a total of 53 copies of existing SCZ subjects in the training set and added Gaussian noise (mean=0, sd=0.2) to all of their numeric features, such that the final training set contained 71 observations of each class.

3. Results

3.1. Behavioral Analysis

We first conducted behavioral analysis on the demographics, psychopathology, and neurocognitive measures of the patients. This is to rule out any confounds such as educational level and learning abilities that might affect their working memory performance and to ascertain whether the control and the schizophrenic group are similar in terms of age, gender, and racial makeup. We found that there are no statistically significant differences in these measures between the two groups. Additionally, psychopathology and neurocognitive measures validate clinical diagnosis of the subjects because there are statistical differences in performance scores between the two groups. More details can be found in **Table 2**.

Table 2. Behavioral Analysis of CON vs. SCZ individuals

Variable	Mean Difference (SCZ-CON)	t	p	p-Bonferroni
Demographics				
age	1.96	1.90	0.06	0.18
years in school	-0.80	-1.19	0.24	0.72
parents' years in school	-0.29	-0.56	0.58	1.74
Psychopathology				
The raw scores from the clinical measures were first standardized by z scores using the means and standard deviations computed across all subjects who have participated in research studies at the Conte				

Table 2. (continued).

Center for the Neuroscience of Mental Disorders (CCNMD) at Washington University, and the z scores from specific measures were then averaged to yield the below 3 domains.				
positive symptoms domain	1.29	1.66	1.1E-15	3E-15***
negative symptoms domain	1.27	10.25	4.32E-17	1.30E-16 ***
disorganized symptoms domain	0.99	-4.71	9.47E-05	2.84E-4 ***
SANS (Scale for Assessment of Negative Symptoms)				
sans8	1.32	7.55	2.47E-11	1.24E-10 ***
sans13	1.12	6.95	4.49E-10	2.25E-9 ***
sans17	2.38	10.40	2.07E-17	1.04E-16 ***
sans22	1.79	8.12	1.61E-12	8.05E-12 ***
sans25	-0.69	2.47	0.02	0.1
SAPS (Scale for Assessment of Positive Symptoms)				
saps7	1.48	6.98	3.77E-10	1.51E-9 ***
saps20	1.74	9.78	4.45E-16	1.78E-15 ***
saps25	0.42	2.69	0.01	0.05 *
saps34	0.86	6.70	1.42E-09	7.10E-9 ***
Neurocognition				
episodic memory domain	-0.97	-6.79	2.62E-08	1.05E-07 ***
working memory domain	-0.75	-4.69	8.88E-06	3.55E-05 ***
IQ domain	-0.57	-3.05	0.004	0.016 *
executive function domain	-0.63	-4.25	4.86E-05	1.94E-04***
Individual Working Memory Tests				
Continuous Performance Task 4-item d'	-0.64	-3.95	2.34E-04	0.0028 **
Trails B Time to Completion	23.13	4.00	3.62E-04	0.0043 **
Wisconsin Card Sort Perseverative Error Score	10.32	2.59	0.016	0.19
Wechsler Abbreviated Intelligence Scale Vocabulary	-2.04	-3.05	0.004	0.048 *
WAIS Matrix Reasoning	-1.37	-2.27	0.029	0.35
Wechsler Memory Scale Logical Memory	-3.21	-4.78	2.49E-05	3.00E-04 ***
WMS Immediate Recall of Family Pictures	-2.51	-3.63	0.001	0.012 *
WMS Letter-number Sequencing	-2.22	-3.64	0.001	0.012 *
Spatial Span Test	-2.46	-3.58	0.001	0.012 *
Digit Span Test	-1.94	-3.43	0.001	0.012 *

Through our behavioral analysis, we found that demographics were statistically similar for both groups and thus did not pose as a confound. Psychopathology tests confirmed the average accuracy of clinical schizophrenia diagnosis. Schizophrenia patients also showed working memory impairment, though there was no statistical difference between executive function between the two groups.

3.2. N-back Task Performance

We compared the letter n-back task target performance between the two groups and found that there are statistically significant differences for the n-back 1 and 2 tasks but not for the n-back 0 task. This makes sense because the n-back 0 task does not use working memory whereas the second 2 tasks do. This shows that working memory is indeed impaired in schizophrenic individuals. See more details in **Table 3** below.

3.3. Classifier Accuracy

We trained all four classifiers with the following three combinations of data types and recorded the average accuracy of 100 iterations: a) behavioral + structural; b) behavioral + functional; c) behavioral + structural + functional.

Behavioral data included psychopathology and neurocognitive measures because two sample t tests have shown that there are statistical differences between the two group's scores in these areas, as well as subject performance in the letter n-back tasks. Gender was also considered as a factor in classifier training since there is a slight gender imbalance in our samples and because it is considered in clinical diagnosis as well. Structural data included anatomical metrics of regions of interest recorded in structural MRI images of the subjects. These metrics were purely structural such as volume, thickness, matter distribution, and shape because these were recorded when subjects were at rest. When accompanied by psychopathology and neurocognitive measures, working memory task performance was an accurate indicator of each subject's cognitive abilities, especially, their working memory functioning.

The purpose of comparing classifier accuracy across multiple types of training data was to investigate which type(s) of data contributed more to classifier decisions and thus can be used to provide more accurate diagnosis. We found that combinations including functional data usually yielded higher accuracies across multiple classifiers.

We also compared diagnostic accuracy across the four classifiers to examine which one works best when trained with the data available. This suggests to future researchers which algorithm(s) might work better. We found that linear SVM has the best performance on average and achieved our highest accuracy of 0.75 when trained with behavioral and functional data. More information regarding our classifier performance can be found in **Table 3** and **Figure 1**.

Table 3. Comparison of Algorithmic Accuracy based on Structural vs. Functional Data (100 iterations)

	Linear SVM			RF			MLP			kNN		
	Acc.	**t**	**p**	**Acc.**	**t**	**p**	**Acc.**	**t**	**p**	**Acc.**	**t**	**p**
Accuracy based on structural data	0.67	28.34	0.00	0.57	8.66	0.00	0.66	26.23	0.00	0.62	21.0	0.00
Accuracy based on functional data	0.75	17.81	0.00	0.58	17.81	0.00	0.72	14.45	0.00	0.61	8.14	0.00
Accuracy based on both	0.64	10.95	0.00	0.52	10.95	0.01	0.68	12.61	0.00	0.57	4.87	0.00

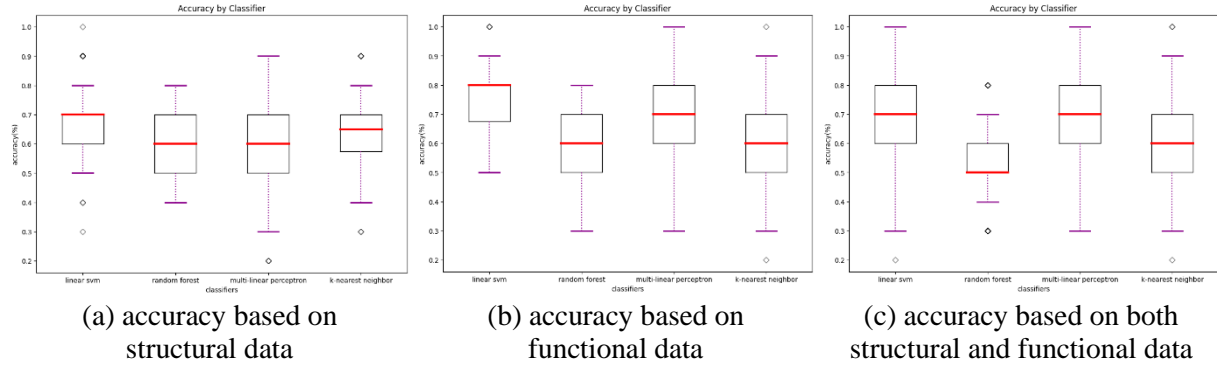| (a) accuracy based on structural data | (b) accuracy based on functional data | (c) accuracy based on both structural and functional data |

Figure 1. Accuracy by classifier

We plotted feature importance to identify the features contributing the most to algorithmic decisions for three classifiers–linear SVM, random forest, and kNN under different data type conditions. We found that there are many overlaps in the top ten features of importance for the three classifiers, which means that the brain regions associated with working memory functioning and other schizophrenia symptoms are largely universal across different algorithms. The result informs physicians of features that should be paid extra attention to during diagnosis. More information about feature importance by classifier can be found in **Figure 2**.

| (a) linear SVM | (b) random forest | (c) kNN |
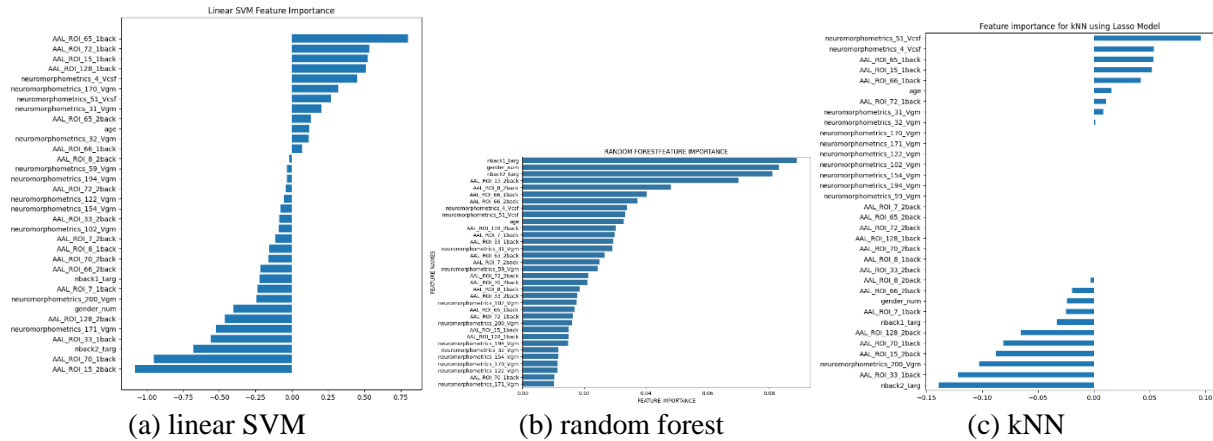
Figure 2. Selected features of importance by classifier including all data types.

4. Discussion

In this study, we demonstrated the possibility of using objective, general-purpose measures of behavior and the brain to diagnose schizophrenia. These behavioral and brain measures are relatively easy to acquire and therefore the implementation may be fairly accessible to the general public. As such, this set of procedures has the potential to aid the diagnosis of schizophrenia early on, when the patients and family experience little symptoms and would not have thought about consulting with psychiatrists.

This study is novel in its comprehensive nature: we both incorporated additional information in training classifiers and compared performance across multiple classifiers. Past research has explored the possibility of using either functional or structural MRI data to detect schizophrenia [53-56], but little research has also included patient background information as would be in a clinical setting. Neither has a lot of research been done to compare the effectiveness of using functional versus structural MRI images in diagnosis, but we found that functional data leads to better classifier performance while the performance based on functional versus both functional and structural data are more similar. This also

indicates that functional data such as fMRI while completing working memory tasks are valuable in clinical diagnosis as well and physicians can consider weighing them more than structural data. The focus on working memory performance itself is relatively novel in the field of machine learning diagnosis of schizophrenia. Additionally, we have only noted one research comparing classifier performance in prediction of hospitalized patients with schizophrenia or other mental health disorders [57]. While covering many classifiers in their research, all of their subjects are either diagnosed with schizophrenia or other mental health disorders. We improved in that area by selecting subjects with or without schizophrenia, which more closely resembles the population in clinical settings.

The results from this study are limited in its small sample size and the number of neurocognitive measures and regions of interest used to train our algorithms. Therefore, we suggest three possible directions for future research. First, we suggest the use of larger sample sizes, when possible, for better classifier training. Since the number of subjects from the public dataset was relatively small, and there was a heavy imbalance of schizophrenic to healthy individuals due to the nature of the previous study, we trained our classifiers using augmented data. However, it will be desirable to use more real subject data both of schizophrenic and healthy individuals for classifier training. Secondly, future studies can try additional behavioral measures, working memory tasks, or other neurocognitive measures to see whether those yield higher diagnosis accuracy. Our results suggest that classification based on neurocognitive measures and fMRI data seems to give higher accuracy, which can help inform future research in choosing their features. Finally, future studies can find a collection of behavioral and neural measures that can be applied to the diagnosis of multiple mental health, neurodegenerative, or psychiatric disorders, such that the utility of the procedures is not limited to schizophrenia.

5. Conclusion

In this study we trained Linear SVM, RF, MLP, and kNN classifiers to diagnose patients as schizophrenic or healthy based on neurocognitive, structural, and functional data. We trained the four classifiers using combinations of afore-mentioned neurocognitive measures and structural or functional MRI data to compare which combination yields higher accuracy. To eliminate confounds in these data, we first conducted a behavioral analysis which yielded that there are no statistical differences in demographic measures between the control and schizophrenic group and the latter does show psychopathology and neurocognitive scores corroborating their clinical diagnosis. In terms of classifier performance, we found that behavioral and functional measures might yield a higher accuracy if linear SVM is used, but that the best combination differs from classifier to classifier. Overall, we found a higher than chance diagnostic accuracy for all classifiers used, with the highest being 75% accuracy using linear SVM based on neurocognitive and functional MRI data. This suggests that algorithms can be effectively used in schizophrenia diagnosis in addition to traditional methods though the technology is not mature enough to replace the traditional methods.

6. Acknowledgements

7. Supplementary materials

7.1. Demographic, clinical, and cognitive measure statistical analysis

We calculated the mean and standard deviation of demographic, clinical and cognitive measures of the control and schizophrenic group to compare the differences between them. Details about these statistics can be found in **Table S1**.

Table S1. Demographic, clinical, and cognitive measures

Measure	Group						
	Healthy Controls (CON)		**Schizophrenic individuals (SCZ)**				
	Mean	**Standard Deviation (SD)**	**Mean**	**Standard Deviation (SD)**			
Age	22.30	4.49	24.26	3.74			
Gender (% male)	56.00%	N/A	73.91%	N/A			
Years in school	13.00	3.09	12.13	1.87			
Parents' years in school	12.93	3.06	14.57	2.48			
Positive symptoms	-0.32	0.30	0.97	1.04			
saps7	0	0	1.48	1.86			
saps20	0	0	1.87	1.39			
saps25	0.15	0.46	0.57	1.08			
saps34	0.01	0.12	0.87	1.10			
Negative symptoms scores	-0.27	0.42	1.00	0.79			
sans8	0.20	0.52	1.52	1.20			
sans13	0.09	0.37	1.22	1.24			
sans17	0.53	0.86	2.91	1.24			
sans22	0.43	0.82	2.22	1.20			
sans25	1.14	1.23	1.83	1.15			
disorganization symptoms	-0.20	0.32	0.79	1.00			
N-BACK PERFORMANCE					**Mean Difference in Accuracy (SCZ-CON)**	**t**	**p-Bonferroni**
0-back accuracy							
target	0.96	0.09	0.92	0.17	-0.04	1.09	0.87
nontarget	0.98	0.10	0.94	0.15			
1-back accuracy							

Table S1. (continued).

target	0.91	0.13	0.79	0.18	-0.12	2.98	0.03 *
nontarget	0.96	0.11	0.90	0.14			
2-back accuracy							
target	0.21	0.17	0.68	0.28	-0.18	2.83	0.03 *
nontarget	0.92	0.85	0.80	0.25			
Neuropsychological assessment							
IQ	-0.30	0.79	-0.87	0.78			
Working memory	0.32	0.68	-0.43	0.64			
Episodic memory	0.17	0.69	-0.80	0.57			
Executive function	0.30	0.54	-0.32	0.82			

Of the 101 participants (99 presented in public dataset) aged 18 or older from whom the original authors collected resting state data, there were 12 who were excluded for poor quality imaging data (4 SCZ, 8 CON), and 15 who were excluded because they did not have N-back data (6 SCZ, 9 CON). Data for subject-03, specifically, had to be discarded in our statistical calculations due to outliers in columns sans17, sans22, sans25, d4prime, LFLUNOVS, and CFLUNOVA, and no data for the remaining neurocognitive and working memory assessments. These outliers are likely due to excessive movement and failure to complete the whole protocol.

7.2. Regions of Interest

We selected structural and functional regions of interest as recorded in **Table S2** and **Table S3**. Both categories were selected based on literature review and all regions of interest were used to train classifiers. Structural regions are given in terms of their neuromorphometrics atlas label. Functional regions are defined based on the Automated Anatomical Labelling atlas 3 (AAL3).

Table S2. Structure MRI Regions of Interest and Correlating AAL3 Labels

Region Name	Neuromorphometrics Label	Structural change
Third ventricle	neuromorphometrics_4_Vcs f	volumetric enlargement
L&R ventricle	neuromorphometrics_51,52_ Vcsf	volumetric enlargement
L&R amygdala	neuromorphometrics_31,32_ Vgm	volumetric reduction (in first episode patients)
L&R hippocampus	neuromorphometrics_47,48_ Vgm (+ Vwm)	volumetric reduction
L&R parahippocampal gyrus	neuromorphometrics_ 170,171_Vgm	volumetric reduction
L&R superior temporal gyrus	neuromorphometrics_200,201_Vgm	volumetric reduction
L&R fusiform gyrus gray matter	neuromorphometrics_ 122,123_Vgm	volumetric reduction
Insula gray matter	neuromorphometrics_ 102, 103, 172, 173_Vgm	volumetric reduction
Right caudate nucleus gray matter	neuromorphometrics_36_Vg m	volumetric excess

Table S2. (continued).

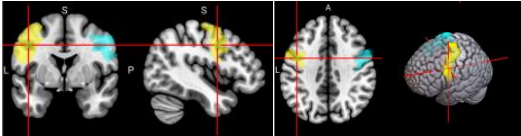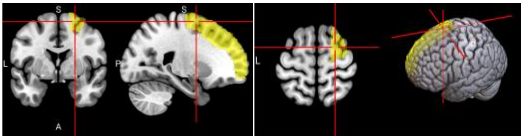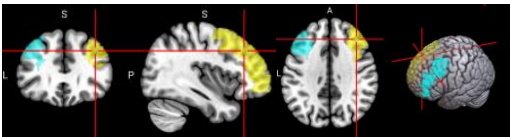| L&R middle temporal gyrus gray matter	neuromorphometrics_ 154,155_Vgm	volumetric excess
L&R supramarginal gyrus	neuromorphometrics_ 194,195_Vgm	volumetric reduction
L&R angular gyrus	neuromorphometrics_ 106,107_Vgm	volumetric increase
L&R thalamus	neuromorphometrics_59,60_ Vgm	volumetric reduction; inward deformation of anterior and posterior regions

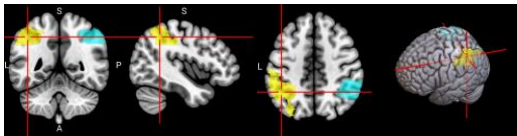**Table S3.** Functional Regions of Interest and CorrespondingAAL3 labels
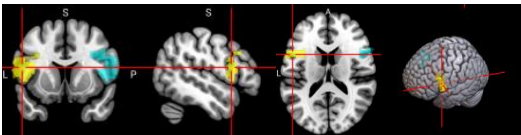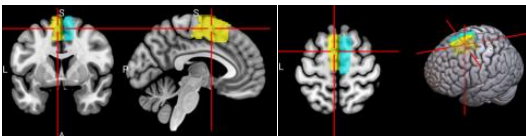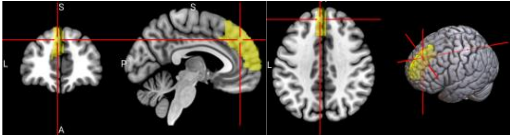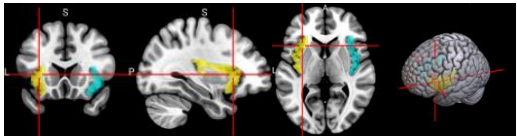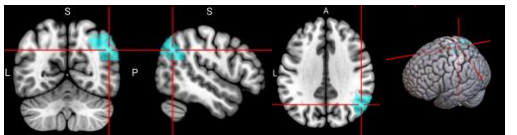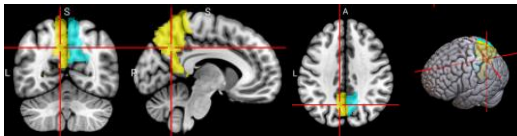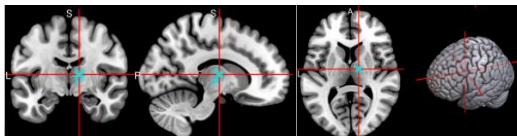
No.	Anatomical Name	Associated Neurocognitive Function	Location in MNI152 Brain Template
1,2	L&R Precentral gyrus (PreCG)	Contains the primary motor cortex, which is responsible for controlling voluntary movements [58].	
4	L Superior frontal gyrus, dorsolateral (SFG)	Left superior frontal gyrus is involved in working memory, spatial processing and other higher cognitive processes [59].	
5,6	L & R Middle frontal gyrus (MFG)	Left middle frontal gyrus is involved in literacy and the right middle frontal gyrus is involved in numeracy [60].	
7,8	L & R Inferior frontal gyrus, opercular part (IFGoperc)	The left inferior frontal gyrus is involved in many cognitive functions including language, executive functioning, social cognition, and inhibition of inappropriate motor responses [61].	
15, 16	L & R Supplementary motor area (SMA)	The SMA plays a role in self-initiated (voluntary) actions and is part of multiple voluntary motor loops [62].	
19	L Superior frontal gyrus, medial (SFGmedial)	The medial superior frontal gyrus is involved in working memory and other higher cognitive functions [59]. The right medial superior frontal gyrus has especially been	

Table S3. (continued).

No.	Anatomical Name	Associated Neurocognitive Function	Location in MNI152 Brain Template
33, 34	L & R Insula (INS)	noted to function in redirecting attention [63]. The insula has many functions including sensorimotor processing, emotion and decision making, attention and salience processing, and speech [64].	
65, 66	L & R Inferior parietal gyrus, excluding supramarginal and angular gyri (IPG)	The inferior parietal gyrus is involved in attentional, semantic, and social cognitive functioning [65].	
70	R Angular gyrus (ANG)	The angular gyrus is the most frequent activated site for semantic processing, and plays an important role in speech, sensory information integration, and semantic processing [66].	
71, 72	L & R Precuneus (PCUN)	The precuneus is involved in recollection and memory, perception, episodic memory, and other cognitive processes [67].	
128	R Ventral Lateral nucleus of the thalamus (tAV)	The ventral lateral nucleus of the thalamus plays an important role in motor control and receiving, integrating, and projecting inputs from cerebellum, striatum, and cortex to the primary motor cortex [68].	

References

[1] American Psychiatric Association, *Schizophrenia Spectrum and Other Psychotic Disorders: DSM-5® Selections*. Washington, D.C., UNITED STATES: American Psychiatric Publishing, 2015. Accessed: Mar. 03, 2024. [Online]. Available: http://ebookcentral.proquest.com/lib/phillipsexeter-ebooks/detail.action?docID=5515120

[2] C. Hollis and J. Rapoport, "Child and adolescent schizophrenia," in *Schizophrenia, 3rd ed*, Hoboken, NJ, US: Wiley Blackwell, 2011, pp. 24–46. doi: 10.1002/9781444327298.ch3.

[3] J. McGrath, S. Saha, D. Chant, and J. Welham, "Schizophrenia: a concise overview of incidence, prevalence, and mortality," *Epidemiol Rev*, vol. 30, pp. 67–76, 2008, doi: 10.1093/epirev/mxn001.

[4] "GBD Results," Institute for Health Metrics and Evaluation. Accessed: Aug. 07, 2023. [Online]. Available: https://vizhub.healthdata.org/gbd-results

[5]   "Mental health systems in selected low- and middle-income countries." Accessed: Aug. 07, 2023. [Online]. Available: https://www.who.int/publications-detail-redirect/9789241547741

[6]   G. Harrison *et al.*, "Recovery from psychotic illness: a 15- and 25-year international follow-up study," *Br J Psychiatry*, vol. 178, pp. 506–517, Jun. 2001, doi: 10.1192/bjp.178.6.506.

[7]   K. Jaeschke, F. Hanna, S. Ali, N. Chowdhary, T. Dua, and F. Charlson, "Global estimates of service coverage for severe mental disorders: findings from the WHO Mental Health Atlas 2017," *Glob Ment Health (Camb)*, vol. 8, p. e27, 2021, doi: 10.1017/gmh.2021.19.

[8]   T. M. Laursen, M. Nordentoft, and P. B. Mortensen, "Excess early mortality in schizophrenia," *Annu Rev Clin Psychol*, vol. 10, pp. 425–448, 2014, doi: 10.1146/annurev-clinpsy-032813-153657.

[9]   A. Baddeley, "Working Memory," *Science*, vol. 255, no. 5044, pp. 556–559, Jan. 1992, doi: 10.1126/science.1736359.

[10]   M. E. Shenton, C. C. Dickey, M. Frumin, and R. W. McCarley, "A review of MRI findings in schizophrenia," *Schizophrenia Research*, vol. 49, no. 1, pp. 1–52, Apr. 2001, doi: 10.1016/S0920-9964(01)00163-3.

[11]   R. W. McCarley *et al.*, "MRI anatomy of schizophrenia," *Biological Psychiatry*, vol. 45, no. 9, pp. 1099–1119, May 1999, doi: 10.1016/S0006-3223(99)00018-9.

[12]   A. Fornito, A. Zalesky, C. Pantelis, and E. T. Bullmore, "Schizophrenia, neuroimaging and connectomics," *NeuroImage*, vol. 62, no. 4, pp. 2296–2314, Oct. 2012, doi: 10.1016/j.neuroimage.2011.12.090.

[13]   A. Meyer-Lindenberg *et al.*, "Evidence for Abnormal Cortical Functional Connectivity During Working Memory in Schizophrenia," *AJP*, vol. 158, no. 11, pp. 1809–1817, Nov. 2001, doi: 10.1176/appi.ajp.158.11.1809.

[14]   "Working memory in healthy and schizophrenic individuals - OpenNeuro." Accessed: Jan. 06, 2024. [Online]. Available: https://openneuro.org/datasets/ds000115/versions/00001

[15]   M. First, R. Spitzer, M. Gibbon, and J. Williams, "Structured clinical interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition," in *(SCID-I/P)*, 2002.

[16]   G. Repovš and D. M. Barch, "Working Memory Related Brain Network Connectivity in Individuals with Schizophrenia and Their Siblings," *Front Hum Neurosci*, vol. 6, p. 137, May 2012, doi: 10.3389/fnhum.2012.00137.

[17]   D. Wechsler, "Wechsler Adult Intelligence Scale--Third Edition." Feb. 11, 2019. doi: 10.1037/t49755-000.

[18]   T. S. Braver, J. D. Cohen, L. E. Nystrom, J. Jonides, E. E. Smith, and D. C. Noll, "A parametric study of prefrontal cortex involvement in human working memory," *Neuroimage*, vol. 5, no. 1, pp. 49–62, Jan. 1997, doi: 10.1006/nimg.1996.0247.

[19]   D. C. Delis, J. Freeland, J. H. Kramer, and E. Kaplan, "California Verbal Learning Test." Jun. 13, 2016. doi: 10.1037/t48844-000.

[20]   R. M. Reitan, "Validity of the Trail Making Test as an Indicator of Organic Brain Damage," *Percept Mot Skills*, vol. 8, no. 3, pp. 271–276, Dec. 1958, doi: 10.2466/pms.1958.8.3.271.

[21]   "Verbal Fluency Test - an overview | ScienceDirect Topics." Accessed: Mar. 06, 2024. [Online]. Available: https://www.sciencedirect.com/topics/nursing-and-health-professions/verbal-fluency-test

[22]   E. A. Berg, "A simple objective technique for measuring flexibility in thinking," *J Gen Psychol*, vol. 39, pp. 15–22, Jul. 1948, doi: 10.1080/00221309.1948.9918159.

[23]   M. Symms, H. R. Jäger, K. Schmierer, and T. A. Yousry, "A review of structural magnetic resonance neuroimaging," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 75, no. 9, pp. 1235–1244, Sep. 2004, doi: 10.1136/jnnp.2003.032714.

[24]   R. Z. Goldstein and N. D. Volkow, "Drug Addiction and Its Underlying Neurobiological Basis: Neuroimaging Evidence for the Involvement of the Frontal Cortex," *Am J Psychiatry*, vol. 159, no. 10, pp. 1642–1652, Oct. 2002.

[25] T. Asami, S. Bouix, T. J. Whitford, M. E. Shenton, D. F. Salisbury, and R. W. McCarley, "Longitudinal loss of gray matter volume in patients with first-episode schizophrenia: DARTEL automated analysis and ROI validation," *NeuroImage*, vol. 59, no. 2, pp. 986–996, Jan. 2012, doi: 10.1016/j.neuroimage.2011.08.066.

[26] K. M. J. Diederen *et al.*, "Deactivation of the Parahippocampal Gyrus Preceding Auditory Hallucinations in Schizophrenia," *AJP*, vol. 167, no. 4, pp. 427–435, Apr. 2010, doi: 10.1176/appi.ajp.2009.09040456.

[27] S. Ehrlich *et al.*, "Associations of White Matter Integrity and Cortical Thickness in Patients With Schizophrenia and Healthy Controls," *Schizophrenia Bulletin*, vol. 40, no. 3, pp. 665–674, May 2014, doi: 10.1093/schbul/sbt056.

[28] M. P. Harms, L. Wang, D. Mamah, D. M. Barch, P. A. Thompson, and J. G. Csernansky, "Thalamic Shape Abnormalities in Individuals with Schizophrenia and Their Nonpsychotic Siblings," *J. Neurosci.*, vol. 27, no. 50, pp. 13835–13842, Dec. 2007, doi: 10.1523/JNEUROSCI.2571-07.2007.

[29] M. J. Kempton, D. Stahl, S. C. R. Williams, and L. E. DeLisi, "Progressive lateral ventricular enlargement in schizophrenia: A meta-analysis of longitudinal MRI studies," *Schizophrenia Research*, vol. 120, no. 1, pp. 54–62, Jul. 2010, doi: 10.1016/j.schres.2010.03.036.

[30] J. S. Kim, C. K. Chung, H. J. Jo, J. M. Lee, and J. S. Kown, "Regional thinning of cerebral cortical thickness in first-episode and chronic schizophrenia," *International Journal of Imaging Systems and Technology*, vol. 22, no. 1, pp. 73–80, 2012, doi: 10.1002/ima.22002.

[31] R. A. Bornstein, S. B. Schwarzkopf, S. C. Olson, and H. A. Nasrallah, "Third-ventricle enlargement and neuropsychological deficit in schizophrenia," *Biological Psychiatry*, vol. 31, no. 9, pp. 954–961, May 1992, doi: 10.1016/0006-3223(92)90121-F.

[32] T. J. Bouchard, D. T. Lykken, M. McGue, N. L. Segal, and A. Tellegen, "Sources of human psychological differences: the Minnesota Study of Twins Reared Apart," *Science*, vol. 250, no. 4978, pp. 223–228, Oct. 1990, doi: 10.1126/science.2218526.

[33] R. Brown *et al.*, "Postmortem Evidence of Structural Brain Changes in Schizophrenia: Differences in Brain Weight, Temporal Horn Area, and Parahippocampal Gyrus Compared With Affective Disorder," *Archives of General Psychiatry*, vol. 43, no. 1, pp. 36–42, Jan. 1986, doi: 10.1001/archpsyc.1986.01800010038005.

[34] D. S. Cheng *et al.*, "Schizophrenia Classification Using Regions of Interest in Brain MRI".

[35] T. J. Crow, "Temporal Lobe Asymmetries as the Key to the Etiology of Schizophrenia," *Schizophrenia Bulletin*, vol. 16, no. 3, pp. 433–443, Jan. 1990, doi: 10.1093/schbul/16.3.433.

[36] L. E. DeLisi, M. Sakuma, W. Tew, M. Kushner, A. L. Hoff, and R. Grimson, "Schizophrenia as a chronic active brain process: a study of progressive brain structural change subsequent to the onset of schizophrenia," *Psychiatry Research: Neuroimaging*, vol. 74, no. 3, pp. 129–140, Jul. 1997, doi: 10.1016/S0925-4927(97)00012-7.

[37] B.-B. Koo, J.-M. Lee, J. S. Kim, J. S. Park, J. S. Kwon, and S. I. Kim, "Assessing spatial probabilistic distributional differences in the common space between schizophrenics and normal controls based on a novel automated probabilistic pattern analysis method," *International Journal of Imaging Systems and Technology*, vol. 18, no. 5–6, pp. 310–324, 2008, doi: 10.1002/ima.20167.

[38] G. R. Kuperberg *et al.*, "Regionally Localized Thinning of the Cerebral Cortex in Schizophrenia," *Archives of General Psychiatry*, vol. 60, no. 9, pp. 878–888, Sep. 2003, doi: 10.1001/archpsyc.60.9.878.

[39] S. M. Lawrie, H. C. Whalley, D. E. Job, and E. C. Johnstone, "Structural and Functional Abnormalities of the Amygdala in Schizophrenia," *Annals of the New York Academy of Sciences*, vol. 985, no. 1, pp. 445–460, 2003, doi: 10.1111/j.1749-6632.2003.tb07099.x.

[40] M. Murakami *et al.*, "Cortical thickness, gray matter volume, and white matter anisotropy and diffusivity in schizophrenia," *Neuroradiology*, vol. 53, no. 11, pp. 859–866, Nov. 2011, doi: 10.1007/s00234-010-0830-2.

[41] O. G. O'Daly, S. Frangou, X. Chitnis, and S. S. Shergill, "Brain structural changes in schizophrenia patients with persistent hallucinations," *Psychiatry Research: Neuroimaging*, vol. 156, no. 1, pp. 15–21, Oct. 2007, doi: 10.1016/j.pscychresns.2007.03.001.

[42] D. G. C. Owens, E. C. Johnstone, T. J. Crow, C. D. Frith, J. R. Jagoe, and L. Kreel, "Lateral ventricular size in schizophrenia: relationship to the disease process and its clinical manifestations," *Psychological Medicine*, vol. 15, no. 1, pp. 27–41, Feb. 1985, doi: 10.1017/S0033291700020900.

[43] E. T. Rolls, C.-C. Huang, C.-P. Lin, J. Feng, and M. Joliot, "Automated anatomical labelling atlas 3," *NeuroImage*, vol. 206, p. 116189, Feb. 2020, doi: 10.1016/j.neuroimage.2019.116189.

[44] J. Sun, J. J. Maller, L. Guo, and P. B. Fitzgerald, "Superior temporal gyrus volume change in schizophrenia: A review on Region of Interest volumetric studies," *Brain Research Reviews*, vol. 61, no. 1, pp. 14–32, Jun. 2009, doi: 10.1016/j.brainresrev.2009.03.004.

[45] Y. Takayanagi *et al.*, "Differentiation of first-episode schizophrenia patients from healthy controls using ROI-based multiple structural brain variables," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 34, no. 1, pp. 10–17, Feb. 2010, doi: 10.1016/j.pnpbp.2009.09.004.

[46] Z. A. Yaple, W. D. Stevens, and M. Arsalidou, "Meta-analyses of the n-back working memory task: fMRI evidence of age-related changes in prefrontal cortex involvement across the adult lifespan," *NeuroImage*, vol. 196, pp. 16–31, Aug. 2019, doi: 10.1016/j.neuroimage.2019.03.074.

[47] F. Zheng, C. Li, D. Zhang, D. Cui, Z. Wang, and J. Qiu, "Study on the sub-regions volume of hippocampus and amygdala in schizophrenia," *Quant Imaging Med Surg*, vol. 9, no. 6, pp. 1025–1036, Jun. 2019, doi: 10.21037/qims.2019.05.21.

[48] "Neuromorphometrics, Inc.: Segmentation Protocols." Accessed: Jan. 08, 2024. [Online]. Available: https://neuromorphometrics.com/Seg/

[49] O. Esteban *et al.*, "fMRIPrep: a robust preprocessing pipeline for functional MRI," *Nat Methods*, vol. 16, no. 1, Art. no. 1, Jan. 2019, doi: 10.1038/s41592-018-0235-4.

[50] M. Monti, "Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach," *Frontiers in Human Neuroscience*, vol. 5, 2011, Accessed: Jan. 08, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2011.00028

[51] "Neurosynth: working memory." Accessed: Jan. 06, 2024. [Online]. Available: https://neurosynth.org/analyses/terms/working%20memory/

[52] "scikit-learn: machine learning in Python — scikit-learn 1.3.2 documentation." Accessed: Jan. 08, 2024. [Online]. Available: https://scikit-learn.org/stable/

[53] Z. Chen *et al.*, "Detecting Abnormal Brain Regions in Schizophrenia Using Structural MRI via Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2020, pp. 1–13, Mar. 2020, doi: 10.1155/2020/6405930.

[54] J. Oh, B.-L. Oh, K.-U. Lee, J.-H. Chae, and K. Yun, "Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm," *Front. Psychiatry*, vol. 11, p. 16, Feb. 2020, doi: 10.3389/fpsyt.2020.00016.

[55] J. Zhang *et al.*, "DETECTING SCHIZOPHRENIA WITH 3D STRUCTURAL BRAIN MRI USING DEEP LEARNING".

[56] Y. Zheng, H. Tong, T. Zhao, X. Guo, H. Xu, and R. Yang, "Support vector machine classification combined with multimodal magnetic resonance imaging in detection of patients with schizophrenia," *IET image process*, vol. 14, no. 11, pp. 2610–2615, Sep. 2020, doi: 10.1049/iet-ipr.2019.1108.

[57] S. Góngora Alonso, G. Marques, D. Agarwal, I. De la Torre Díez, and M. Franco-Martín, "Comparison of Machine Learning Algorithms in the Prediction of Hospitalized Patients with Schizophrenia," *Sensors*, vol. 22, no. 7, Art. no. 7, Jan. 2022, doi: 10.3390/s22072517.

[58] L. Banker and P. Tadi, "Neuroanatomy, Precentral Gyrus," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2024. Accessed: Feb. 24, 2024. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK544218/

[59] F. du Boisgueheneuc *et al.*, "Functions of the left superior frontal gyrus in humans: a lesion study," *Brain*, vol. 129, no. Pt 12, pp. 3315–3328, Dec. 2006, doi: 10.1093/brain/awl244.

[60] R. M. El-Baba and M. P. Schury, "Neuroanatomy, Frontal Cortex," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2024. Accessed: Feb. 24, 2024. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK554483/

[61] B. Ishkhanyan *et al.*, "Anterior and Posterior Left Inferior Frontal Gyrus Contribute to the Implementation of Grammatical Determiners During Language Production," *Front Psychol*, vol. 11, p. 685, Apr. 2020, doi: 10.3389/fpsyg.2020.00685.

[62] P. Nachev, C. Kennard, and M. Husain, "Functional role of the supplementary and pre-supplementary motor areas," *Nat Rev Neurosci*, vol. 9, no. 11, Art. no. 11, Nov. 2008, doi: 10.1038/nrn2478.

[63] S. Japee, K. Holiday, M. D. Satyshur, I. Mukai, and L. G. Ungerleider, "A role of right middle frontal gyrus in reorienting of attention: a case study," *Frontiers in Systems Neuroscience*, vol. 9, 2015, Accessed: Feb. 24, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnsys.2015.00023

[64] L. Q. Uddin, J. S. Nomi, B. Hebert-Seropian, J. Ghaziri, and O. Boucher, "Structure and function of the human insula," *J Clin Neurophysiol*, vol. 34, no. 4, pp. 300–306, Jul. 2017, doi: 10.1097/WNP.0000000000000377.

[65] O. Numssen, D. Bzdok, and G. Hartwigsen, "Functional specialization within the inferior parietal lobes across cognitive domains," *eLife*, vol. 10, p. e63591, doi: 10.7554/eLife.63591.

[66] M. L. Seghier, "The Angular Gyrus," *Neuroscientist*, vol. 19, no. 1, pp. 43–61, Feb. 2013, doi: 10.1177/1073858412440596.

[67] A. E. Cavanna and M. R. Trimble, "The precuneus: a review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. Pt 3, pp. 564–583, Mar. 2006, doi: 10.1093/brain/awl004.

[68] A. P. TLAMSA and J. C. BRUMBERG, "Organization and morphology of thalamocortical neurons of mouse ventral lateral thalamus," *Somatosens Mot Res*, vol. 27, no. 1, pp. 34–43, 2010, doi: 10.3109/08990221003646736.