

Article

Four-Class Classification of Neuropsychiatric Disorders by Use of Functional Near-Infrared Spectroscopy Derived Biomarkers

Sinem Burcu Erdoğan *  and Gülnaz Yükselen 

Department of Biomedical Engineering, Acıbadem Mehmet Ali Aydınlar University, Istanbul 34684, Turkey; gulnaz.yukselen@acibadem.edu.tr

* Correspondence: sinem.erdogan@acibadem.edu.tr

Abstract: Diagnosis of most neuropsychiatric disorders relies on subjective measures, which makes the reliability of final clinical decisions questionable. The aim of this study was to propose a machine learning-based classification approach for objective diagnosis of three disorders of neuropsychiatric or neurological origin with functional near-infrared spectroscopy (fNIRS) derived biomarkers. Thirteen healthy adolescents and sixty-seven patients who were clinically diagnosed with migraine, obsessive compulsive disorder, or schizophrenia performed a Stroop task, while prefrontal cortex hemodynamics were monitored with fNIRS. Hemodynamic and cognitive features were extracted for training three supervised learning algorithms (naïve bayes (NB), linear discriminant analysis (LDA), and support vector machines (SVM)). The performance of each algorithm in correctly predicting the class of each participant across the four classes was tested with ten runs of a ten-fold cross-validation procedure. All algorithms achieved four-class classification performances with accuracies above 81% and specificities above 94%. SVM had the highest performance in terms of accuracy ($85.1 \pm 1.77\%$), sensitivity ($84 \pm 1.7\%$), specificity ($95 \pm 0.5\%$), precision ($86 \pm 1.6\%$), and F1-score ($85 \pm 1.7\%$). fNIRS-derived features have no subjective report bias when used for automated classification purposes. The presented methodology might have significant potential for assisting in the objective diagnosis of neuropsychiatric disorders associated with frontal lobe dysfunction.

Keywords: fNIRS; BCI; classification; schizophrenia; obsessive compulsive disorder; migraine; Stroop test



Citation: Erdoğan, S.B.; Yükselen, G. Four-Class Classification of Neuropsychiatric Disorders by Use of Functional Near-Infrared Spectroscopy Derived Biomarkers. *Sensors* **2022**, *22*, 5407. <https://doi.org/10.3390/s22145407>

Academic Editor: Rebecca Re

Received: 23 May 2022

Accepted: 22 June 2022

Published: 20 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In clinical practice, the majority of neuropsychiatric disorders are diagnosed with a clinician-dependent interpretation of patient information, which is obtained through a variety of subjectively biased sources such as clinical interviews, self-reports, observational data, and behavioral measures [1–5]. The potential of introducing subjectivity during both interpretation and acquisition of these diagnostic measures may have a prominent impact on the final clinical decision and highlights the critical need for developing accurate, objective, and reliable clinical decision support systems. Such decision support systems should ideally analyze objective and quantitative measures of the distinct characteristics of the neurobiological changes that are gradually induced by each neuropsychiatric disorder [6,7].

Within this context, various functional brain imaging modalities, such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG), and functional near-infrared spectroscopy (fNIRS), have been utilized for characterizing the neurobiological underpinnings of a variety of neuropsychiatric disorders [8–10]. Among these modalities, fNIRS systems have stepped forward for extracting informative, neuronally induced hemodynamic markers of cognition during altered brain states in naturalistic settings [11,12]. Consequently, fNIRS systems have also received increasing interest in the field of psychiatry for assisting diagnosis, prognosis, and follow-up of treatment procedures thanks to their: (1) portability, (2) non-invasive nature, (3) modest

equipment size, (4) robustness to electrogenic or motion artifacts, (5) low operating cost, (6) quick set-up time and calibration, (7) ability to collect biological information at any desired frequency and duration, and (8) ease of application in ecologically valid settings to a broad range of patient populations involving children and elderly adults [13].

Indeed, recent studies have presented compelling evidence that a wide variety of neuropsychiatric disorders can be characterized by functional alterations in the hemodynamic activity of the prefrontal cortex (PFC), which can be detected with fNIRS [13]. For instance, hypoactivation in frontal lobe regions has been detected in patients with schizophrenia (SCZ) and major depressive disorder (MDD) during verbal fluency tasks when compared to their healthy counterparts [5,14]. Similarly, hyper- and hypo-connectivity between different brain regions during resting state have been identified in patients with schizophrenia (SCZ) [15–19] and major depressive disorder (MDD) [8,9,14], while decreased cerebral blood flow in bilateral symmetric regions of the inferior PFC has been detected in patients with obsessive compulsive disorder (OCD) when compared to their healthy counterparts [20,21]. PFC dysfunction in the form of hypo- or hyper-connectivity during resting state or hypoactivation during various cognitive tests (e.g., Stroop and verbal fluency test) has been extensively observed and reported in patient groups diagnosed with a variety of major neuropsychiatric disorders, which include SCZ, MDD, bipolar disorder (BD), post-traumatic stress disorder (PTSD), and attention deficiency and hyperactivity disorder (ADHD). Results from meta-analysis studies indicated that the topographical distributions of functional abnormalities observed in these patient groups are likely to have disorder-specific patterns [14]. Overall, these studies have highlighted the potential of exploring PFC-based neurofunctional features as objective and distinctive biomarkers of various major neuropsychiatric disorder states. They also showed that information from practical and preferably field-deployable cerebral physiology monitoring tools such as fNIRS systems can quantify and parameterize abnormalities in frontal lobe function and may have a great potential for assisting in the objective diagnosis and classification of major psychiatric disorders which, in most cases, have overlapping behavioral symptoms across each other and are difficult to distinguish when decisions are based solely on observation, self-report, interview, and/or rating scales.

Considering the critical demand to integrate more objective measures of neurophysiological alterations into diagnostic clinical decision processes, the presented study aimed to assess the feasibility and applicability of an fNIRS-based automated classification approach for accurate prediction and objective identification of the presence of three distinct neuropsychiatric or neurological disorder states which are known to induce alterations in frontal lobe function. In our recent work, we demonstrated the feasibility and applicability of an fNIRS-assisted automated classification approach for accurate prediction of the presence of impulsivity in adolescents [22]. More specifically, our results suggested that training computationally efficient supervised learning algorithms with informative features obtained from clinical, behavioral, and fNIRS-derived hemodynamic measures could serve as a decision support system for recognizing the presence of impulsivity in individuals. However, the clinical features included in the feature sets still had the potential to present subjective bias when used for algorithm training purposes because there always existed some probability that the subjects could provide false reports in the clinical interviews.

Based on the promising performance of integrating fNIRS-derived features and clinical features for recognizing the presence of impulsivity in our recent work, the objective of this study was to introduce a more reliable, machine learning-based classification approach for correct identification of the presence of three distinct neuropsychiatric and/or neurological disorder states. The proposed machine learning-based classification approach involved training three supervised learning algorithms with (i) fNIRS-derived informative biomarkers only and (ii) a combination of fNIRS-derived biomarkers and performance measures obtained during a cognitive test, named the Stroop task. We tested the feasibility of the proposed approach with three distinct supervised learning algorithms and by extending our classification problem to include four classes of subjects. The ultimate goal

was to demonstrate the feasibility of an fNIRS-based automated classification methodology for predicting the presence of a neuropsychiatric disease, where input features are of pure biological origin and can be derived non-invasively in naturalistic settings by use of ergonomic fNIRS probes. For this purpose, hemodynamic information obtained from concurrent fNIRS recordings during a Stroop task was processed to extract global efficiency metrics which are indicative of the strength of functional connectivity among different PFC regions. The efficacy of training three distinct supervised machine learning algorithms, namely naive Bayes (NB), linear discriminant analysis (LDA), and support vector machines (SVM), with (a) fNIRS-derived neuronally induced biomarkers and (b) a combination of fNIRS-derived biomarkers and cognitive performance measures obtained during the Stroop task, was evaluated. The performance metrics of possible combinations of each classification algorithm and feature set combination were assessed by whether each subject was correctly labeled among the four classes, which included healthy controls (HC), patients diagnosed with migraine without aura (MIG), schizophrenia (SCZ), and obsessive compulsive disorder (OCD).

Our study presents the following novelties with respect to the current literature: To date, there have been no studies that attempted to identify the presence of a neuropsychiatric or neurological disorder by use of a four-class automated classification scheme based on a combination of fNIRS-derived neuronally induced metrics obtained during a neuropsychological test and supervised machine learning methods. The efficacy of an automated classification approach which aims to correctly label a neuropsychiatric disorder into one of four categories has not been evaluated before with structural or functional neuroimaging measures. The efficacy of combining fNIRS-derived global efficiency metrics of the PFC as sole informative features of a neuropsychiatric and/or neurological disorder with supervised learning methods has also not been evaluated before.

2. Materials and Methods

2.1. Subjects

In this study, 13 healthy control (HC) subjects (6 female (F), mean age 26), 20 migraine (MIG) patients without aura (12 F, mean age 27), 26 patients with obsessive compulsive disorder (OCD) (11 F, mean age 29), and 21 schizophrenia (SCZ) patients (10 F, mean age 28) participated. Each subject provided informed consent before participating in the experiment. The study protocol was approved by the Ethics Committee of Pamukkale University, Denizli, Turkey. All experiments were conducted according to the latest Declaration of Helsinki. Parts of these datasets have been utilized in previous works performed by our group and coworkers [23–29].

2.2. Experimental Protocol

During the experiments, subjects sat on a comfortable chair in front of a computer screen which was placed approximately 1 m away from their eyes. All experiments were carried out in a dimly illuminated, silent room. The experimental protocol was briefly explained to each subject prior to the onset of each experiment. They were requested to sit relaxed and refrain from moving their head during the fNIRS recordings. During the experiment, their task was to carefully complete a color–word Stroop task which was adapted to Turkish from a pioneer protocol proposed by Zysset et al. [30]. Each experiment began with 30 s of a baseline recording followed by presentation of alternating blocks of 3 stimulus conditions which consisted of neutral (N), congruent (C), and incongruent (IC) stimuli (Figures 1 and 2). There was a total of 5 stimulus blocks for each condition (i.e., N, C, IC) and all task blocks were presented in a randomized order that changed for every experimental session. Each stimulus block consisted of 6 different trials of the same condition. Within a block, each trial appeared on the screen for 2.5 s followed by a 4 s blank screen. Task blocks were separated with 20 s periods of rest (Figure 2).

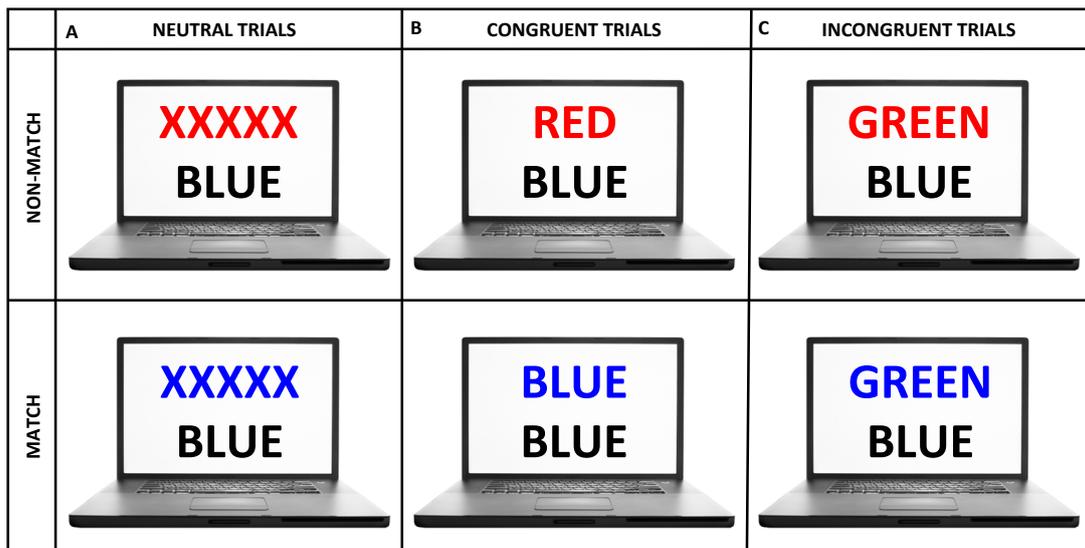


Figure 1. Types of stimuli that were presented within the color–word Stroop experiment. Samples of trial presentations are schematically represented for match (top row) and non-match conditions (bottom row) of (A) neutral, (B) congruent, and (C) incongruent stimuli.

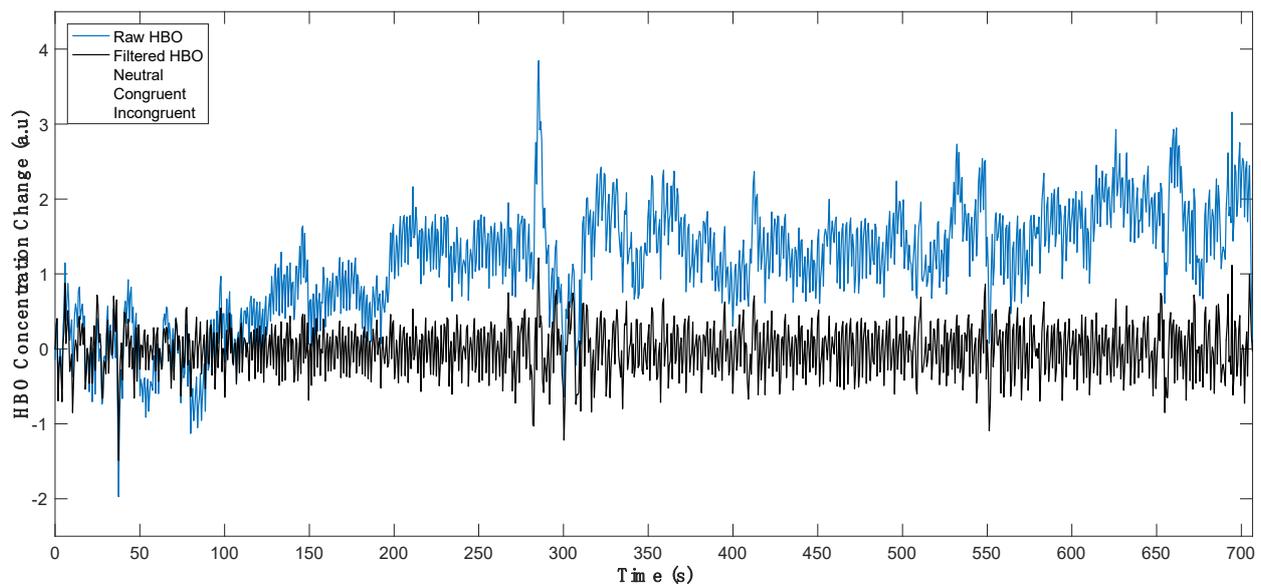


Figure 2. Experimental protocol depicting stimuli timing for neutral, congruent, and incongruent stimuli blocks in a sample session. HBO time series from a representative channel of a subject are plotted before (black) and after an 8th order Butterworth high-pass filter is applied with a cut-off frequency of 0.009 Hz (blue). a.u.: arbitrary units.

During each stimulus presentation, two rows of letters were displayed on the screen. The task was to evaluate whether the color of the letters displayed at the top row matched with the meaning of the word displayed at the bottom row. Subjects were asked to press the left button of the mouse if the color of the upper row letters matched with the meaning of the bottom row word. These cases were called match cases (Figure 1, top panel). They were asked to press the right button if the color of the upper row letters did not match with the meaning of the bottom row word for non-match cases (Figure 1, bottom panel). The letters in all trials were printed in one of four basic colors, which were yellow, red, blue, or green. In N trials, top row letters were written in yellow, red, blue, or green but did not form a meaningful word, and a color name was typed in black on the bottom row. For C trials, a word with the meaning of a color was typed in the same congruent color in the

2.4. Data Analysis

2.4.1. Processing of fNIRS Signals

The fNIRS-HBO signals are composed of neuronally and systemically induced hemodynamic components which are intermixed with each other over a broad range of frequencies. The neuronally induced hemodynamic variations in the HBO signal are caused by both spontaneous and task-related neuronal activity, while the systemic physiological activity-related hemodynamic components have multiple origins, which include variations in heartbeat, respiration, blood pressure, and vascular tone. Hence, prior to obtaining correlation-based functional connectivity metrics between HBO signals of different channel pairs, the impact of common, global systemic effects of non-neuronal origin inherent in both channel data had to be reduced. Such a procedure is necessary to isolate the extent of correlation caused by only neuronally induced hemodynamic effects, since common physiological effects of non-neuronal origin could inflate the correlation between signals of channel pairs. A partial correlation approach was adapted from the works of Akin [28] and Akin [29] to reduce the impact of common systemic interference to Pearson's correlation coefficients calculated between HBO signals of each channel pair. Similar to these works, HBO signals of all channels were initially high-pass-filtered with a cut-off frequency of 0.009 Hz using an 8th order Butterworth filter. The high-pass-filtered HBO signals were then averaged to have a single global signal regressor, which was utilized as the partial regressor for modeling and removing the impact of common systemic noise from the correlations between each channel pair in the subsequent step of the analysis [22,24–29].

Time traces used for computing the correlation between each channel pair were obtained as follows. For each channel, HBO signals corresponding to each stimulus block were truncated from the onset to the end of that block. These time segments were then concatenated in time to obtain a single task-related HBO signal for each channel of each subject. Similarly, the partial correlation regressor was obtained by truncating and concatenating the time segments belonging to all task blocks in the global signal regressor. Then, 16-by-16 partial correlation (PC)-corrected functional connectivity (FC) matrices for each subject were generated after removing the impact of this partial correlation regressor [22,29].

2.4.2. Computation of Cognitive Quotient and Global Efficiency Features

Two groups of features were extracted from the behavioral and hemodynamic data obtained during the Stroop task. Similar to our previous work [22], the behavioral performance was quantified with a feature named the cognitive quotient (CQ), which could be considered as a generalized cognitive performance indicator of each subject during the Stroop task. The accuracy and reaction time metrics obtained from all trials of the Stroop experiment were fused in this single metric by dividing the overall accuracy performance (i.e., percentage of correct answers over all trials) with the average reaction time for all trials.

Regarding the hemodynamic features, a relatively novel functional connectivity metric called global efficiency (GE) was obtained from the 16-by-16 partial correlation-corrected FC matrices obtained for each subject. The GE metric was obtained from a graph theoretical network analysis approach, and its efficacy in demonstrating the degree of connectedness and information transfer between cortical regions during various cognitive tasks has been shown in previous studies [22,27–29].

After the partial correlation-corrected FC matrices were obtained for each subject, these matrices were decomposed into two matrices, which represent the degree of connectedness of the default mode (DM) and the cognitive mode (CM) networks of the brain. This decomposition was established by applying principal component analysis to the FC matrix, the details of which are extensively explained in the recent work of Akin [29]. Briefly, principal component (PC) decomposition was applied to the 16 by 16 FC matrices and the weights of the PCs were thresholded using an optimization procedure described in [29]. The DM and CM components of the FC matrices were reconstructed by weighting and summing the PC regressors that had weights below and above the threshold, separately.

GE values of the DM and the CM components were computed separately for each subject by using the formula of Latora and Marchiori [45]. The GE feature for the DM network was named GE_{dm} , and similarly, the GE feature for the CM was named GE_{cm} (Figure 4).

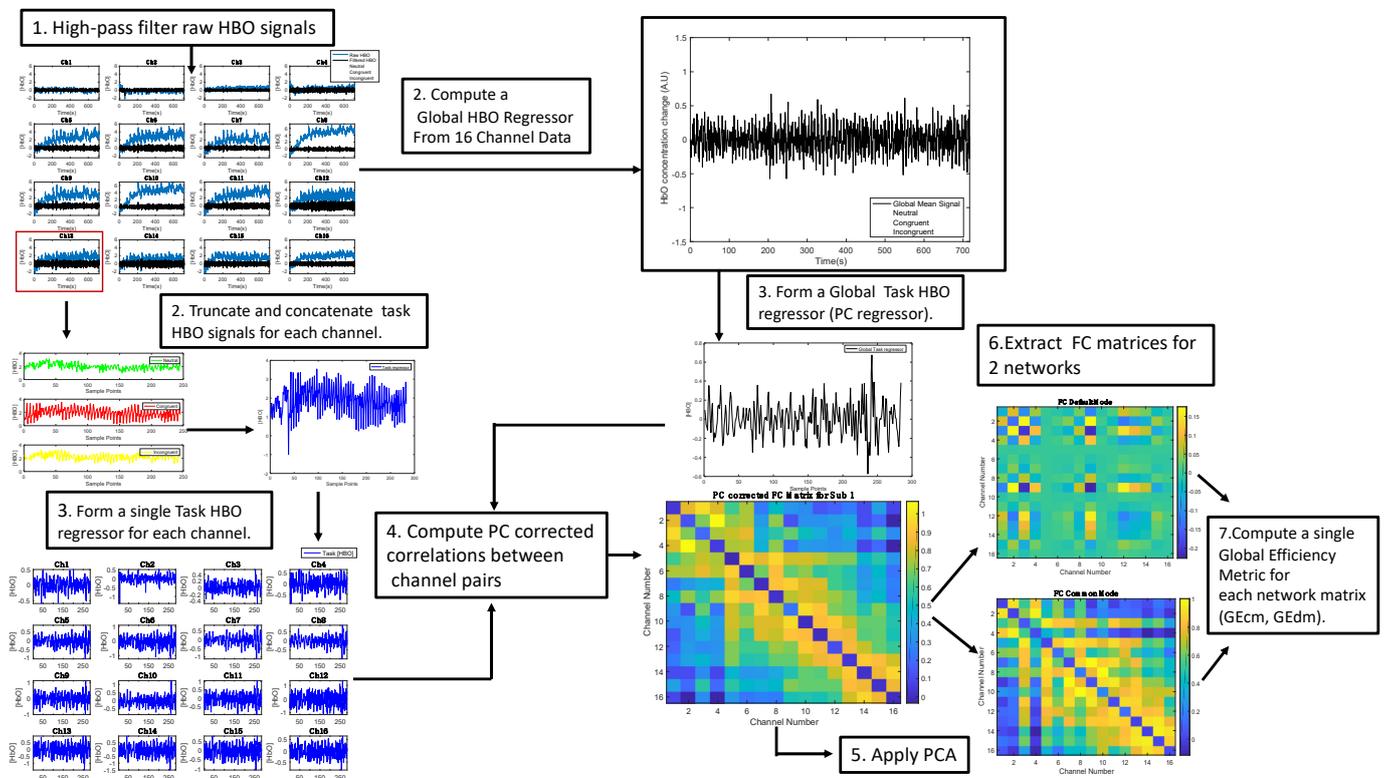


Figure 4. Pipeline for extracting fNIRS derived features.

2.4.3. Classification Methods

The feasibility of training fNIRS-derived GE features with machine learning classifiers for correct identification of the presence of a disorder in each subject was evaluated and compared for three distinct algorithms. These algorithms were naive Bayes (NB), linear discriminant analysis (LDA), and support vector machines (SVM). These algorithms were selected for several reasons: (1) They have been shown to perform well with small sample sizes ($n < 200$) [5,15,22,46–57]. (2) Their computational cost is low. (3) They have performed successfully in a variety of classification problems where fNIRS features extracted during cognitive and motor tasks were utilized [25,26,28–31]. (4) Their good performance for classification at the single subject level has been reported for previous neuropsychiatry studies with similar sample sizes, but a lower number of classes [1,5,15,58–62]. The mathematical architecture of these algorithms has been extensively explained in previous work performed by our group [22,46] and others [52–56].

Each classification algorithm was constructed by using the libraries of the WEKA platform (version 3.8.5) [63]. The sequential minimal optimization (SMO) algorithm was utilized for training the SVM classifier [64]. SMO was run with the Pearson VII universal kernel [65], also known as the PUK kernel. To avoid overfitting, the regularization parameter (C) of SMO and PUK kernel parameters (i.e., omega (ω) and sigma (σ)) was optimized by maximizing the accuracy with a grid-search procedure. Assigning $C = 10$ and $\omega = \sigma = 1$ yielded the best results. LDA and NB classifiers were constructed with the default parameters implemented in the WEKA software. A brief flowchart of the processing pipeline is demonstrated in Figure 5.

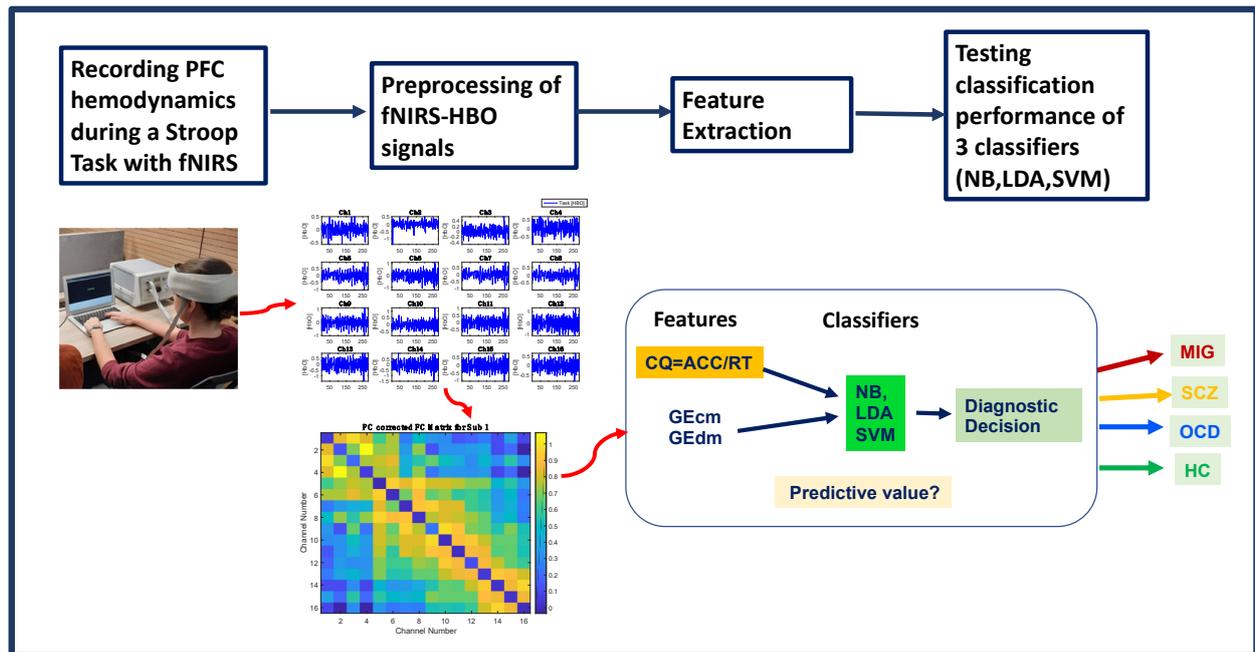


Figure 5. Flowchart of the system design.

2.4.4. Performance Evaluation

To evaluate and compare the classification performances of each algorithm, accuracy, precision, sensitivity, specificity, and F1-score were calculated through a comparison of the actual and predicted labels of test data [61]. For each algorithm, performance metrics were obtained after 10 runs of a 10-fold cross-validation (C.V) procedure, where in each run, 1/10th of the subject data were separated for testing the algorithm and the remainder of subject data were used for training, and this procedure was repeated 10 times. For each performance metric, the mean scores across all runs and their standard deviation were computed (Tables 1 and 2). This procedure was conducted for cases when each algorithm was trained with (i) fNIRS only features (i.e., GE_{cm} and GE_{dm}) and (ii) a combination of fNIRS-derived features (i.e., GE_{cm} , GE_{dm}) and a behavioral feature (i.e., CQ). All features were computed for each subject separately.

Table 1. Four-class classification performances of NB, LDA, and SVM algorithms when trained with fNIRS only features (i.e., GE_{cm} and GE_{dm}). Each performance metric is represented in percentages (%) as the mean value across all runs \pm standard deviation of the mean.

Method	Accuracy	Precision	Recall	Specificity	F1-Score
NB	81.77 \pm 1.06	82.1 \pm 1	81 \pm 0.01	94 \pm 0.004	81 \pm 1
LDA	83.8 \pm 1	85 \pm 0.01	83 \pm 0.01	95 \pm 0.01	84 \pm 0.01
SVM	81 \pm 0.84	80 \pm 0.01	79 \pm 0.01	94 \pm 0.003	80 \pm 0.008

Pairwise comparisons between the performance metrics obtained from each possible algorithm (i.e., NB, LDA, SVM) and feature set (i.e., $GE_{cm} + GE_{dm}$ or $GE_{cm} + GE_{dm} + CQ$) combination were performed with two-tailed, two-sample *t*-tests. Comparisons of each performance metric (i.e., accuracy, precision, recall, specificity, and F1-score) among different combinations of algorithm and feature set choices aimed to assess: (i) whether training each algorithm with only fNIRS features resulted in a statistically significantly different classification performance when compared to training the same algorithm with a combination of fNIRS and behavioral features, and (ii) whether there exists an algorithm and feature set combination with a statistically significantly higher performance when compared to all other options.

Table 2. Four-class classification performances of NB, LDA, and SVM when trained with fNIRS and behavioral features (i.e., GE_{cm} , GE_{dm} , and CQ). Each performance metric is represented in percentages (%) as the mean value across all runs \pm standard deviation of the mean. Bold-typed results denote significantly higher performance of the corresponding algorithm with respect to the results when the algorithm is fed with fNIRS only features.

Method	Accuracy	Precision	Recall	Specificity	F1-Score
NB	84.68 \pm 1.3	85 \pm 0.01	83 \pm 0.01	95 \pm 0.01	84 \pm 0.01
LDA	83.8 \pm 1.6	84 \pm 1.1	83 \pm 1.4	94 \pm 0.04	84 \pm 1.2
SVM	85 \pm 1.77	86 \pm 1.6	84 \pm 1.7	95 \pm 0.5	85 \pm 1.7

3. Results

Table 1 presents the four-class classification performances of NB, LDA, and SVM classifiers when they were trained with two fNIRS-derived features (i.e., GE_{cm} , GE_{dm}). All algorithms achieved accuracy, precision, recall, and F1-score performances above 81%, while the specificity scores were all above 94%. It should be noted that LDA performed significantly higher than both SVM and NB (Figure 6) in terms of accuracy (83.8 \pm 1%, $p < 0.05$), precision (85 \pm 0.01%, $p < 0.05$), recall (83 \pm 0.01%, $p < 0.05$), specificity (95 \pm 0.01%, $p < 0.05$), and F1-score (84 \pm 0.01%, $p < 0.05$). The performances of NB and SVM were not statistically significantly different in terms of the reported metrics.

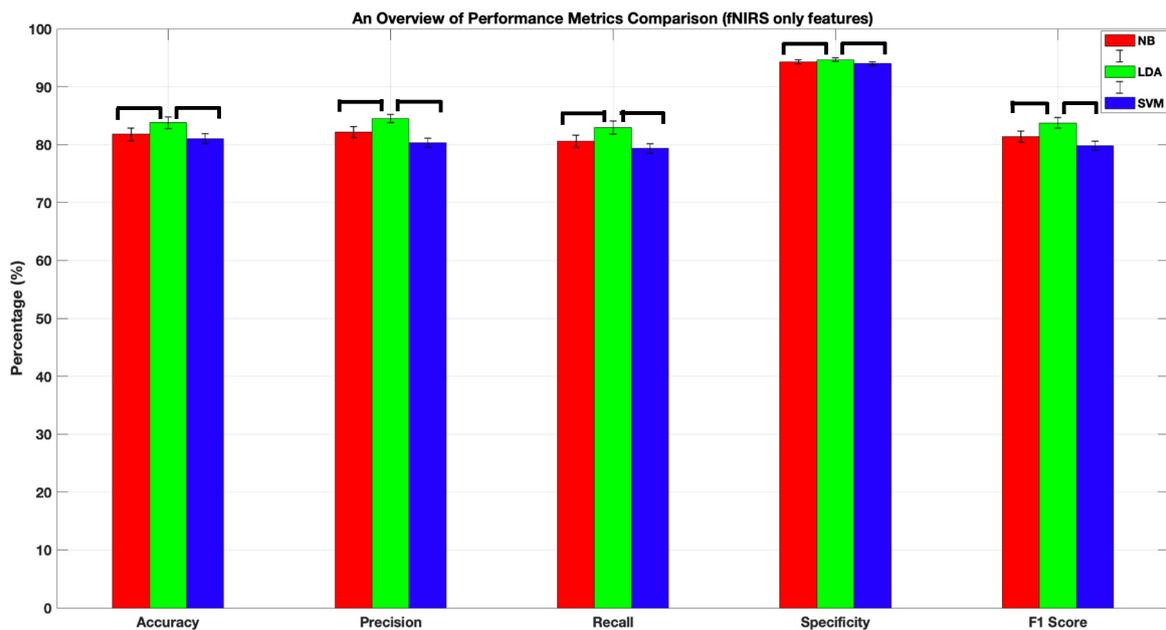


Figure 6. Classification performances of NB, LDA, and SVM algorithms after being trained with fNIRS-derived features (i.e., GE_{cm} and GE_{dm}). Horizontal lines depict statistically significant differences between performances of different algorithm pairs. All algorithms achieved accuracy, precision, recall, and F1-score performances above 80%, while the specificity scores were above 94%. LDA performed significantly higher than both SVM and NB in terms of accuracy, precision, recall, specificity, and F1-score. The error bars represent standard error of the mean performance after 10 runs of a 10-fold C.V.

Table 2 presents the four-class classification performances of NB, LDA, and SVM classifiers when they were trained with fNIRS and behavioral features (i.e., GE_{cm} , GE_{dm} , and CQ). Comparisons between the performance of each tabulated algorithm with respect to the corresponding performance obtained with fNIRS only features (Table 1) were performed with paired t -tests, and bold-typed results (Table 2) denote significantly higher performance of the corresponding algorithm when compared to the results when the same

algorithm is fed with fNIRS only features. All algorithms achieved accuracy, precision, recall, and F1-score performances above 83%, while the specificity scores were all above 94%. Feeding NB and SVM with a combination of fNIRS and behavioral features resulted in a statistically significantly higher performance in each metric when compared to the performance obtained by training the same algorithm with fNIRS only features. However, LDA achieved a similar performance in each metric regardless of the type of feature set combination utilized for training. There were no statistically significant differences in accuracy, recall, specificity, and F1-scores among the three algorithms. Nonetheless, the precision score obtained with SVM was statistically significantly higher than both LDA and NB ($86 \pm 1.6\%$, $p < 0.05$, Table 2 and Figure 7).

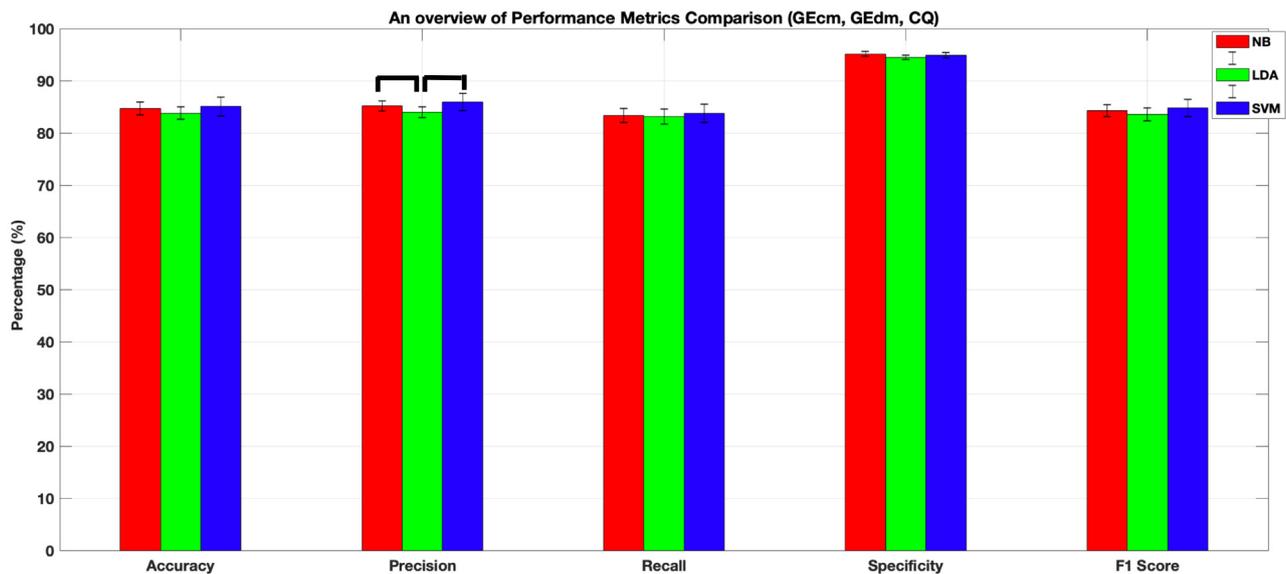


Figure 7. Classification performances of NB, LDA, and SVM algorithms after being trained with a combination of fNIRS and behavioral features (i.e., GE_{cm} , GE_{dm} , and CQ). Horizontal lines depict statistically significant differences between performances of different algorithm pairs. All algorithms achieved accuracy, sensitivity, precision, and recall performances above 83%, while the specificity scores were all above 94%. There was no statistically significant difference among accuracy, recall, specificity, and F1-score performances of the three algorithms. The error bars represent standard error of the mean performance after 10 runs of a 10-fold C.V.

Training LDA with fNIRS only features resulted in a comparable performance with the performance metrics obtained when the same algorithm was trained with a combination of fNIRS and behavioral features. A statistical comparison of the performance of the best-performing algorithm (LDA) and fNIRS only feature set combination of Table 1 with the performance metrics of NB and SVM of Table 2 demonstrated that no significant difference existed between any algorithm pair for accuracy, recall, specificity, and F1-scores.

To sum up, we conclude that training LDA with fNIRS only features results in a comparable performance with training the three supervised algorithms with a combination of fNIRS and behavioral features. Regarding the best performance, although there were no statistically significant differences among the three algorithms for accuracy, recall, specificity, and F1-scores (Figure 6), we should still note that SVM had the best performance in all metrics when trained with a combination of fNIRS and behavioral features obtained during the Stroop task (Tables 1 and 2 and Figure 7).

Figure 8 presents the confusion matrices for each algorithm, which demonstrate the true-positive and false-negative predictions attributed to each class. All algorithms achieved classification accuracies above 70% for each class. All algorithms demonstrated the highest true-positive prediction rate for SCZ patients, which was followed by OCD, HC, and MIG. SCZ and OCD subjects were not misclassified as HCs for any of the algorithms.

This result is significant as these two patient groups are expected to have the most distinct alterations in cognitive performance and cerebral hemodynamic activity during the Stroop task when compared to the HC group [13,17,66–71]. The fact that HC subjects were not misclassified as OCD or SCZ for any of the algorithms suggests the distinctive and physiology-related informative power of the selected features. However, HC subjects could be falsely attributed to the MIG class (SVM: 1.05%, NB: 4.74%, LDA: 4.21%) regardless of the algorithm type. This result is not surprising as MIG subjects were tested during the interictal period while they were exempt from attacks, hence their cognitive performance and the relevant spatial and topographic distribution of functional activation might have been similar to HCs during the interictal period. The consistencies in the classification performance patterns of the three algorithms as well as the consistency of performance results with physiology-related information highlight the distinctive power and biologically informative nature of the fNIRS-derived features utilized in the study. It can be concluded that training NB, LDA, and SVM with fNIRS-derived metrics demonstrates a differential diagnosis potential, regardless of the mathematical architecture of the algorithm.

		Predicted Label				
		HC	MIG	OCD	SCH	
Actual Label	Class	HC	MIG	OCD	SCH	
	HC	83.85%	16.15%	0.00%	0.00%	
	MIG	1.05%	70.53%	17.89%	10.53%	
	OCD	0.00%	14.29%	80.95%	4.76%	
	SCZ	0.00%	0.38%	0.00%	99.62%	
		SVM				
		Predicted Label				
		HC	MIG	OCD	SCH	
Actual Label	Class	HC	MIG	OCD	SCH	
	HC	77.69%	22.31%	0.00%	0.00%	
	MIG	4.74%	79.47%	10.53%	5.26%	
	OCD	0.00%	11.90%	83.33%	4.76%	
	SCZ	0.00%	3.08%	3.85%	93.08%	
		NB				
		Predicted Label				
		HC	MIG	OCD	SCH	
Actual Label	Class	HC	MIG	OCD	SCH	
	HC	87.69%	12.31%	0.00%	0.00%	
	MIG	4.21%	70.53%	17.37%	7.89%	
	OCD	0.00%	17.14%	78.10%	4.76%	
	SCZ	0.00%	0.00%	3.85%	96.15%	
		LDA				

Figure 8. Confusion matrices depicting the true-positive (shaded in yellow) and false-negative predictions of (A) SVM, (B) NB, (C) LDA when they are trained with a combination of fNIRS and behavioral features.

4. Discussion

The current diagnostic model for a majority of neuropsychiatric disorders relies on evaluation of measures which include clinical, observational, and/or behavioral scales that are obtained through interviews, questionnaires, observations, self-reports, and/or neuropsychiatric test batteries [3–5]. However, subjectivity introduced during both collection and clinical interpretation of these multi-domain measures brings forth the demand for more objective diagnostic markers. The high variability in clinical decisions for similar cases observed across different clinicians, cultures, and countries highlights the critical need for developing more objective decision support systems for diagnosis, which should ideally be based on quantitative measures of the neurophysiological alterations underlying each disorder.

Taking this critical demand into consideration, the presented study aimed to assess the feasibility and applicability of an fNIRS-based automated classification approach for accurate prediction and objective identification of the presence of three distinct neuropsychiatric or neurological disorder states which are known to induce alterations in frontal lobe function. The proposed machine learning-based classification approach involved training various supervised learning algorithms with (i) novel fNIRS-derived informative biomarkers and (ii) a combination of fNIRS-derived biomarkers and performance measures obtained during a neuro-cognitive test, named the Stroop task. We tested and compared the efficacy of training three commonly employed and computationally efficient supervised learning algorithms with these neuronally induced biomarkers, and their comparably high

performances were demonstrated with accuracy, precision, recall, specificity, and F1-scores. The performance of each algorithm in the correct identification of the presence of a disorder in each subject was evaluated by whether the subject was correctly labeled among the four classes, which included HCs, MIG, SCZ, and OCD. Hence, four-class brain–computer interface system designs were formulated which simply included the collection of hemodynamic signals with an fNIRS system while the subject was engaged in a Stroop task. Two global efficiency features were obtained from the PFC HBO signals, and accuracy and reaction rate performance obtained during the Stroop task were fused in a single behavioral feature, named the cognitive quotient (CQ). The comparably high performance scores obtained with the three classification algorithms, which have distinct mathematical architectures, highlighted the informative nature of these neuronally induced features. They also demonstrated the promising nature of integrating fNIRS-derived features together with cognitive performance scores from neuropsychiatric test measures and multivariate pattern analysis (MVPA) approaches for accurate recognition of neuropsychiatric disorder states. Our methodological approach resulted in increased classification accuracy when compared to the brain–computer interface (BCI) study designs conducted with fNIRS for other classification purposes, such as decoding mental thought processes or motor imagery signals [71,72].

In the following sections, we first evaluate the efficacy of NB, LDA, and SVM in correct identification of the presence of a disorder at the single subject level and we discuss the differential diagnostic potential of the proposed approach. We then highlight the importance of our findings, discuss the limitations of our study, and propose recommendations for future work.

4.1. Comparison of the Classification Performances of LDA, NB, and SVM

Training NB, LDA, and SVM with two fNIRS-derived functional connectivity metrics resulted in accuracy, precision, recall, and F1-score performances above 81%, while the specificity scores were all above 94%. While the performance metrics obtained with each algorithm had a very close range, it should be noted that LDA performed significantly higher than both SVM and NB in terms of accuracy ($83.8 \pm 1\%$, $p < 0.05$), precision ($85 \pm 0.01\%$, $p < 0.05$), recall ($83 \pm 0.01\%$, $p < 0.05$), specificity ($95 \pm 0.01\%$, $p < 0.05$), and F1-score ($84 \pm 0.01\%$, $p < 0.05$) when trained with fNIRS only features. A statistical comparison of the performance of the best-performing algorithm (LDA) and fNIRS only feature set combination of Table 1 with the performance metrics obtained by training each algorithm with a combination of fNIRS and behavioral features demonstrated that no significant difference existed between the performances of any algorithm pair for accuracy, recall, specificity, and F1-scores. Hence, we conclude that training LDA with fNIRS only features results in a comparable performance with training the three supervised algorithms with a combination of fNIRS and behavioral metrics. Regarding the best performance, we should note that SVM had the best performance in all metrics when trained with a combination of fNIRS and behavioral features obtained during a Stroop task (Tables 1 and 2 and Figure 6). However, we should also note that SVM did not have a statistically significantly higher performance than the rest of the algorithms for the majority of the performance metrics (i.e., accuracy, recall, specificity, and F1-scores reported in Figure 6). Hence, we can conclude that the utilized features are distinctive in nature as they performed well with all three classifiers regardless of the mathematical architecture of the algorithm. Obtaining a high classification performance with all classifiers highlights the feasibility and applicability of feeding machine learning-based methods with fNIRS-derived neuro-cognitive biomarkers for classification of disorder states associated with alterations in frontal lobe function.

With the recent advances in the computational power of daily used computers, MVPA methods have received increasing interest for automated identification and objective recognition of neurological and neuropsychiatric disorder states by use of structural and functional neuroimaging features. The majority of these studies examined the diagnostic

potential of utilizing multivariate features for: (i) correct identification of the presence of a disease state, (ii) rating the severity of a clinical state, or (iii) differentiating subgroups of patients. Arabshirani et al. provided an excellent review of previous neuroimaging studies that aimed at single-subject prediction of neurological, neurodegenerative, or neuropsychiatric disorders by use of structural and functional imaging features [61], while Orru et al. presented an extensive summary of the previous studies that utilized SVM for differentiating a neuropsychiatric disease state from a healthy state [73]. Regarding automated recognition of SCZ, Steardo et al. provided a review of classification studies that utilized a combination of SVM and neuroimaging markers [58]. The majority of these studies reported binary classification performances for differentiating a disorder state from a healthy state and the reported accuracies ranged from 67% to 100%. Regarding differentiation of OCD from a healthy state, the highest performance metrics were reported by Sen et al., who proposed the utility of resting state functional connectivity-derived network features with SVM [74]. They achieved 80% accuracy, 81% sensitivity, and 77% specificity with a relatively small sample size ($n = 16$ for OCD and $n = 13$ for HC). Similarly, three studies utilized MVPA methods and MRI-based neuroimaging markers for accurate prediction of the presence of migraine by use of two-class classification schemes, and the reported accuracies ranged between 80% and 96% [75–77].

Among three-class classification studies, Yu et al. reported a study where they used several MVPA methods to discriminate healthy controls ($n = 38$), schizophrenic patients ($n = 32$), and patients diagnosed with major depression disorder ($n = 19$). They achieved a correct classification rate of 81% using functional connectivity features from resting state fMRI scans [59]. Their sample size was also similar to our study. Kawazaki et al. built a binary classification model for differentiating SCZ from HC utilizing voxel-based morphometry features from MRI with a small dataset ($n = 30$ per class). Their classification accuracy performance was 80% [78]. Yassin et al. performed a three-class classification study where they trained several machine learning algorithms for accurate identification of autism spectrum disorder, healthy controls, and SCZ patients. The best results were achieved with MRI-derived cortical thickness parameters using a logistic regression (LR) classifier. Their overall maximum classification accuracy was reported as 69%. The maximum binary classification accuracies between different class pairs were less than 80% when tested with several classifiers, including SVM [79].

We should note that an objective comparison of our performance results with the performances reported in previous studies is complicated since the study designs differed in terms of sample size, number of classes, type and number of features, disorder types, C.V procedure, and the selected classifiers (Table 3). Nonetheless, we can still conclude that the performance metrics achieved with our four-class classification methodology fall in the high-performance spectrum among the performance metrics reported in previous studies, which targeted classification of various neuropsychiatric populations from healthy counterparts by use of structural and functional neuroimaging measures.

Table 3. Comparison of the classification performances of the discussed studies.

Author/s (Year)	Sample Size	Classifier(s)	Number of Classes	Features	Mean Accuracy (%)
Sen et al. (2017) [74]	16 OCD, 13 HC	SVM	2	Resting state network features derived from fMRI data	80
Chong et al. (2016) [75]	58 MIG, 50 HC	Quadratic Discriminate Analysis	2	Resting state network features derived from fMRI data	86
Yang et al. (2018) [76]	21 MIG without aura, 15 MIG with aura, 28 HC	Convolutional Neural Networks	2 and 3	Resting state network features derived from fMRI data	85–99 (2 class), 87(3 class)
Hernandez et al. (2014) [77]	15 HC, 20 MIG, 19 Medication Abuse	SVM	2	Graph theoretical features derived from fMRI data	87

Table 3. Cont.

Author/s (Year)	Sample Size	Classifier(s)	Number of Classes	Features	Mean Accuracy (%)
Yu et al.(2013) [59]	32 SCZ, 19 MDD, 38 HC	SVM	3	Resting state network features derived from fMRI data	81
Kawasaki et al.(2007) [78]	30 SCZ, 30 HC	Multivariate Linear Model	2	Voxel based morphometry features extracted from MRI data	80
Yassin et al. (2020) [79]	64 SCZ, 36 ASD, 106 HC	Logistic Regression	3	Cortical thickness and subcortical volume features derived from MRI data	69
Pardo et al. (2006) [80]	10 SCZ, 10 BP, 8 HC	LDA	3	Neuropsychiatric test scores and structural metrics derived from MRI data	96
Present work	20 MIG, 26 OCD, 21 SCZ, 13 HC	LDA, SVM, NB	4	Cognitive quotient and Global Efficiency metrics derived from fNIRS data	84.7 (LDA), 83.8(NB), 85 (SVM)

4.2. Potential of the Proposed Methodology for Differential Diagnosis

Comorbidities often exist among major neuropsychiatric disorders in the form of overlapping behavioral symptoms and similar neurobiological alterations. Hence, one of the major challenges for a precise diagnostic decision is to be able to differentially diagnose neuropsychiatric disorders which have overlapping symptoms, such as SCZ, MDD, and BD [14,57]. While differential diagnosis of the patient groups presented in this study would be easy to decipher at the clinical stage, we should emphasize the fact that our work serves as a proof-of-concept study to demonstrate the utility of combining fNIRS-derived functional connectivity metrics obtained during a cognitive test with machine learning-based classification methods for assisting accurate classification and objective identification of neuropsychiatric disorder states associated with frontal lobe functional abnormalities.

Recent studies have presented compelling evidence that a wide variety of neuropsychiatric disorders are characterized with alterations in the neural activity of the PFC [13]. However, whether there exists a distinct topographical distribution of functional abnormalities specific to each neuropsychiatric disorder and whether each neuropsychiatric disorder can be associated with a distinct abnormality in cerebral activation that can be recognized by fNIRS during a cognitive test remains unclear. In our study, all algorithms achieved classification accuracies above 70% for each class. All algorithms demonstrated the highest true-positive prediction rate for SCZ patients, which was followed by OCD, HC, and MIG. HC subjects were not misclassified as OCD or SCZ for any of the algorithms. These two patient groups are expected to have the most distinct alterations in cognitive performance and cerebral hemodynamic activity during the Stroop tasks when compared to the HC group [13,17,66–70]. Hence, the fact that HC subjects were not misclassified as OCD or SCZ for any of the algorithms suggests the distinctive and physiology-related informative power of the selected features. HC subjects could be falsely attributed to the MIG class. This result is not surprising as MIG subjects were tested during the interictal period which might be cognitively similar to a healthy state, and hence the spatial and topographic distribution of their functional activation might not be significantly different from HCs during the Stroop task. OCD and SCZ subjects were not misclassified as HCs for any of the algorithms. The consistencies in the classification performance patterns of the three algorithms as well as the consistency of performance results with physiology-related information highlight the distinctive power and biologically informative nature of the fNIRS-derived features utilized in the study.

To sum up, our results suggest that training NB, SVM, or LDA with the fNIRS-derived global efficiency metrics obtained during a Stroop task demonstrates a differential diagnosis potential, regardless of the mathematical architecture of the algorithm. Our findings also support the notion that some novel neuro-biological features obtained with fNIRS

methodology during cognitive tasks might serve as distinct signatures of the spatiotemporal characteristics of different neuropsychiatric disorder states which are associated with frontal lobe function abnormalities. Exploration of such informative and biologically derived features and combining them with machine learning-based classification approaches may have significant potential for differential diagnoses of psychopathologies which have comorbidities and overlapping symptoms.

4.3. Limitations of the Study and Recommendations for Future Work

We should note that the sample sizes of our subject groups were still small, although they exceeded the sample sizes reported in many of the previously reported classification studies in neuropsychiatry literature [58,61,73–80]. As a continuation of this study, we will test the performance of our methodology on a larger subject cohort. Our classification problems will include a higher number of disorder types and we will test the efficacy of identifying patients with comorbidities. We will also test the informative power of extracting hemodynamic and cognitive features from concurrent fNIRS recordings taken during a variety of neuropsychological tests which target different aspects of cognition.

Deep learning (DL) techniques have a great potential to improve the performance of fNIRS-based BCI systems if sufficiently large training sets are available [81,82]. The major advantages of these techniques rely on their ability to capture the complexity of neural information embedded in the HBO signal patterns through optimization of the network structures [81]. Indeed, there exists some successfully implemented DL classifiers with fNIRS and EEG signals [82–86]. However, we avoided testing the utility of DL algorithms in the presented work because of the limited cohort size of each group. Models constructed with DL algorithms have a tendency to overfit data when they are trained with small sample sizes (i.e., $n < 5000$) [81]. Future work will involve testing the efficacy of DL algorithms for addressing the presented classification problem in a larger cohort size and by utilizing data augmentation procedures.

In the presented study, clinical diagnosis of each participant was performed by experienced psychiatrists after careful follow-up procedures, and their final clinical decision was considered the golden standard. Hence, we could test and report the performance of each algorithm by whether it could correctly predict the final clinical decision of an experienced psychiatrist whose decision is considered as ground truth. Although the participants included in the study were reported to have strong and distinct symptoms and the clinicians had good clinical expertise for making a correct diagnosis, there still exists a possibility that some of the patients might have been given a different diagnosis by a different group of clinicians and might be incorrectly labeled. Hence, we can only report the value and high performance of combining fNIRS only markers and supervised learning algorithms in correctly predicting the clinical decision of an experienced clinician. Nonetheless, such a decision support system still might assist young clinicians who have not gained enough expertise with patients.

While the differential diagnosis of the patient classes reported in this study might not be a difficult problem in the clinics, we should note that this is a proof-of-concept study for demonstrating the potential of predicting a clinical decision through analysis of informative hemodynamic features obtained noninvasively in a clinical setting with a wearable and ergonomic fNIRS system design. Hemodynamic information can be collected with similar system designs during similar cognitive tests or vasomechanical challenges and can be processed to extract biomarkers which can be used for differential diagnosis of neurological or neuropsychiatric disorders that are known to induce abnormalities in PFC function.

5. Conclusions

The overarching goal of this study was to test the feasibility of an fNIRS-based BCI system design for accurate and objective identification of the presence of neuropsychiatric or neurological disorders. Our results demonstrate the potential of training supervised learning algorithms with fNIRS-derived hemodynamic and cognitive features for precise

recognition of the presence of a neurological or neuropsychiatric disorder at the single-subject level. They also highlight the promise of exploring PFC-based neurofunctional features as distinctive and objective biomarkers of neuropsychiatric or neurological disorders which are associated with alterations in frontal lobe function. Neuronally induced biomarkers can be easily obtained in clinical settings with portable, wearable fNIRS systems. Such system designs might also have great potential for objective classification and differential diagnosis of major neuropsychiatric disorders which, in most cases, have overlapping behavioral symptoms across each other and are hard to distinguish when decisions are based solely on observation, self-report, interview, and/or rating scales.

Author Contributions: Conceptualization, S.B.E. and G.Y.; methodology, S.B.E. and G.Y.; validation, S.B.E. and G.Y.; formal analysis, S.B.E. and G.Y.; investigation, S.B.E. and G.Y.; data curation, S.B.E. and G.Y.; writing—original draft preparation, S.B.E.; writing—review and editing, S.B.E. and G.Y.; visualization, S.B.E. and G.Y.; supervision, S.B.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Pamukkale University (protocol code B.30.2.PAU.0.01.00.00-200/130-0146, 15 January 2007).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is available upon request.

Acknowledgments: The authors would like to thank Ata Akin for his generosity in sharing the data set and the fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Klöppel, S.; Abdulkadir, A.; Jack, C.R.; Koutsouleris, N.; Mourão-Miranda, J.; Vemuri, P. Diagnostic neuroimaging across diseases. *Neuroimage* **2012**, *61*, 457–463. [[CrossRef](#)] [[PubMed](#)]
2. Singh, I.; Rose, N. Biomarkers in psychiatry. *Nature* **2009**, *460*, 202–207. [[CrossRef](#)] [[PubMed](#)]
3. World Health Organization. *International Statistical Classification of Diseases*, 10th ed.; World Health Organization: Geneva, Switzerland, 1992.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed.; American Psychiatric Press: Washington, DC, USA, 1994.
5. Azechi, M.; Iwase, M.; Ikezawa, K.; Takahashi, H.; Canuet, L.; Kurimoto, R.; Nakahachi, T.; Ishii, R.; Fukumoto, M.; Ohi, K.; et al. Discriminant analysis in schizophrenia and healthy subjects using prefrontal activation during frontal lobe tasks: A near-infrared spectroscopy. *Schizophr. Res.* **2010**, *117*, 52–60. [[CrossRef](#)] [[PubMed](#)]
6. Aboraya, A.; Rankin, E.; France, C.; El-Missiry, A.; John, C. The Reliability of Psychiatric Diagnosis Revisited: The Clinician's Guide to Improve the Reliability of Psychiatric Diagnosis. *Psychiatry* **2006**, *3*, 41–50.
7. Ward, C.H.; Beck, A.T.; Mendelson, M.; Mock, J.E.; Erbaugh, J.K. The psychiatric nomenclature. Reasons for diagnostic disagreement. *Arch. Gen. Psychiatry* **1962**, *7*, 198–205. [[CrossRef](#)]
8. Drevets, W.C. Neuroimaging studies of mood disorders. *Biol. Psychiatry* **2000**, *48*, 813–829. [[CrossRef](#)]
9. Okada, G.; Okamoto, Y.; Yamashita, H.; Ueda, K.; Takami, H.; Yamawaki, S. Attenuated prefrontal activation during a verbal fluency task in remitted major depression. *Psychiatry Clin. Neurosci.* **2009**, *63*, 423–425. [[CrossRef](#)]
10. Weyandt, L.; Swentosky, A.; Gudmundsdottir, B.G. Neuroimaging and ADHD: fMRI, PET, DTI Findings, and Methodological Limitations. *Dev. Neuropsychol.* **2013**, *38*, 211–225. [[CrossRef](#)]
11. Ferrari, M.; Quaresima, V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* **2012**, *63*, 921–935. [[CrossRef](#)]
12. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Pavia, J.M.; Wolf, U.; Wolf, M. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* **2014**, *85*, 6–27. [[CrossRef](#)]
13. Ehlis, A.-C.; Schneider, S.; Dresler, T.; Fallgatter, A.J. Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage* **2014**, *85*, 478–488. [[CrossRef](#)] [[PubMed](#)]
14. Yeung, M.K.; Lin, J. Probing depression, schizophrenia, and other psychiatric disorders using fNIRS and the verbal fluency test: A systematic review and meta-analysis. *J. Psychiatr. Res.* **2021**, *140*, 416–435. [[CrossRef](#)] [[PubMed](#)]

15. Song, H.; Chen, L.; Gao, R.; Bogdan, I.I.M.; Yang, J.; Wang, S.; Dong, W.; Quan, W.; Dang, W.; Yu, X. Automatic schizophrenic discrimination on fNIRS by using complex brain network analysis and SVM. *BMC Med. Inform. Decis. Mak.* **2017**, *17* (Suppl. S3), 166. [[CrossRef](#)] [[PubMed](#)]
16. Garrity, A.G.; Pearlson, G.D.; McKiernan, K.; Lloyd, D.; Kiehl, K.A.; Calhoun, V.D. Aberrant “Default Mode” Functional Connectivity in Schizophrenia. *Am. J. Psychiatry* **2007**, *164*, 450–457. [[CrossRef](#)] [[PubMed](#)]
17. Woodward, T.S.; Leong, K.; Sanford, N.; Tipper, C.M.; Lavigne, K.M. Altered balance of functional brain networks in Schizophrenia. *Psychiatry Res. Neuroimaging* **2016**, *248*, 94–104. [[CrossRef](#)]
18. Sharma, A.; Weisbrod, M.; Kaiser, S.; Markela-Lerenc, J.; Bender, S. Deficits in fronto-posterior interactions point to inefficient resource allocation in schizophrenia. *Acta Psychiatr. Scand.* **2010**, *123*, 125–135. [[CrossRef](#)]
19. Venkataraman, A.; Whitford, T.J.; Westin, C.-F.; Golland, P.; Kubicki, M. Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophr. Res.* **2012**, *139*, 7–12. [[CrossRef](#)]
20. Narayanaswamy, J.; Hazari, N.; Venkatasubramanian, G. Neuroimaging findings in obsessive-compulsive disorder: A narrative review to elucidate neurobiological underpinnings. *Indian J. Psychiatry* **2019**, *61* (Suppl. S1), S9–S29. [[CrossRef](#)]
21. Fajnerova, I.; Gregus, D.; Francova, A.; Noskova, E.; Koprivova, J.; Stopkova, P.; Hlinka, J.; Horacek, J. Functional Connectivity Changes in Obsessive-Compulsive Disorder Correspond to Interference Control and Obsessions Severity. *Front. Neurol.* **2020**, *11*, 568. [[CrossRef](#)]
22. Erdoğan, S.B.; Yükselen, G.; Yegül, M.M.; Usanmaz, R.; Kiran, E.; Derman, O.; Akin, A. Identification of impulsive adolescents with a functional near infrared spectroscopy (fNIRS) based decision support system. *J. Neural Eng.* **2021**, *18*, 056043. [[CrossRef](#)]
23. Çiftçi, K.; Sankur, B.; Kahya, Y.P.; Akin, A. Multilevel Statistical Inference from Functional Near-Infrared Spectroscopy Data during Stroop Interference. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 2212–2220. [[CrossRef](#)]
24. Aydore, S.; Mihcak, M.K.; Ciftci, K.; Akin, A.; Mihak, M.; Cifti, K. On Temporal Connectivity of PFC via Gauss–Markov Modeling of fNIRS Signals. *IEEE Trans. Biomed. Eng.* **2009**, *57*, 761–768. [[CrossRef](#)] [[PubMed](#)]
25. Dadgostar, M.; Setarehdan, S.K.; Shahzadi, S.; Akin, A. Functional connectivity of the PFC via partial correlation. *Optik* **2016**, *127*, 4748–4754. [[CrossRef](#)]
26. Einalou, Z.; Maghooli, K.; Setarehdan, S.K.; Akin, A. Effective channels in classification and functional connectivity pattern of prefrontal cortex by functional near infrared spectroscopy signals. *Optik* **2016**, *127*, 3271–3275. [[CrossRef](#)]
27. Einalou, Z.; Maghooli, K.; Setarehdan, S.K.; Akin, A. Graph theoretical approach to functional connectivity in prefrontal cortex via fNIRS. *Neurophotonics* **2017**, *4*, 041407. [[CrossRef](#)] [[PubMed](#)]
28. Akin, A. Partial correlation-based functional connectivity analysis for functional near-infrared spectroscopy signals. *J. Biomed. Opt.* **2017**, *22*, 126003. [[CrossRef](#)]
29. Akin, A. fNIRS-derived neurocognitive ratio as a biomarker for neuropsychiatric diseases. *Neurophotonics* **2021**, *8*, 035008. [[CrossRef](#)]
30. Zysset, S.; Müller, K.; Lohmann, G.; von Cramon, D. Color-Word Matching Stroop Task: Separating Interference and Response Conflict. *Neuroimage* **2001**, *13*, 29–36. [[CrossRef](#)]
31. Erdogan, S.B.; Özsarfatı, E.; Dilek, B.; Sogukkanlı Kadak, K.; Hanoglu, L.; Akin, A. Classification of motor imagery and execution signals with population-level feature sets: Implications for probe design in fNIRS based BCI. *J. Neural Eng.* **2019**, *16*, 026029. [[CrossRef](#)]
32. Akgül, C.B.; Akin, A.; Sankur, B. Extraction of cognitive activity-related waveforms from functional near-infrared spectroscopy signals. *Med. Biol. Eng. Comput.* **2006**, *44*, 945–958. [[CrossRef](#)]
33. Erdoğan, S.B.; Yucel, M.; Akin, A. Analysis of task-evoked systemic interference in fNIRS measurements: Insights from fMRI. *Neuroimage* **2014**, *87*, 490–504. [[CrossRef](#)] [[PubMed](#)]
34. Pollonini, L.; Bortfeld, H.; Oghalai, J.S. PHOEBE: A method for real time mapping of optodes-scalp coupling in functional near-infrared spectroscopy. *Biomed. Opt. Express* **2016**, *7*, 5104–5119. [[CrossRef](#)] [[PubMed](#)]
35. Jobsis, F.F. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* **1977**, *198*, 1264–1267. [[CrossRef](#)]
36. Firbank, M.; Eiji, O.; Delpy, D.T. A theoretical study of the signal contribution of regions of the adult head to near-infrared spectroscopy studies of visual evoked responses. *Neuroimage* **1998**, *8*, 69–78. [[CrossRef](#)] [[PubMed](#)]
37. Cope, M.; Delpy, D.T. System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Med. Biol. Eng. Comput.* **1988**, *26*, 289–294. [[CrossRef](#)]
38. Delpy, D.T.; Cope, M.; Zee, P.V.D.; Arridge, S.; Wray Sand Wyatt, J. Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys. Med. Biol.* **1988**, *33*, 1433–1442. [[CrossRef](#)]
39. Boas, D.A.; Dale, A.M.; Franceschini, M.A. Diffuse optical imaging of brain activation: Approaches to optimizing image sensitivity, resolution, and accuracy. *Neuroimage* **2004**, *23*, S275–S288. [[CrossRef](#)]
40. Sutoko, S.; Monden, Y.; Tokuda, T.; Ikeda, T.; Nagashima, M.; Funane, T.; Atsumori, H.; Kiguchi, M.; Maki, A.; Yamagata, T.; et al. Atypical dynamic-connectivity recruitment in attention-deficit/hyperactivity disorder children: An insight into task-based dynamic connectivity through an fNIRS study. *Front. Hum. Neurosci.* **2020**, *14*, 3. [[CrossRef](#)]
41. Al-Shargie, F.; Kiguchi, M.; Badruddin, N.; Dass, S.C.; Hani AF, M.; Tang, T.B. Mental stress assessment using simultaneous measurement of EEG and fNIRS. *Biomed. Opt. Express* **2016**, *7*, 3882–3898. [[CrossRef](#)]

42. Al-Shargie, F.; Tang, T.B.; Kiguchi, M. Assessment of mental stress effects on prefrontal cortical activities using canonical correlation analysis: An fNIRS-EEG study. *Biomed. Opt. Express* **2017**, *8*, 2583–2598. [[CrossRef](#)]
43. Hoshi, Y. Functional near-infrared optical imaging: Utility and limitations in human brain mapping. *Psychophysiology* **2003**, *40*, 511–520. [[CrossRef](#)] [[PubMed](#)]
44. Hoshi, Y.; Kobayashi, N.; Tamura, M. Interpretation of near-infrared spectroscopy signals: A study with a newly developed perfused rat brain model. *J. Appl. Physiol.* **2001**, *90*, 1657–1662. [[CrossRef](#)] [[PubMed](#)]
45. Latora, V.; Marchiori, M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* **2001**, *87*, 198701. [[CrossRef](#)]
46. Cui, X.; Bray, S.; Reiss, A.L. Speeded near infrared spectroscopy (NIRS) response detection. *PLoS ONE* **2010**, *5*, e15474. [[CrossRef](#)] [[PubMed](#)]
47. Hawkins, D.M.; Basak, S.C.; Mills, D. Assessing model fit by cross-validation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 579–586. [[CrossRef](#)] [[PubMed](#)]
48. Tai, K.; Chau, T. Single-trial classification of NIRS signals during emotional induction tasks: Towards a corporeal machine interface. *J. Neuroeng. Rehabil.* **2009**, *6*, 39. [[CrossRef](#)]
49. Sitaram, R.; Zhang, H.; Guan, C.; Thulasidas, M.; Hoshi, Y.; Ishikawa, A.; Shimizu, K.; Birbaumer, N. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *Neuroimage* **2007**, *34*, 1416–1427. [[CrossRef](#)]
50. Tanaka, H.; Katura, T. Classification of change detection and change blindness from near-infrared spectroscopy signals. *J. Biomed. Opt.* **2011**, *16*, 087001. [[CrossRef](#)]
51. Misawa, T.; Takano, S.; Shimokawa, T.; Hirobayashi, S. A brain-computer interface for motor assist by the prefrontal cortex. *Electron. Commun. Jpn.* **2012**, *95*, 1–8. [[CrossRef](#)]
52. Naseer, N.; Hong, K.-S. fNIRS-based brain-computer interfaces: A review. *Front. Hum. Neurosci.* **2015**, *9*, 3. [[CrossRef](#)]
53. Naseer, N.; Qureshi, N.K.; Noori, F.M.; Hong, K.-S. Analysis of Different Classification Techniques for Two-Class Functional Near-Infrared Spectroscopy-Based Brain-Computer Interface. *Comput. Intell. Neurosci.* **2016**, *2016*, 5480760. [[CrossRef](#)] [[PubMed](#)]
54. Khan, R.A.; Naseer, N.; Qureshi, N.K.; Noori, F.M.; Nazeer, H.; Khan, M.U. fNIRS-based Neurorobotic Interface for gait rehabilitation. *J. Neuroeng. Rehabil.* **2018**, *15*, 7. [[CrossRef](#)] [[PubMed](#)]
55. Shin, J.; Jeong, J. Multiclass classification of hemodynamic responses for performance improvement of functional near-infrared spectroscopy-based brain-computer interface. *J. Biomed. Opt.* **2014**, *19*, 67009. [[CrossRef](#)] [[PubMed](#)]
56. Shoushtarian, M.; Alizadehsani, R.; Khosravi, A.; Acevedo, N.; McKay, C.M.; Nahavandi, S.; Fallon, J.B. Objective measurement of tinnitus using functional near-infrared spectroscopy and machine learning. *PLoS ONE* **2020**, *15*, e0241695. [[CrossRef](#)]
57. Cearns, M.; Hahn, T.; Baune, B.T. Recommendations and future directions for supervised machine learning in psychiatry. *Transl. Psychiatry* **2019**, *9*, 271. [[CrossRef](#)]
58. Steardo, L.; Carbone, E.A.; De Filippis, R.; Pisanu, C.; Segura-Garcia, C.; Squassina, A.; De Fazio, P. Application of Support Vector Machine on fMRI Data as Biomarkers in Schizophrenia Diagnosis: A Systematic Review. *Front. Psychiatry* **2020**, *11*, 588. [[CrossRef](#)]
59. Yu, Y.; Shen, H.; Zeng, L.-L.; Ma, Q.; Hu, D. Convergent and Divergent Functional Connectivity Patterns in Schizophrenia and Depression. *PLoS ONE* **2013**, *8*, e68250. [[CrossRef](#)]
60. Sakai, K.; Yamada, K. Machine learning studies on major brain diseases: 5-year trends of 2014–2018. *Jpn. J. Radiol.* **2018**, *37*, 34–72. [[CrossRef](#)]
61. Arbabshirani, M.R.; Plis, S.; Sui, J.; Calhoun, V.D. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage* **2016**, *145*, 137–165. [[CrossRef](#)]
62. Stephan, K.E.; Schlagenhauf, F.; Huys, Q.; Raman, S.; Aponte, E.; Brodersen, K.; Rigoux, L.; Moran, R.; Daunizeau, J.; Dolan, R.; et al. Computational neuroimaging strategies for single patient predictions. *Neuroimage* **2016**, *145*, 180–199. [[CrossRef](#)]
63. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: An update SIGKDD. *Explorations* **2009**, *11*, 10–18. [[CrossRef](#)]
64. Keerthi, S.S.; Shevade, S.K.; Bhattacharyya, C.; Murthy, K.R.K. Improvements to Platt’s SMO algorithm for SVM classifier design. *Neural Comput.* **2001**, *13*, 637–649. [[CrossRef](#)]
65. Üstün, B.; Melssen, W.J.; Buydens, L.M. Facilitating the application of support vector regression by using a universal Pearson VII function based kernel. *Chemom. Intell. Lab. Syst.* **2006**, *81*, 29–40. [[CrossRef](#)]
66. Zhang, S.; Yang, G.; Ou, Y.; Guo, W.; Peng, Y.; Hao, K.; Zhao, J.; Yang, Y.; Li, W.; Zhang, Y.; et al. Abnormal default-mode network homogeneity and its correlations with neurocognitive deficits in drug-naïve first-episode adolescent-onset schizophrenia. *Schizophr. Res.* **2019**, *215*, 140–147. [[CrossRef](#)] [[PubMed](#)]
67. Yamamuro, K.; Kimoto, S.; Iida, J.; Kishimoto, N.; Tanaka, S.; Toritsuka, M.; Ikawa, D.; Yamashita, Y.; Ota, T.; Makinodan, M.; et al. Distinct patterns of blood oxygenation in the prefrontal cortex in clinical phenotypes of schizophrenia and bipolar disorder. *J. Affect. Disord.* **2018**, *234*, 45–53. [[CrossRef](#)] [[PubMed](#)]
68. Zhang, Z.; Wang, Y.; Zhang, Q.; Zhao, W.; Chen, X.; Zhai, J.; Chen, M.; Du, B.; Deng, X.; Ji, F.; et al. The effects of CACNA1C gene polymorphism on prefrontal cortex in both schizophrenia patients and healthy controls. *Schizophr. Res.* **2019**, *204*, 193–200. [[CrossRef](#)] [[PubMed](#)]
69. Okada, K.; Ota, T.; Iida, J.; Kishimoto, N.; Kishimoto, T. Lower prefrontal activity in adults with obsessive-compulsive disorder as measured by near-infrared spectroscopy. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2013**, *43*, 7–13. [[CrossRef](#)]

70. Ota, T.; Iida, J.; Sawada, M.; Suehiro, Y.; Yamamuro, K.; Matsuura, H.; Tanaka, S.; Kishimoto, N.; Negoro, H.; Kishimoto, T. Reduced Prefrontal Hemodynamic Response in Pediatric Obsessive–Compulsive Disorder as Measured by Near-Infrared Spectroscopy. *Child Psychiatry Hum. Dev.* **2012**, *44*, 265–277. [[CrossRef](#)]
71. Ahn, S.; Jun, S.C. Multi-Modal Integration of EEG-fNIRS for Brain-Computer Interfaces—Current Limitations and Future Directions. *Front. Hum. Neurosci.* **2017**, *11*, 503. [[CrossRef](#)]
72. Batula, A.M.; Kim, Y.E.; Ayaz, H. Virtual and Actual Humanoid Robot Control with Four-Class Motor-Imagery-Based Optical Brain-Computer Interface. *BioMed Res. Int.* **2017**, *2017*, 1463512. [[CrossRef](#)]
73. Orru, G.; Pettersson-Yeo, W.; Marquand, A.F.; Sartori, G.; Mechelli, A. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neurosci. Biobehav. Rev.* **2012**, *36*, 1140–1152. [[CrossRef](#)] [[PubMed](#)]
74. Sen, B.; Bernstein, G.A.; Xu, T.; Mueller, B.A.; Schreiner, M.W.; Cullen, K.R.; Parhi, K.K. Classification of obsessive-compulsive disorder from resting-state fMRI. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3606–3609. [[CrossRef](#)]
75. Chong, C.D.; Gaw, N.; Fu, Y.; Li, J.; Wu, T.; Schwedt, T.J. Migraine classification using magnetic resonance imaging resting-state functional connectivity data. *Cephalalgia* **2016**, *37*, 828–844. [[CrossRef](#)] [[PubMed](#)]
76. Yang, H.; Zhang, J.; Liu, Q.; Wang, Y. Multimodal MRI-based classification of migraine: Using deep learning convolutional neural network. *Biomed. Eng. Online* **2018**, *17*, 138. [[CrossRef](#)] [[PubMed](#)]
77. Jorge-Hernandez, F.; Chimenó, Y.G.; Garcia-Zapirain, B.; Zubizarreta, A.C.; Beldarrain, M.A.G.; Fernandez-Ruanova, B. Graph theory for feature extraction and classification: A migraine pathology case study. *Bio-Med. Mater. Eng.* **2014**, *24*, 2979–2986. [[CrossRef](#)]
78. Kawasaki, Y.; Suzuki, M.; Kherif, F.; Takahashi, T.; Zhou, S.-Y.; Nakamura, K.; Matsui, M.; Sumiyoshi, T.; Seto, H.; Kurachi, M. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage* **2007**, *34*, 235–242. [[CrossRef](#)]
79. Yassin, W.; Nakatani, H.; Zhu, Y.; Kojima, M.; Owada, K.; Kuwabara, H.; Gonoi, W.; Aoki, Y.; Takao, H.; Natsubori, T.; et al. Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl. Psychiatry* **2020**, *10*, 278. [[CrossRef](#)]
80. Pardo, P.J.; Georgopoulos, A.P.; Kenny, J.T.; Stuve, T.A.; Findling, R.L.; Schulz, S.C. Classification of adolescent psychotic disorders using linear discriminant analysis. *Schizophr. Res.* **2006**, *87*, 297–306. [[CrossRef](#)]
81. D’souza, R.N.; Huang, P.Y.; Yeh, F.C. Structural Analysis and Optimization of Convolutional Neural Networks with a Small Sample Size. *Sci. Rep.* **2020**, *10*, 834. [[CrossRef](#)]
82. Wickramaratne, S.D.; Mahmud, M.S. Conditional-GAN Based Data Augmentation for Deep Learning Task Classifier Improvement Using fNIRS Data. *Front. Big Data* **2021**, *4*, 659146. [[CrossRef](#)]
83. Khalil, K.; Asgher, U.; Ayaz, Y. Novel fNIRS study on homogeneous symmetric feature-based transfer learning for brain-computer interface. *Sci. Rep.* **2022**, *12*, 3198. [[CrossRef](#)]
84. Lyu, B.; Pham, T.; Blaney, G.; Haga, Z.; Sassaroli, A.; Fantini, S.; Aeron, S. Domain adaptation for robust workload level alignment between sessions and subjects using fNIRS. *J. Biomed. Opt.* **2021**, *26*, 022908. [[CrossRef](#)] [[PubMed](#)]
85. Hennrich, J.; Herff, C.; Heger, D.; Schultz, T. Investigating Deep Learning for fNIRS Based BCI. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 2844–2847. [[CrossRef](#)]
86. Chiarelli, A.M.; Croce, P.; Merla, A.; Zappasodi, F. Deep Learning for Hybrid EEG-fNIRS Brain-Computer Interface: Application to Motor Imagery Classification. *J. Neural Eng.* **2018**, *15*, 036028. [[CrossRef](#)] [[PubMed](#)]