

# **A Graph-Based Nonlinear Dynamic Characterization of Motor Imagery Toward an Enhanced Hybrid BCI**

Sarah M. I. Hosni<sup>1</sup>, Seyyed. B. Borgheai<sup>1</sup>, John McLinden<sup>1</sup>, Shaotong Zhu<sup>2</sup>, Xiaofei Huang<sup>2</sup>, Sarah Ostadabbas<sup>2</sup>, and Yalda Shahriari<sup>1\*</sup>

<sup>1</sup> Department of Electrical, Computer & Biomedical Engineering; University of Rhode Island (URI), Kingston, RI 02881 USA; and <sup>2</sup> Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA.

\* Corresponding author; E-mail: [yalda\\_shahriari@uri.edu](mailto:yalda_shahriari@uri.edu)

## **Acknowledgment**

This study was supported by the National Science Foundation (NSF-1913492, NSF-2006012) and the Institutional Development Award (IDeA) Network for Biomedical Research Excellence (P20GM103430).

**Abstract.** Decoding neural responses from multimodal information sources, including electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS), has the transformative potential to advance hybrid brain-computer interfaces (hBCIs). However, existing modest performance improvement of hBCIs might be attributed to the lack of computational frameworks that exploit complementary synergistic properties in multimodal features. This study proposes a multimodal data fusion framework to represent and decode synergistic multimodal motor imagery (MI) neural responses. We hypothesize that exploiting EEG nonlinear dynamics adds a new informative dimension to the commonly combined EEG-fNIRS features and will ultimately increase the synergy between EEG and fNIRS features toward an enhanced hBCI. The EEG nonlinear dynamics were quantified by extracting graph-based recurrence quantification analysis (RQA) features to complement the commonly used spectral features for an enhanced multimodal configuration when combined with fNIRS. The high-dimensional multimodal features were further given to a feature selection algorithm relying on the least absolute shrinkage and selection operator (LASSO) for fused feature selection. Linear support vector machine (SVM) was then used to evaluate the framework. The mean hybrid classification performance improved by up to 15% and 4% compared to the unimodal EEG and fNIRS, respectively. The proposed graph-based framework substantially increased the contribution of EEG features for hBCI classification from 28.16% up to 52.9% when introduced the nonlinear dynamics and improved the performance by approximately 2%. These findings suggest that graph-based nonlinear dynamics can increase the synergy between EEG and fNIRS features for an enhanced MI response representation that is not dominated by a single modality.

*Keywords.* Hybrid brain-computer interface (hBCI), multimodal data fusion, nonlinear dynamics, motor imagery (MI), graph-based feature extraction, recurrence quantification analysis (RQA)

# 1. Introduction

The progress in acquiring multimodal neuroimaging data opens new frontiers in systematically discovering a discriminative multimodal representation of neural responses to convey users' intent through hybrid brain computer interfaces (hBCIs). Incorporating multiple neural data sources, including electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS), into a hBCI framework expands the information content of the acquired neural signatures. This allows extracting a holistic multimodal electrical and vascular-hemodynamic representation of the underlying neural response across different spectral-temporal scales and has the potential of advancing BCI research. Traditionally, most motor imagery (MI) neural responses in BCI systems have been characterized using spectral analysis of EEG sensorimotor oscillatory variations in the  $\mu$  (8–12 Hz) and  $\beta$  (13–25 Hz) frequency bands for communication in EEG-based MI-BCIs (Kübler et al. 2005; Pfurtscheller and Lopes Da Silva 1999). However, despite continuous research efforts, EEG-based MI-BCIs still fall far short of achieving satisfactory performance levels. This is likely attributable to several inherent limitations of EEG, including its complex nonstationary nature, its poor spatial resolution, low signal-to-noise ratio (SNR), and potential disease-specific abnormalities in patients' electrical responses that impose additional challenges in extracting discriminative features from their MI responses (S. M. Hosni et al. 2019; Kasahara et al. 2012). Recently, fNIRS opened new horizons to characterize and decode the MI neural response for communication, with remarkably promising results for both healthy and patient populations (S M Hosni et al. 2020; Naseer and Hong 2015). fNIRS is an optical neuroimaging modality that measures changes in oxygenated (HbO<sub>2</sub>) and deoxygenated (HbR) hemoglobin concentrations on the cortical surface (Ayaz et al. 2013). Similar to EEG, fNIRS is a noninvasive, safe, portable, and cost-effective neuroimaging modality suitable for less restrictive recordings at patients' bedside (Ayaz et al. 2013). With EEG's instantaneous measure of neural activity, which complements fNIRS' metabolic-based spatial specificity, and its robustness to various types of artifacts, EEG and fNIRS represent excellent candidates for simultaneous multimodal recording that capitalizes on their similarities and maximizes the benefits from their complementary features. However, to exploit the potential benefits of combining both modalities in a hybrid MI-BCI framework fully, systematic approaches are needed to characterize the fused EEG-fNIRS representation of the MI neural responses to ensure that the discriminative features that capture the underlying neural dynamics embedded in these multimodal signals are extracted completely, and translated

meaningfully into efficient means of communication (Ahn and Jun 2017; Deligani et al. 2021).

To date, several MI-BCI studies have explored the combination of EEG with fNIRS to improve MI-BCI performance. Hybrid MI-BCI systems were first investigated in (Fazli et al. 2012) in which integrating fNIRS features with EEG spectral band-power features was suggested and enhanced the classification accuracy by 5% on average. However, in their study, the hybrid EEG-fNIRS configuration was based upon “decision-level” fusion of both modalities, in which the unimodal features were fed into separate classifiers on an individual level and the final outcome was selected optimally based upon the union of both modalities through a meta-classifier (i.e., the final decision is based upon either EEG or fNIRS). Moreover, only the mean changes in HbO<sub>2</sub> and HbR concentrations in fNIRS were used as fNIRS features, which did not capture the MI hemodynamic response temporal characteristics fully.

Other studies have investigated “feature-level” fusion, in which EEG and fNIRS features are concatenated and selected optimally before training the classifier to provide a broader range of information (Buccino, Keles, and Omurtag 2016; Chiarelli et al. 2018; Khan, Hong, and Hong 2014). This allows greater synergy between modalities as the classifier can learn a fused electrical-hemodynamic representation of the MI neural response. For example, in (Buccino, Keles, and Omurtag 2016) slope indicators of HbO<sub>2</sub> and HbR were combined with their mean to generate a set of features extracted from fNIRS signals and regularized common spatial patterns (RCSP) estimated separately for  $\mu$ - and  $\beta$ -filtered EEG signals, which resulted in 85% and 92% classification accuracies for EEG and fNIRS, respectively, in a Movement-Rest recognition task. In their work, simple concatenation of EEG and fNIRS features achieved a 2% improvement in accuracy over fNIRS unimodal classification, and no feature selection strategy was used, as their feature set was small. A hybrid EEG-fNIRS configuration was proposed in (Khan, Hong, and Hong 2014) to increase the number of control commands to four and eight commands, respectively. The control commands were decoded from ME tasks in EEG and mental arithmetic (MA) tasks in fNIRS. The MI features were extracted based upon the  $\beta$ -filtered EEG peak amplitudes, and only the mean values of HbO<sub>2</sub> and HbR were used as fNIRS features. They reported high classification accuracies across subjects that ranged from a mean of 80-95%. However, the objective of their study was to increase the number of control commands, and hence, the classification accuracy was evaluated individually for each modality and each control command without considering EEG-fNIRS data fusion. Chiarelli et al. (2018) achieved a significant 10% mean increase in performance

across subjects using multimodal EEG-fNIRS recording when compared to standalone modalities. The study used an advanced non-linear deep learning classification procedure to learn complex synergistic structures in the fused data on the feature-level. However, similar to previous studies, classical linear spectral power features were extracted from EEG, while only the mean HbO<sub>2</sub> and HbR concentrations were extracted as fNIRS features. While the improvement in performance over unimodal classification was promising, the extracted fNIRS features did not capture the hemodynamic response's temporal characteristics fully, which potentially degrades the unimodal classification accuracy and affects the fused EEG-fNIRS representation in the hBCI framework. In another study by Yin et al. (2015), a hybrid EEG-fNIRS framework was used to decode the force and speed of hand clenching to increase the number of classified commands in MI-BCI. The authors investigated the effect of broadening the information content within each modality on the classification accuracy of both the unimodal and fused levels. Band-power, amplitude, phase, and frequency features were extracted from  $\mu$  and  $\beta$  frequency bands in EEG rather than relying solely on the classical band-power features, while the difference between HbO<sub>2</sub> and HbR concentrations (HbD) was proposed as a single feature to increase the fNIRS classification accuracy. The mean classification accuracy improved by 18% over classical power features for EEG and 1% over individual HbO<sub>2</sub> and HbR features. To optimize the fused EEG-fNIRS feature selection, a feature optimization method based upon joint mutual information was proposed to remove redundant information that might affect the classification accuracy. However, the improvement in the hybrid framework's performance was 13% and 1% over unimodal fNIRS and EEG, respectively, when averaged over all subjects. To avoid a priori feature selection, deep learning was adopted in (Saadati, Nelson, and Ayaz 2020b) to extract the fused discriminative EEG-fNIRS representation directly and optimize classification performance. The decrease/increase in power spectral density, i.e., event-related de/synchronization (ERD/ERS) in the  $\mu$  and  $\beta$  frequency bands were extracted as discriminative features for EEG, while only the mean HbO<sub>2</sub> and HbR were extracted from fNIRS from all the recording channels. Promising hybrid performance improvements of 8% and 18% were achieved over unimodal fNIRS and EEG, respectively. Nevertheless, a large EEG dataset was required in their study to apply such deep-learning techniques (Shin et al. 2018).

In addition to the applications of EEG-fNIRS in MI-BCIs aforementioned, many studies have considered EEG-fNIRS feature-level fusion procedures to improve the performance in other hBCI applications. For example, Nguyen et al. (2017) classified driver drowsiness during long-term

simulated driving using combined EEG and fNIRS features and achieved a mean 5.5% improvement in accuracy when compared to single modality features. In (Saadati, Nelson, and Ayaz 2020a), a deep learning procedure was used to classify mental workload from temporal and spectral features extracted from fNIRS and EEG data with a mean improvement in classification of 8% and 23% over unimodal fNIRS and EEG features, respectively. EEG-fNIRS feature-level fusion was performed in (Al-Shargie, Tang, and Kiguchi 2017) using canonical correlation analysis (CCA) and resulted in 7.9% and 12.1% improved accuracy in a mental stress assessment problem over unimodal EEG and fNIRS, respectively. In (Deligani et al. 2021), the mutual information criterion was used as a powerful mathematical tool for feature selection to minimize the redundancy between high-dimensional multimodal EEG-fNIRS features to classify neural responses of a visuo-mental paradigm adopted from the P300 BCI speller application. The mutual information-based feature selection resulted in a 16% improvement in accuracy over hybrid classification with no feature selection, and 12% and 23% improvements over single modal classification using EEG and fNIRS, respectively.

Despite these considerable efforts to merge EEG with fNIRS for BCI applications, a significant gap remains before the benefits envisioned for multimodal hybrid techniques to achieve robust and desirable performance in practical environments can be obtained. This is likely attributable to the need of a methodological computational approach that optimizes the representation and selection of fused electrical and vascular-hemodynamic features in an integrative manner that exploits EEG and fNIRS's unique properties from the multimodal high-dimensional features efficiently (Ahn and Jun 2017). This is crucial to ensure a synergistic multimodal representation that is not dominated by a single modality in a hBCI context and to avoid redundant information in the fused feature space that can potentially degrade classification performance (Yin et al. 2015). To realize this, a systematic approach is needed to characterize the underlying dynamics of the MI neural response fully within each single modality before exploiting their complementary synergistic features. This is considered a crucial step before the selection of discriminative fused features to ensure that the information content of the fused MI response representation in the feature space is maximized. However, an additional challenge related to overfitting is imposed due to the increased dimensionality of the multimodal feature vector resulting from concatenating multimodal features from multiple channels, together with the limited number of samples available from simultaneous EEG-fNIRS recordings (Lotte et al. 2007). To address this, various feature selection approaches

have been suggested for MI-BCIs. However, least absolute shrinkage and selection operator (LASSO) feature selection algorithms, which relies on obtaining a refined model of the data through compressing some regression coefficients while setting others to zero based on a defined penalty function, have demonstrated practicality and superiority for MI-BCIs compared to other methods, particularly for relatively small datasets where deep learning techniques cannot be adopted (Jiang et al. 2020). Therefore, Lasso is advantageous for processing biased estimates with complex data and retains the advantage of subset shrinkage (Jiang et al. 2020).

In order to explore the underlying dynamics of MI responses within each single modality fully, the mean HbO<sub>2</sub> and HbR concentrations may not be sufficient to characterize the MI hemodynamic response's temporal dynamics completely. Including other statistical features can therefore broaden the information range of the hemodynamic feature space, and hence, improve the discrimination of fNIRS modality. Similarly, for EEG, the inherent complex nonstationary nature of the signals suggests investigating novel analysis methods to discriminate the MI response better beyond traditional linear spectral analysis features. Generally, from the perspective of nonlinear dynamics, the brain is a very complex dynamic system at all levels, from the nonlinear modeling of a single neuron's burst patterns to the macroscopic measurement of the activity of large groups of neurons measured with EEG from the surface of the scalp (McKenna, McMullen, and Shlesinger 1994). Rooted in chaos theory and the nonlinear dynamic systems literature, the recurrence quantification analysis (RQA) approach is a powerful nonlinear analytic tool developed for chaotic time series. RQA has been applied successfully to measure numerous biological signals' complexity, particularly when traditional techniques fail, including heart rate variability (Acharya, Chua, et al. 2011; Norbert Marwan et al. 2002; Zbilut, Thomasson, and Webber 2002) muscle (Bauer et al. 2017; Ikegawa et al. 2000), as well as epileptic EEG (Ngamga et al. 2016). Because they are more suitable for the analysis of short, noisy, and nonstationary time series, RQA complexity measures were proposed as a new way to analyze event-related potentials by identifying transitions in the brain process during surprising moments on a single trial level, rather than the traditional averaging of many trials, which emphasizes RQA's robustness (Norbert Marwan and Meinke 2004). Complex system's recurrent behavior involves transitions between periods of regularities to more complex irregular cycles, as well as chaos to chaos transitions (Norbert Marwan and Meinke 2004).

Recent evidence from our group and others suggested that nonlinear RQA features are sensitive

to transitions between motor tasks and rest in EEG (Ismail Hosni et al. 2021; Pitsik et al. 2020). For example, in a previous study (Ismail Hosni et al. 2021), nonlinear RQA and graph-based features were evaluated for an EEG-based MI-BCI with a mean improvement of 5.8% when compared to commonly used linear spectral features. To date, nonlinear features have not been explored for hybrid MI-BCIs. Exploring these nonlinear properties of EEG will broaden the information range decoded from the MI response. In this respect, RQA and graph-based features may provide a novel dimension of characteristic nonlinear features for hybrid MI-BCIs. These features characterize the underlying nonlinear dynamics of the complex sensorimotor neural system and their corresponding graph-based topological information from the EEG time series observed during motor tasks.

The goal of this study is to propose a systematic multimodal data fusion framework to represent and decode MI neural responses for hBCIs. The framework extracts high-dimensional multimodal linear and nonlinear features that expand the information content within the single modalities, and then adopts a fused feature selection strategy to identify discriminative synergistic EEG-fNIRS fused features to improve MI response discrimination. Our purpose is to extract a fused EEG-fNIRS feature set that will unveil a unique complementary representation that is not primarily dominated by a single modality toward an enhanced MI neural classification. For this purpose, we propose to analyze the changes in the nonlinear dynamics and recurrence patterns underlying the MI-based EEG neural responses and extract graph-based nonlinear features as an additional information dimension to the commonly used EEG linear spectral band-power features. Further, we extract a set of statistical features for the temporal characterization of fNIRS MI responses. We hypothesize that the EEG nonlinear features will complement the commonly used EEG linear spectral features that characterize the sensorimotor oscillatory variations in the electrical MI response combined with fNIRS features that characterize the temporal characteristics of the hemodynamic MI response in a hybrid EEG-fNIRS BCI fashion. The fusion of the extracted multimodal features is evaluated to determine the effect of the proposed features on the improvement in hybrid EEG-fNIRS MI-BCI's performance. In addition, a Fusion Level (*FL*) metric is proposed to quantify the balance of proportion of (EEG/fNIRS) feature contribution to the total number of selected fused feature set. This will evaluate the effect of the proposed features on extracting a synergistic multimodal representation that is not dominated by a single modality. Through the proposed analyses, we intend to investigate a new informative dimension to decode a



synergistic representation of MI neural signatures that will ultimately enhance hybrid MI-based BCI applications' performance. Simultaneous EEG-fNIRS data were recorded from eight healthy participants as they performed a MI-Rest task. RQA analysis and complex network theory were used to extract the nonlinear dynamics within the  $\mu$  and  $\beta$  frequency bands. Nonlinear graph-based RQA features were extracted from the recurrence plots (RPs) reconstructed from each  $\mu$ - and  $\beta$ -filtered one-dimensional EEG time series measured at each channel and its adjacency matrix reinterpretation. Spectral features were extracted using the mean power spectral density over the corresponding frequency bands. The fNIRS response's temporal characteristics were captured fully using various features (i.e., slope, mean, maximum, variance, skewness, kurtosis, and the difference between the mean and minimum activity) extracted from fNIRS HbO<sub>2</sub> and HbR. To overcome the challenge of high-dimensional multimodal feature vectors, we applied the LASSO algorithm to select the most informative fused features. Three types of EEG-fNIRS fused data were evaluated to analyze the effect of broadening the information content of the MI neural response on the performance of hybrid MI-BCI. The performance of the unimodal techniques was evaluated further for comparison through a 5-fold cross-validation procedure using a linear support vector machine (SVM).

## 2. Materials and Methods

A graphical illustration/flow diagram of the proposed computational multimodal framework is presented in Fig. 1, including the EEG and fNIRS data acquisition, the data preprocessing, and the multimodal feature fusion and classification procedure.

### 2.1. Data Acquisition, Participants, and Experimental Protocol

EEG and fNIRS signals were recorded simultaneously using a single cap mounted with both EEG electrodes and fNIRS optodes. EEG was recorded from 13 Ag/AgCl electrodes (i.e., channels) referenced to the left earlobe and amplified using a g.USBamp amplifier (g.tec medical engineering). The signals were digitized at 256 Hz and zero-phase bandpass filtered (1–45 Hz). The EEG channels covered the pre-motor (FC3, FC4), primary motor (C1, C3, Cz, C2, C4), sensorimotor (CP1, CP3, CP2, CP4), and parietal (P3, P4) areas of the brain according to the 10–5 system. An additional electrode was placed at FCz as the ground electrode. fNIRS data were recorded using NIRScout (NIRX Inc., NY, USA), with two near-infrared light wavelengths (760

nm and 850 nm) to acquire HbR and HbO<sub>2</sub> responses. The fNIRS probe layout resulted in 14 fNIRS channels covering the pre/frontal cortex in addition to the primary motor cortex. The signals were digitized at 15.6 Hz, and the optode montage was configured using 16 probes, 8 sources, and 8 detectors, with a separation distance of ~3 cm to maintain acceptable signal quality and sensing depth Fig. 2-left shows a schematic head montage model of the EEG-fNIRS sensors layout to capture the electrical-hemodynamic MI response (Buccino, Keles, and Omurtag 2016; S M Hosni et al. 2020; S. M. Hosni et al. 2019) Data acquisition for EEG and fNIRS and the design of the MI paradigm were handled by BCI2000 software (Schalk et al. 2004) and NIRX software (NIRX Inc., NY, USA).

Eight healthy participants with no reported history of neurological disease attended two MI data recording sessions on separate days. Each session contained three runs separated by approximately 5 minutes of rest. During each run, subjects were instructed to respond to visual cues presented on-screen with either left-hand motor imagery when the cue appears on the left side of the screen; right-hand motor imagery when the cue appears on the right side of the screen; and resting when the cue appears in the middle of the screen (Fig. 2-right). Each run consisted of 20 trials for each type of MI tasks randomly, with Rest trials in between (20 MI trials and 20 Rest trials per run). The resting cue was a green circle positioned in the middle of the screen to help them relax, allowing hemodynamic responses to return to baseline. None of the participants had previous BCI experience. The first session was used to familiarize the subjects with the task and the second session was used for data analysis.

## **2.2. Data Preprocessing**

Eye movement artifacts were removed from EEG data using the extended Infomax Independent Component Analysis (ICA) algorithm using the EEGLAB toolbox (Brunner, Delorme, and Makeig 2013). The artifact-free signal was then reconstructed after removing the predominant artifactual components identified by visual inspection. The data were then zero-phase bandpass filtered into the  $\mu$  (8–12 Hz) and  $\beta$  (13–25 Hz) frequency bands for further analysis. EEG data were re-referenced offline using a common average reference (CAR) (McFarland et al. 1997). For fNIRS data, the modified Beer-Lambert Law was used to calculate changes in the concentrations of HbO<sub>2</sub> and HbR using recorded alterations in the reflected light attenuation (Sassaroli and Fantini 2004). fNIRS data were then band-pass filtered at 0.01-0.09 Hz to eliminate physiological noise

caused by respiration ( $\sim 0.3$  Hz), cardiac activities ( $\sim 1$  Hz), and Mayer waves ( $\sim 0.1$  Hz). As fNIRS signal quality can be heavily compromised by poor coupling of optodes to the head, due to optical interference from dense and heavily pigmented hair, the quality of the signal was automatically evaluated through the signal-to-noise-ratio (SNR) of each channel using NIRScout. Further, an exclusion criterion was considered based on a correlation threshold between  $\text{HbO}_2$  and  $\text{HbR}$ , indicating a high-level physiological motion artifact (Cui, Bray, and Reiss 2010). The running correlation between  $\text{HbO}_2$  and  $\text{HbR}$  was calculated for each channel, and if it exceeded a 0.5 threshold, or was strictly  $-1$ , the channel was discarded. The data from both modalities (i.e., EEG and fNIRS) were segmented into 10-sec trials synchronized with the appearance of the visual stimulus cues (Rest/LMI/RMI). Individual MI trials that contained artifacts were automatically rejected based on subject-specific thresholds from both modalities. For MI vs. Rest classification, the trials were combined to form two sets with 60 trials for each condition of MI and Rest representing the two classes (i.e. 20 trials MI and 20 trials Rest for each run).

## **2.3. Data Analysis**

### **2.3.1. Linear Data Analysis**

For the EEG spectral features, the average power spectral density (PSD) was calculated using Welch's method from the filtered EEG signals giving  $\text{PSD-}\mu$  and  $\text{PSD-}\beta$  extracted from each channel. This resulted in a total of 26 linear EEG spectral features extracted from each trial from all the 13 EEG channels from both frequency bands (Kübler et al. 2005). In order to capture the characteristics of the MI hemodynamic response fully, seven discriminative features were extracted from each  $\text{HbO}_2$  and  $\text{HbR}$  response, corresponding to MI and Rest trials (S M Hosni et al. 2020), including slope ( $\text{SlopeHbO}_2$ ,  $\text{SlopeHbR}$ ), mean ( $\text{MeanHbO}_2$ ,  $\text{MeanHbR}$ ), maximum ( $\text{MaxHbO}_2$ ,  $\text{MaxHbR}$ ), variance ( $\text{VarHbO}_2$ ,  $\text{VarHbR}$ ), skewness ( $\text{SkewHbO}_2$ ,  $\text{SkewHbR}$ ), kurtosis ( $\text{KurtHbO}_2$ ,  $\text{KurtHbR}$ ), and the difference between the mean and the minimum ( $\text{DMMHbO}_2$ ,  $\text{DMMHbR}$ ). This resulted in a total of 196 fNIRS features extracted from each trial from both  $\text{HbO}_2$  and  $\text{HbR}$  (i.e., seven features were extracted from each fNIRS concentration change from each of the 14 channels). The features were extracted from several window sizes as follows: Considering the high temporal resolution for EEG, [0-2], [0-5], and [0-10] sec post-stimulus windows were considered for each frequency band. For slower hemodynamic response in fNIRS, [0-5], [2-7], [4-9], and [0-10] sec post-stimulus windows were considered for both fNIRS

concentration variations. These windows were chosen based on the typical hemodynamic response pattern of rising HbO<sub>2</sub> levels coupled with a decrease in HbR response approximately in the same time of the rise corresponding to an approximate hemodynamic response delay of 4 to 6 seconds relative to MI stimulus onset (S M Hosni et al. 2020). The optimized response windows were then selected for each modality based on the global peak of a nested 5-fold cross-validation classification procedure as explained in section 2.4.

### 2.3.2. Graph-based Recurrence Quantification Analysis and Complex Network Features

In order to approximate the nonlinear neural dynamics underlying the MI and Rest tasks within each  $\mu$  and  $\beta$  frequency bands separately, the bandpass filtered one-dimensional EEG signal measured at each frequency band, each channel, and each 10-sec MI/Rest trial was projected to a multi-dimensional phase space based on Takens' theorem of time-delay embedding (Norbert Marwan et al. 2007) using the following equation (Takens 1981):

$$X_k = (x_k, x_{k+\tau}, \dots, x_{k+(m-1)\tau}) \quad (1)$$

where  $X_k$  is the reconstructed phase space vector based on the observation  $x_k$  of the bandpass filtered EEG time series  $(x_1, x_2, \dots, x_L)$ ,  $L$  is the number of samples in the EEG time series,  $\tau$  is the time delay, and  $m$  is the embedding dimension. The time-delay parameter ( $\tau$ ) and the embedding dimension ( $m$ ) were estimated using the average mutual information (AMI) and the false nearest neighbor (FNN) methods respectively (Eckmann, Oliffson Kamphorst, and Ruelle 1987). The time delay  $\tau$  and the embedding dimension  $m$ , were directly calculated for  $\mu$  and  $\beta$  frequency bands using only the training set of each of the 5 cross-validation folds as explained in section 2.4. The phase space reconstruction can be represented as an  $N \times m$  trajectory matrix  $X = (X_1, X_2, \dots, X_N)^T$  where  $N = L - (m - 1)$  is the number of states in time, and  $L$  is the number of samples in the EEG time series. Next, the recurrence plots (RPs) were created to visualize and quantify the recurrence patterns of the  $m$ -dimensional phase space trajectory  $X$  corresponding to each trial within each frequency band in a 2-dimensional plot (Eckmann, Oliffson Kamphorst, and Ruelle 1987). RPs were constructed by considering an  $\varepsilon$ -neighborhood of states in phase space as follows:

$$RP_{i,j}(\varepsilon) = \Theta(\varepsilon - \|X_i - X_j\|) \quad i, j = 1, \dots, N \quad (2)$$

where  $RP$  is the  $N \times N$  recurrence plot,  $N$  is the number of states in time,  $\Theta$  is the Heaviside function,  $\varepsilon$  is the recurrence threshold determining the size of the neighborhood in state space,  $\|\bullet\|$  is the Euclidean norm, and  $X$  is the reconstructed phase space vector. The recurrence exists when  $RP_{i,j}=1$ , (i.e., when the state space vectors at time  $i$  and  $j$  are within the same  $\varepsilon$ -neighborhood). The choice of the  $\varepsilon$ -neighborhood threshold was based on previous studies' recommendation and should not exceed 10% of the maximum phase space diameter (Norbert Marwan et al. 2007). Therefore, the value of  $\varepsilon$  was optimized for each participant by choosing from four different thresholds, namely 3%, 5%, 7%, and 10% of the maximum phase space diameter, for each frequency band, based on the global peak of a nested 5-fold cross-validation classification procedure as explained in section 2.4. The steps of the EEG nonlinear analysis can be visualized in Fig. 3 for a 30-sec Rest-MI-Rest task to illustrate the changes in the recurrence patterns across tasks as visualized in the RPs.

Features characterizing the recurrence patterns in each trial were extracted using graph-based RQA and complex network representations of the recurrence plots were reconstructed from each one-dimensional EEG time series measured at each channel. As it is common to find small distances between points in the reconstructed phase space that are close in time, the Theiler window in this study was set to a value of  $(m - 1)\tau$  so that only points that are farther than  $(m - 1)\tau$  from the diagonal were taken into account in the evaluation of the RQA measures (Javorka et al. 2009).

The recurrence patterns were quantified using the vertical and diagonal line structures of the  $RPs$  using the nonlinear RQA features. 12 RQA features were extracted, namely, recurrence rate ( $RR$ ), determinism ( $DET$ ), the mean length of a diagonal line ( $LMEAN$ ), the maximum length of a diagonal line ( $LMAX$ ), the maximum vertical length ( $VMAX$ ), the trapping time ( $TT$ ), the laminarity ( $LAM$ ), the entropy of diagonal line length distribution ( $ENTR$ ), the entropy of vertical line length distribution ( $ENTRV$ ), the recurrence time entropy ( $RTE$ ), and the recurrence times of first type ( $RT1$ ) and second type ( $RT2$ ) (Norbert Marwan et al. 2007),(Webber, Jr. and Marwan 2015),(N. Marwan 2013). In addition, two features from complex network theory, namely the global clustering coefficient ( $CC$ ) and transitivity ( $T$ ), were extracted from the adjacency matrix reinterpretation of the  $RP$  to include the topological characteristics of the recurrence patterns. The features were extracted from the  $RP$  corresponding with  $\mu$  and  $\beta$  frequency bands separately using the following equations(Webber, Jr. and Marwan 2015).

$$RR = \frac{100}{N^2} \sum_{i,j=1}^N RP_{i,j} \quad (3)$$

where  $RR$  is the recurrence rate, which is a measure of the density of recurrence points in  $RP$ , and  $N$  is the number of states.  $DET$  is a relative measure of the system's regularity defined as the percentage of recurrence points forming diagonal structures with respect to all recurrence points in  $RP$  as follows:

$$DET = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=1}^N lP(l)} \quad (4)$$

where  $P(l)$  is the frequency distribution of diagonal lines of length  $l$ , and  $l_{min} = 2$  is the length of the shortest diagonal (Norbert Marwan et al. 2007).  $LMAX$  is the maximum length of diagonal structures defined as follows:

$$LMAX = \max(\{l_i: i = 1 \dots N_l\}) \quad (5)$$

where  $l_i$  is the length of diagonal line  $i$ , and  $N_l$  is the total number of diagonal lines.  $LMEAN$  is the average diagonal line length defined as follows:

$$LMEAN = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=l_{min}}^N P(l)} \quad (6)$$

where  $P(l)$  and  $l_{min} = 2$  remain as defined in the explanation of  $DET$ .  $LAM$  is the laminarity, representing the probability of occurrence of laminar states in the system defined as follows:

$$LAM = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{v=1}^N vP(v)} \quad (7)$$

where  $P(v)$  is the frequency distribution of vertical lines of length  $v$ , and  $v_{min}=2$  is the length of the shortest vertical line (Norbert Marwan and Meinke 2004).  $VMAX$  is the maximum length of vertical structures, i.e., the longest duration of the laminar states defined as follows:

$$VMAX = \max(\{v_i: i = 1 \dots N_v\}) \quad (8)$$

where  $v_i$  is the length of vertical line  $i$ , and  $N_v$  is the total number of vertical lines.  $TT$  represents the average length of vertical lines defined as follows:

$$TT = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{v=v_{min}}^N P(v)} \quad (9)$$

where  $P(v)$  and  $v_{min}=2$  is the same as explained for  $LAM$ .  $ENTR$  refers to the Shannon entropy and it is defined as a complexity measure of the deterministic diagonal line structures in the  $RP$  as follows:

$$ENTR = - \sum_{l=l_{min}}^N p(l) \ln p(l) \quad (10)$$

where  $p(l) = P(l)/N_l$  is the estimated probability of finding a diagonal line of length  $l$ . Similarly,

the *ENTRV* refers to the entropy of vertical lines, reflecting the complexity of the vertical lines structure, defined as follows:

$$ENTRV = - \sum_{v=v_{min}}^N p(v) \ln p(v) \quad (11)$$

where  $p(v) = P(v)/N_v$  is the estimated probability of finding a vertical line of length  $v$ . *RTE* is the recurrence time entropy, i.e., the entropy of the “white” (non-recurrent) vertical lines indicating recurrence times  $t_w$  defined as follows:

$$RTE = - \sum_{t_w=1}^{T_{max}} p(t_w) \ln p(t_w) \quad (12)$$

where  $p(t_w) = P(t_w)/N_w$  is the estimated probability of a recurrence time  $t_w$ ,  $P(t_w)$  is the distribution of recurrence times, and  $T_{max}$  is the maximum recurrence time (Pitsik et al. 2020). The recurrence times of first type (**RT1**) and second type (**RT2**) are defined according to (Gao 1999; Norbert Marwan et al. 2007). First, denote the set of points that are defined as  $\varepsilon$ -neighborhood recurrences of an arbitrary phase space vector  $X_i$  as  $\mathfrak{R}_i = \{X: |X - X_i| \leq \varepsilon\}$ . The elements of this set correspond to the recurrence points  $j$  of the  $i$ th column  $\{R_{i,j}\}_{j=1}^N$  of the *RP*. Then, the recurrence times of the first type are defined as follows:

$$\{RT1_k = j_{k+1} - j_k\}_{k \in N} \quad (13)$$

where  $RT1_k$  represent the recurrence times corresponding to the recurrence point  $j$  in  $\mathfrak{R}_i$ , and  $k$  is the recurrence point index from 1 to  $N$ . The recurrence times of the second type (i.e., Poincare recurrence times) are calculated after removing all consecutive recurrence points with  $RT1_k = 1$  from the set  $\mathfrak{R}_i$ . This results in a new set  $\mathfrak{R}_i'$  of remaining recurrence points  $j'$ . Then, the corresponding recurrence times are calculated as follows:

$$\{RT2_k = j'_{k+1} - j'_k\}_{k \in N} \quad (14)$$

where  $RT2_k$  represent the recurrence times corresponding to the recurrence points  $j'$  in  $\mathfrak{R}_i'$ , and  $k$  is the recurrence point index from 1 to  $N$  as previously explained. Hence, *RT2* measures vertically the time distance between the beginning of (vertically) subsequent recurrence structures in the *RP* (Norbert Marwan et al. 2007)

To extract the topological characteristics of the trajectory in phase space, an undirected unweighted recurrence network whose elements are denoted by a binary adjacency matrix  $A_{i,j}$ ,  $i, j = 1, \dots, N$ .  $A$  is defined in terms of the associated *RP* as follows (Donner et al. 2011):

$$A_{i,j} = RP_{i,j} - \delta_{i,j} \quad (15)$$

where  $\delta_{i,j}$  is the Kronecker delta. Each node of the network corresponds to an EEG sample in time, and edges are conveniently represented by the recurrence links based on the  $\epsilon$ -neighborhood in phase space (Donner et al. 2010).  $A_{i,j} = 1$  if vertex  $i$  connects to vertex  $j$ , and  $A_{i,j} = 0$  if the edge  $(i,j)$  does not exist, i.e., there is no recurrence of the system's state at time  $i$  and  $j$ . Since we considered a Theiler window in our study, the  $RP$  was regarded to be the adjacency matrix  $A$  for further analysis. From  $RP$ , two graph-based features were defined as follows (Norbert Marwan et al. 2009):

$$CC = \sum_{v=1}^N \frac{C_v}{N} \quad (16)$$

where  $CC$  is the global clustering coefficient, introducing a new recurrence aspect of the  $RP$  as it represents the probability that two recurrences of any state are also neighbors, and  $C_v$  is the local clustering coefficient defined for each node  $v$  (Norbert Marwan et al. 2009). Then, transitivity  $T$  which provides an effective measure of the global dimensionality of the underlying dynamical system is defined as follows (Feldhoff et al. 2013):

$$T = \frac{\sum_{i,j,k=1}^N A_{i,j} A_{j,k} A_{k,i}}{\sum_{i,j,k=1}^N A_{i,j} A_{k,i}} \quad (17)$$

In summary, the nonlinear analysis resulted in a total of 364 nonlinear graph-based RQA and complex network features extracted from each EEG trial from both  $\mu$  and  $\beta$  frequency bands (i.e., 14 features were extracted from 2 frequency bands from each of the 13 channels) to quantify the nonlinear dynamics underlying the MI-Rest tasks. All RQA related computations were performed using custom MATLAB (R2016b) code adapted from the CRP Toolbox (N. Marwan 2013). The features were extracted from several window sizes, similar to linear EEG features, and the optimized response window was selected within each frequency band based on the global peak of a nested 5-fold cross-validation classification procedure as explained in section 2.4. The illustration of the sensitivity of the extracted graph-based RQA and complex network features to the transition from the background neuronal activity (Rest) to MI through a time-dependent quantification of the feature values during 30-sec Rest-MI-Rest task is shown in Fig. S.1 in the supplementary section.

## 2.4. Multimodal Feature Fusion and Classification Procedure

To evaluate the effect of multimodal EEG-fNIRS fusion on classification performance and



investigate the effect of nonlinear features in complementing and increasing the synergy between both the linear PSD features in EEG and the characteristic temporal features in fNIRS, three types of EEG-fNIRS data fusion were evaluated; namely, EEG (linear)-fNIRS, EEG (nonlinear)-fNIRS, and EEG (linear+nonlinear)-fNIRS. For comparison, the performance of each of the three types of extracted features, i.e., EEG (linear), EEG (nonlinear) and fNIRS, were individually evaluated on a unimodal level.

Linear SVM was used to evaluate performance for each subject using a nested 5-fold cross-validation procedure to avoid biased estimation of the generalization error. To account for the high response variability in the neural data, hyper-parameter optimization was performed independently for each of the 5 outer-folds based on the global peak of the nested 5-fold cross-validation procedure (i.e., inner-folds) within each of the 5 outer-folds. The nonlinear RQA parameters (i.e., the time delay  $\tau$ , the embedding dimension  $m$ , and the  $\epsilon$ -neighborhood threshold) as well as the classification parameters (i.e., optimized response window and optimized number of selected features) were estimated and simultaneously optimized using only the training set of each of the outer-folds within the nested 5-fold cross-validation procedure. As the MI response dynamics vary across modalities (EEG/fNIRS), feature types (linear/nonlinear), and frequency bands for EEG, the response windows were optimized independently for EEG (linear), EEG (nonlinear) and fNIRS features. For unimodal EEG (linear/nonlinear) classification, three post stimulus windows were considered for the optimization testing for both the  $\mu$  and  $\beta$  frequency bands ([0-2], [0-5], and [0-10] sec). To do so, for each window, the features were extracted from the window corresponding to each band and then concatenated to constitute a single (linear/nonlinear) EEG feature vector containing all extracted features from all channels corresponding to the combined response windows. As each frequency band has its own separate dynamics, with the  $\beta$  frequency band being the higher of the two, combinations of longer  $\mu$ -band response windows with shorter  $\beta$ -band response windows were also considered resulting in a total of six possible EEG (linear/nonlinear) combined response windows for optimization testing corresponding to the six combinations of  $\mu$  and  $\beta$  response windows ( $\mu$ [0-2] sec+ $\beta$ [0-2] sec,  $\mu$ [0-5] sec+ $\beta$ [0-5] sec,  $\mu$ [0-10] sec+ $\beta$ [0-10] sec,  $\mu$ [0-5] sec+ $\beta$ [0-2]sec,  $\mu$ [0-10] sec+ $\beta$ [0-2] sec, and  $\mu$ [0-10] sec+ $\beta$ [0-5] sec). These response windows were considered in the hyper-parameter optimization of unimodal EEG (linear/nonlinear) classification and for the multimodal EEG-fNIRS classification. Due to the relatively slower dynamics of fNIRS, four possible response windows ([0-5], [2-7], [4-9], and [0-10] sec) were

considered for the nested 5-fold cross-validation optimization testing. The seven aforementioned fNIRS features were extracted from both HbO<sub>2</sub> and HbR and then concatenated resulting in four possible combinations of fNIRS response windows (HbO<sub>2</sub>[0-5] sec+HbR[0-5] sec, HbO<sub>2</sub>[2-7] sec+HbR[2-7] sec, HbO<sub>2</sub>[4-9] sec+HbR[4-9] sec, and HbO<sub>2</sub>[0-10] sec+HbR[0-10] sec). In addition, for the unimodal EEG (nonlinear) classification, the time delay  $\tau$  and the embedding dimension  $m$  were directly calculated for the  $\mu$  and  $\beta$  frequency bands using only the training set of each of the 5 cross-validation outer-folds. However, the choice of  $\varepsilon$  for the nonlinear features extraction was simultaneously optimized with the choice of the response windows within the nested 5-fold cross validation procedure. This resulted in 96 possible unimodal nonlinear EEG feature vectors to be considered for optimization testing (6 response windows  $\times$  4  $\times$  4 possible combinations of  $\varepsilon$ -neighborhood threshold for  $\mu$  and  $\beta$  frequency bands).

For multimodal classification, the EEG (linear/nonlinear) feature vector was concatenated with the fNIRS feature vector for EEG (linear/nonlinear)-fNIRS feature fusion. Similarly, for EEG (linear+nonlinear)-fNIRS feature fusion, all of the extracted features were concatenated resulting in a single multimodal feature vector completely characterizing the linear and nonlinear EEG dynamics as well as the fNIRS temporal dynamics. For hyper-parameter optimization, all the possible combinations of multimodal feature vectors corresponding to all possible choices of combined multimodal response windows were considered. For EEG (linear)-fNIRS, 24 possible multimodal feature vectors corresponding to the multimodal response windows were considered within the nested 5-fold cross-validation procedure. The multimodal response windows were based on all the possible combinations of response windows within each modality (i.e., 6 EEG response windows  $\times$  4 fNIRS response windows = 24 multimodal response windows). As for the EEG (nonlinear)-fNIRS and the EEG (linear+nonlinear)-fNIRS hyper-parameter optimization, the choice of  $\varepsilon$  for the nonlinear features extraction was simultaneously optimized with the choice of the multimodal response windows, resulting in 384 possible multimodal feature vectors corresponding to all possible combinations of parameters (24 multimodal response windows  $\times$  16 possible combinations of  $\varepsilon$ -neighborhood threshold for  $\mu$  and  $\beta$  frequency bands= 384 multimodal feature vector).

Due to the high dimensionality of the multimodal EEG-fNIRS feature vectors constructed using 26 EEG linear spectral features (PSD- $\mu$  and PSD- $\beta$  extracted from 13 EEG channels), 364 EEG nonlinear features (14  $\mu$  and  $\beta$  nonlinear features extracted from 13 EEG channels), and 196

fNIRS temporal features (7 fNIRS features extracted from each of HbO<sub>2</sub> and HbR from 14 fNIRS channels), fused feature selection represents a crucial step to ensure the selection of a discriminative fused EEG-fNIRS representation of the MI neural response. To do so, we adopted the LASSO feature selection scheme, which has been proven to exceed in performance for MI-BCIs when compared to other feature selection methods especially for relatively small datasets (Jiang et al. 2020). Therefore, a fused multimodal electrical-vascular representation of the MI response was selected on a multimodal level from the aforementioned constructed high-dimensional feature vectors. Similarly, for unimodal classification, discriminative features were selected using LASSO from the unimodal feature vectors to optimize the selected features within each unimodal technique. For all the constructed unimodal and multimodal feature vectors, the number of selected features was optimized within the nested 5-fold cross-validation procedure by considering different numbers of selected features from each feature vector (ranging from 5 features to 23 features with steps of 2). Finally, the optimized classification results for each subject were averaged over all the 5 outer cross-validation folds and reported along with the results of the nested 5-fold cross-validation procedure for all types of unimodal and hybrid classifications.

## 2.5. Fusion Level Quantification

In order to quantify the balance of proportion of (EEG/fNIRS) unimodal features in the three EEG-fNIRS data fusion types, we define the Fusion Level (*FL*) percentage to be the ratio of the number of selected features from each modality as below:

$$FL = \frac{\%Min-Contribution}{\%Max-Contribution} \quad (18)$$

where *%Min – Contribution* represents the smaller percentage of features contributed from either EEG or fNIRS to the total number of fused selected features and *%Max – Contribution* represents the larger percentage of features contributed from either EEG or fNIRS in the total number of fused selected features. The *FL* quantifies the balance of the contribution between both modalities such that a 100% *FL* indicates a perfectly balanced fusion (i.e., each modality contributed 50% of the selected fused features) and 0% indicates that one modality completely dominated the classification accuracy (i.e., equivalent to unimodal classification).

## 3. Results

In order to show the overall trend of the hyper parameter optimization for the unimodal feature

classification, Fig. 4 shows a series of bar plots comparing the optimized nested inner-loop 5-fold cross-validation accuracy averaged across the 5 outer-folds for the three types of unimodal features and different numbers of selected features for each subject. The plots show that performance of unimodal MI classification varies across modalities and feature types for each subject. As it is seen, for S-1, S-3, S-5, S-6, and S-7, the temporal fNIRS features outperformed both the linear and nonlinear EEG features. The nonlinear EEG features overall performed better than both linear EEG and fNIRS features for S-2 and S-4 and outperformed linear EEG features for almost all the participants except S-3 and S-8. However, the linear EEG features were the most discriminative for S-8. These plots highlight the importance of fully characterizing the hemodynamic fNIRS MI response with proper statistical features to improve the classification performance. Furthermore, the plots compare the nonlinear RQA and graph-based features to the linear PSD features, illustrating the discriminative ability of the nonlinear features for the majority of the participants. This emphasizes the importance of systematically characterizing the MI response dynamics across modalities and highlights the discriminative ability of the nonlinear recurrence patterns as a new informative dimension for MI EEG responses. **The figure also illustrates the effect of hyperparameter optimization, particularly the number of selected features, on the classification accuracy for all modalities. This highlights the importance of accounting for the inter-subject variability commonly observed in neural responses.** The details of the optimized parameters used for each of the 5 outer-folds for each subject are shown in Tables S.1, S.2, and S.3 in the supplementary section.

Table 1 shows the optimized unimodal classification performance for each subject, comparing the fNIRS, EEG (linear), and EEG (nonlinear) unimodal classification illustrating the optimized averaged 5-fold classification accuracy (outer-folds) and the median of the optimized number of selected features across folds for all feature types. The reported results are based on the optimized classification parameters for each fold (i.e., the response window, the number of selected features and the nonlinear RQA parameters if any) related to each subject's MI neural response in each modality. As shown in Table 1, the obtained average accuracies were  $92.1\% \pm 6.5\%$ ,  $81.2\% \pm 6.5\%$ , and  $85.8\% \pm 6.2\%$  using fNIRS, EEG (linear) and EEG (nonlinear) features respectively. Overall, the classification outcomes show that the performance of fNIRS features is superior in discriminating the MI neural response when compared to EEG (linear/nonlinear) features with  $\sim 11\%$  and  $\sim 6\%$  average performance improvement over EEG

(linear) and EEG (nonlinear) features respectively. This highlights the important role of discriminative fNIRS features (i.e., statistical features of HbO<sub>2</sub> and HbR) in classification accuracy and MI neural characterization. Moreover, the classification performance of the EEG (nonlinear) features indicates an overall average performance improvement of ~5% when compared to EEG (linear) features with an overall relatively higher number of selected features. This indicates that the complex nature of EEG signals, particularly MI responses, encompasses discriminative information beyond what is provided by EEG (linear) PSD features. These results support our principal hypothesis that EEG nonlinear graph-based RQA features can complement linear spectral features and fNIRS temporal features in a holistic fused representation for MI neural responses in order to improve the performance of MI classification in a hybrid BCI framework.

In order to show the overall trend of the hyper-parameter optimization for the hybrid multimodal classification, Fig. 5 shows a series of bar plots comparing the optimized nested inner-loop 5-fold cross-validation accuracy across the 5 outer-folds for the three types of EEG-fNIRS data fusion with different numbers of selected features for each subject. The figure also illustrates the effect of the number of selected features on the classification accuracy of the fused EEG-fNIRS data. The details of the optimized parameters used for each of the 5 outer-folds and the frequency of selection of the features in the optimized classification performance for each subject are shown in Tables S.4, S.5, S.6, and Fig. S.2 in the supplementary section. Unlike the unimodal features, the differences between the performance of the three types of fusion are relatively closer; however, the EEG (nonlinear)-fNIRS and the EEG (linear+nonlinear)-fNIRS outperformed the EEG (linear)-fNIRS for most of the subjects, indicating a substantial contribution of nonlinear EEG features. This shows the effect of complementing the linear EEG spectral features with the nonlinear graph-based RQA features and the effect of incorporating the temporal hemodynamic fNIRS features on the fusion classification accuracy for all the subjects.

Table 2 shows the optimized hybrid multimodal classification performance for each subject, comparing the EEG (linear)-fNIRS, EEG (nonlinear)-fNIRS, and EEG (linear+nonlinear)-fNIRS types of hybrid classification. In this table, the optimized averaged 5-fold classification accuracy (outer-folds) and the median of the optimized number of selected features across folds for all feature types are shown. The reported results are based on the optimized classification parameters for each fold (i.e., the hybrid response windows, the number of fused selected features and the nonlinear RQA parameters if any) related to each subject's MI neural response in each type of

fusion.

As shown in Table 2 the obtained average accuracies were  $93.8\% \pm 5.4\%$ ,  $96.1\% \pm 3.5\%$ , and  $96.1\% \pm 3.5\%$  using EEG (linear)-fNIRS, EEG (nonlinear)-fNIRS and EEG (linear+nonlinear)-fNIRS fused features respectively. This shows the discrimination ability of the nonlinear EEG features in complementing the fNIRS features when compared to the linear PSD features. Overall, the classification outcomes show that the hybrid EEG-fNIRS multimodal classification performs better than the unimodal (EEG/fNIRS) classification. Fusing linear EEG features to fNIRS features improved performance by  $\sim 2\%$  over fNIRS and  $\sim 13\%$  over EEG (linear) alone. Fusing EEG (nonlinear) with fNIRS features improved the performance by  $4\%$  over fNIRS and  $\sim 10\%$  over EEG (nonlinear) features, while fusing EEG (linear), EEG (nonlinear), and fNIRS features achieved the same accuracy as EEG (nonlinear)-fNIRS and improved the performance by  $4\%$ ,  $\sim 15\%$  and  $\sim 10\%$  over fNIRS, EEG (linear), and EEG (nonlinear) features respectively. This highlights the importance of decoding the hemodynamic fNIRS response features fully, in addition to complementing the classical EEG features with nonlinear features for improved performance of hBCIs for MI classification. In order to show the level of contribution of EEG features for each type of EEG-fNIRS fusion, Table 2 shows the percentage of EEG features contributing to the total number of selected fused multimodal features in EEG (linear)-fNIRS, EEG (nonlinear)-fNIRS, and EEG (linear+nonlinear)-fNIRS averaged across the 5 folds for each subject. The contribution of EEG linear features in EEG (linear)-fNIRS data fusion was  $\sim 28\%$  averaged over all subjects, whereas the contribution of EEG (nonlinear) features in EEG (nonlinear)-fNIRS data fusion was  $\sim 50\%$ . ~~This highlights the synergetic complementary and discriminative relationship between the nonlinear EEG features and the temporal fNIRS features, which is also illustrated by the performance improvement achieved by combining them for data fusion (i.e., EEG (nonlinear)-fNIRS achieved  $\sim 2\%$  performance improvement over EEG (linear) fNIRS with an increased contribution of EEG features).~~ Incorporating all feature types in EEG (linear+nonlinear)-fNIRS data fusion resulted in an increased contribution of EEG features to  $\sim 53\%$  averaged over all subjects with the same classification performance compared to EEG (nonlinear)-fNIRS data fusion.

Fig. 6 illustrates the fusion level (*FL*) quantification of each type of EEG-fNIRS data fusion. The figure compares the proportion of EEG/fNIRS features' balance across the fusion types and illustrates the classification accuracy corresponding to each *FL* (on top of each bar plot). This

figure illustrates the synergetic and complementary relationship between the fused EEG and fNIRS modalities by quantifying the balance percentage of their fusion in relation to hybrid classification accuracy. For all subjects, the *FL* for EEG (linear)-fNIRS data fusion was the least among all the other data fusion types. This indicates that adding the EEG (nonlinear) features as proposed results in an increased complementarity between EEG and fNIRS features, contributing to classification improvement. For most subjects the increased *FL* resulted in an increase in the classification accuracy. This highlights the importance of EEG-fNIRS complementary features for performance improvement; however, it further emphasizes the importance of feature-level fusion by employing a feature selection scheme to identify an optimized discriminative set of fused features from the high dimensional multimodal feature set without redundancy or contrast of information, and hence an optimal *FL* for performance improvement. Specifically, for S-3, the maximum classification accuracy was 100% for EEG (linear)-fNIRS data fusion with a *FL* of ~14%. This indicates that for this subject this classification accuracy was dominated by one modality rather than a result of the data fusion complementary improvement. However, EEG (nonlinear)-fNIRS and EEG (linear+nonlinear)-fNIRS data fusion achieved ~99% accuracy with ~69% and ~63% *FL* respectively for the same subject.

Fig. 7 compares the optimum unimodal (EEG/fNIRS) classification accuracy for each subject with the optimum hybrid classification accuracy. The plot indicates that the optimum hybrid multimodal classification performs better than the optimum unimodal classification for all subjects which supports the main hypothesis of this study. For subjects S-3 and S-7, 100% accuracy was achieved using unimodal classification, and the optimum multimodal classification did achieve the same 100% classification accuracy.

## 4. Discussion

Characterizing neural responses from multiple neuroimaging sources is important in BCI research, as it introduces complementary characteristic features from various neural perspectives to improve decoding and performance. However, current hybrid BCIs that rely on MI fall short of their true potential for multiple reasons, including the absence of systematic data fusion frameworks that characterize the MI neural response dynamics within each modality fully to reveal their complementary synergistic aspects, as well as the need to select the discriminative fused features from high-dimensional multimodal feature vectors properly. This study proposed an EEG-fNIRS

data fusion framework to decode and represent MI neural dynamics for a binary hBCI classification task fully. The framework relies on exploiting the underlying linear and nonlinear dynamics of neural responses across modalities and provides a comprehensive set of high-dimensional electrical-vascular multimodal features. A set of graph-based RQA features characterized the nonlinear recurrence patterns underlying the electrical MI neural response in the  $\mu$  and  $\beta$  spectral bands were characterized in phase space to complement the commonly used linear ERD/ERS spectral band-power features. The fNIRS-based MI response was characterized using a set of statistical features to quantify the morphological variations in the MI hemodynamic temporal response. Through the fused feature selection strategy adopted and the selection of the most discriminative subset of complementary features from all of the features evaluated, the proposed framework improved the hybrid MI-BCI's mean classification performance by approximately 15%, 10%, and 4% over the unimodal EEG (linear), EEG (nonlinear), and fNIRS features, respectively. Interestingly, the performance when EEG (linear+nonlinear)-fNIRS features were fused was similar to EEG (nonlinear)-fNIRS data fusion on average over all subjects, which indicates that including the EEG linear features in addition to the EEG (nonlinear) and fNIRS features did not increase the performance. This suggests an overall discriminative nature of the nonlinear dynamics of EEG when combined with the temporal hemodynamics features of the fNIRS response. However, it also indicates that the data fusion and fused feature selection scheme adopted select the most discriminative complementary fused features across modalities and feature types successfully, as, although there was an increase in feature dimensionality and potential redundant information, the performance did not degrade when compared to EEG (nonlinear)-fNIRS fusion.

Our findings suggest that the proper characterization of the underlying dynamics within each modality improved the performance of MI-BCI classification on the unimodal level and provided a comprehensive set of features with which to decode the response on a multimodal level systematically through the proposed multimodal data fusion framework. Moreover, the accuracy improvement was also realized by relying on hyper-parameter optimization, including the number of selected discriminative features, to account for the inter- and intra-subject variations commonly observed in neural responses. Our results indicate that optimizing the number of synergistic multimodal features can contribute toward an enhancement of hybrid BCI performance. The proposed framework and optimization approaches can be implemented based on BCI system



calibration sessions commonly used in the BCI field.

On the unimodal level, our results highlight the important role of discriminative fNIRS features (i.e., statistical features of HbO<sub>2</sub> and HbR) in classification accuracy and MI neural characterization. Moreover, the classification performance of the EEG (nonlinear) features indicates an overall average performance improvement of ~5% when compared to EEG (linear) features with an overall relatively higher number of selected features (Table 1). This indicates that the complex nature of EEG signals, particularly MI responses, encompasses discriminative information beyond what is provided by EEG (linear) PSD features. These results support our principal hypothesis that EEG nonlinear graph-based RQA features can complement linear spectral features and fNIRS temporal features in a holistic fused representation for MI neural responses in order to improve the performance of MI classification in a hybrid BCI framework.

Particularly for fNIRS, extracting a set of statistical features to characterize the MI hemodynamic response had a remarkable effect on the performance of the unimodal fNIRS approach. Based upon the hyper-parameter optimization trends of the unimodal features overall (Fig. 4) and the corresponding final results (Table 1), fNIRS performance was higher than that of EEG (linear and nonlinear) for most of the subjects. The unimodal fNIRS results indicated an ~11% improvement in performance compared to EEG (linear) features and ~6% compared to EEG (nonlinear). This highlights the discriminative power of the statistical fNIRS temporal features when compared with EEG spectral features and EEG graph-based nonlinear features. These results are consistent with previous findings that increasing the number of statistical features used to characterize the hemodynamic response and investigating the optimum feature combination influences classification accuracy substantially (Hong et al. 2017; Hong, Khan, and Hong 2018; S M Hosni et al. 2020; Naseer et al. 2016; Naseer and Hong 2013, 2015; Qureshi et al. 2017). For example, in (Naseer and Hong 2013), adding the slope of the fNIRS signal to the signal mean, as well as confining the response window, improved the classification accuracy significantly. In another study, Naseer et al. (2016) highlighted the importance of investigating the optimal feature combination to improve fNIRS classification performance. Their study considered all possible two- and three-feature combinations of signal slope, mean, variance, peak, kurtosis, and skewness, and they concluded that signal mean and peak was the optimum feature combination for their dataset. Qureshi et al. (Qureshi et al. 2017) investigated multiple two-feature combinations of the same set of statistical features to improve classification. Their results suggested that the signal

mean and skewness is an optimal feature combination. In (S M Hosni et al. 2020), Hosni et al. investigated two to eight possible feature combinations to optimize MI classification performance. Their findings suggested the importance of identifying subject-specific features, channels, and response windows of the fNIRS response to achieve optimal classification performance. The discrepancy in optimal feature combinations across subjects can be interpreted with respect to the variability identified commonly in fNIRS hemodynamic responses, and illustrates the need for a large number of statistical features to capture the potential subject-specific changes (Holper et al. 2011; Hong et al. 2017; S M Hosni et al. 2020). Our proposed framework accounts for fNIRS MI responses' subject-specific characteristics by capturing the temporal morphological variations in the hemodynamic response fully with a complete set of statistical features, and adopting hyperparameter optimization and feature selection to ensure that subject-specific discriminative features are identified to optimize performance. Notably, most fNIRS studies rely on channel selection approaches in addition to extracting discriminative features for classification (Al-Shargie et al. 2016; S M Hosni et al. 2020; Hu et al. 2010; Li et al. 2017; von Lühmann et al. 2020; Santosa, Hong, and Hong 2014). In this study, the features extracted were aggregated across all of the recording channels, which eliminated the need to select channels and rely on the feature selection scheme adopted to identify a subject-specific discriminative feature set across channels. Given fNIRS features' effect on a multimodal level, the multimodal results of fusing fNIRS features with EEG (linear) features improved the performance by  $\sim 13\%$  and  $\sim 2\%$  over unimodal EEG (linear) and unimodal fNIRS features, respectively. Although it is hard to make a direct comparison with the percentage of improvement of multimodal classification in other studies due to the differences in the number of extracted features, these results are consistent with those of previous studies and confirm that fNIRS is a unique modality that improves the performance of EEG in a hybrid MI-BCI system, and is capable to achieve competitive improvements over unimodal EEG-based systems (Buccino, Keles, and Omurtag 2016; Chiarelli et al. 2018; Fazli et al. 2012; Saadati, Nelson, and Ayaz 2020b; Yin et al. 2015). Further, it highlights the importance of exploiting various features across modalities for complementary representation of MI neural responses. However, considering the relatively slight improvement in performance when compared to the use of unimodal fNIRS features raises a question with respect to the extent to which the fusion of EEG PSD features with fNIRS exploits the full potential improvement of combining both modalities in a hybrid MI-BCI context.

By exploiting the discriminative features of the fNIRS MI response fully, an emerging challenge was to extract representative EEG dynamics that can complement the fNIRS features synergistically and yet to improve the performance on a multimodal level. This is especially important considering the challenging nature of classifying subject-specific neural responses to decode motor imagery signatures, particularly for severely disabled patients for whom BCI systems are originally designed. For these target groups, exploiting the neural information content fully from EEG and fNIRS is required to compensate for potential disease-related neural alterations that might affect one modality more than the other and to avoid relying on just one modality. To the best of our knowledge, this is the first study to propose nonlinear features for multimodal hybrid BCI applications that rely on EEG-fNIRS neuroimaging modalities. Here we demonstrated that by incorporating the nonlinear dynamics to complement the EEG linear PSD features, the proposed framework can expand the information content extracted from EEG successfully and achieve improved performance on the multimodal level with a high level of synergy between modalities. On a unimodal level, characterizing the MI neural response's nonlinear dynamics influenced the classification results remarkably. Given the unimodal EEG nonlinear features, the unimodal results indicated an improvement in performance of ~5% when compared to EEG linear features. These results verify previous findings further and demonstrate that extracting nonlinear EEG features improves MI-BCI's performance (Ismail Hosni et al. 2021). Our results are also consistent with those of Pitsik et al. (2020), who investigated the feasibility of using RQA features to characterize the nonlinear dynamics of EEG motor execution responses. Their findings revealed an increase in predictability and determinism during motor execution coupled with a decrease in complexity and chaos which transitioned back to increased complexity and reduced regularity during the background neuronal dynamics in rest. This was similarly observed in our study illustrated in the time-dependent quantification of all the extracted nonlinear features (Fig. S.1 in the supplementary section). However, our study is the first to investigate RQA features' ability to characterize motor imagery responses to improve BCI performance. Our findings confirmed the feasibility of using RQA features to characterize the neural responses that correspond with motor imagery, as they achieved improved classification performance for MI-BCIs both on the unimodal and multimodal levels. The success of RQA in characterizing the nonlinear dynamics of transitions between MI and rest in EEG can be explained by several studies that investigated nonlinear techniques for EEG as a time series that derives from a nonlinear

dynamic system. For example, in (Norbert Marwan and Meinke 2004), Marwan et al. used RQA to identify chaos-chaos transitions in brain potentials that are caused by stimulus events in single-trial event-related potential data. In (Acharya, Sree, et al. 2011), Acharya et al. quantified the nonlinear dynamics of EEG to classify epileptic EEG signals. Successfully RQA features could characterize the increase in regularity from “normal” EEG to “interictal” and then to “ictal” activity because of increased underlying rhythmicity. Recently, Baghdadi et al. (Baghdadi et al. 2021) suggested investigating the neuro-cognitive EEG responses to roughness stimuli using RQA. They used nonlinear analysis in their study, because the representation of texture features in cortical responses was proposed to follow a nonlinear model and RQA was found to characterize tactile sensations in EEG effectively.

Although EEG (nonlinear) features performed better than the EEG (linear) PSD features in our study, unimodal fNIRS features outperformed EEG (nonlinear) features ( $\sim 7\%$  improvement). This demonstrates further the importance of hemodynamic characteristics when decoding the MI neural response in addition to characterizing its nonlinear dynamics. Further, it highlights the importance of combining fNIRS with EEG-based BCIs to create an improved hybrid system. However, on a multimodal level, fusing the EEG (nonlinear) features with the fNIRS features improved the performance by  $\sim 10\%$  and  $4\%$  compared to unimodal EEG (nonlinear) and fNIRS features, respectively. These findings demonstrate the advantage of adopting graph-based RQA features to discriminate the MI neural response both on the unimodal and multimodal levels. The overall hyper-parameter optimization trends of the multimodal fused features (Fig. 5) and the corresponding final results (Table 2) demonstrate a relatively higher performance for EEG (nonlinear)-fNIRS and EEG (linear+nonlinear)-fNIRS fused features compared to EEG (linear)-fNIRS for most subjects. These results demonstrated the effect of incorporating the nonlinear dynamics of MI responses in the proposed computational framework to enhance performance.

Complementing EEG linear PSD features with the nonlinear graph-based features allowed more discriminative characteristics of the MI neural responses to be represented in the multimodal feature set, and thus, resulted in better performance overall on the multimodal level. The increased degree of contribution of the nonlinear EEG features in the EEG (nonlinear)-fNIRS fusion compared to the contribution of the linear features in EEG (linear)-fNIRS fusion support this result further (Table 2). This observation further highlights the synergetic complementary and discriminative relationship between the nonlinear EEG features and the temporal fNIRS features,

which is also illustrated by the slight performance improvement achieved by combining them for data fusion (i.e., EEG (nonlinear)-fNIRS achieved ~2% performance improvement over EEG (linear)-fNIRS with an increased contribution of EEG features).

—Our fusion level (*FL*) quantification results revealed clearly a decrease in the proportion of EEG/fNIRS features' balance in EEG (linear)-fNIRS fusion when compared to the other fusion types (Fig. 6). This indicates that for EEG (linear)-fNIRS fusion, the classification performance was dominated largely by fNIRS, given the smaller percentage of the EEG features' contribution in the fused feature selection scheme of EEG (linear)-fNIRS fusion (Table 2). The nonlinear EEG feature's increased percentage of contribution was illustrated in the increase of *FL* for the other fusion types (Fig. 6). This indicates that decoding and representing the underlying neural response appropriately can elucidate more synergistic dynamics and discriminative feature characteristics across modalities, and thus enhance classification performance at a hybrid level. Complementing EEG (linear) with EEG (nonlinear) features ensured that both EEG and fNIRS modalities are represented completely in the multimodal feature space, thus, providing an opportunity for increased synergy and complementarity between features for better performance. Overall, the inclusion of the discriminative EEG (nonlinear) features increased the *FL* for all subjects, which contributed to better performance for most participants (S-2, S-5, S-6, S-7, S-8). For the remaining subjects (S-1, S-3, and S-4), the increase in *FL* did not achieve the maximum performance. This is likely because of sub-optimal selection of the fused feature set, which emphasizes the importance of selecting a fused representative feature subset from the high-dimensional multimodal feature set properly to achieve improved performance.

In general, the results of this study demonstrated that multimodal EEG-fNIRS classification of MI neural responses performed better than relying solely on a single modality for all subjects (Fig. 7). Even in the case when unimodal classification dominated the classification accuracy (S-3 and S-7), increasing the number of features in multimodal classification did not degrade the accuracy for at least one of the fusion types evaluated. Hence, a systematic approach to feature-level fusion represents a crucial step to ensure robust classification performance, particularly with the increased number of features when all feature types are fused. Our study adopted a fused feature selection scheme based upon LASSO, which was supported by previous promising results, particularly for MI-BCIs and relatively small datasets (Jiang et al. 2020).

## 4.1. Limitations & Future Directions

One of the current limitations of this study is the low number of subjects due to the relative difficulty of recruiting human subjects for neural data recording sessions. This can affect the interpretability of the mean as an informative performance metric considering the variability of performance across subjects. Future work can consider addressing the challenge of neural response variability commonly faced in BCI research, potentially by increasing the number of subjects, including end users with severe motor impairment, and exploring other advanced methods, such as deep learning-based approaches to model subject-specific neural response variability. ~~One of the current limitations of this study~~ Another limitation is that the proposed framework did not consider other feature selection schemes for feature-level fusion. Comparing various feature selection algorithms was outside of the scope of this study; however, future work may consider adopting other approaches, including those based upon mutual information to explore other potential feature selection approaches to determine an optimal fused feature set (Deligani et al. 2021). Feature-level fusion allows synergistic complementary features across modalities to be selected, which has an advantage over decision-level fusion because of the potential adverse effect of cross-modality inconsistencies on the decision-level (Deligani et al. 2021; Wu et al. 2019). Further, future work could consider an early-fusion level that investigates the joint analysis of both modalities using approaches that represent the multimodal data in a joint representation space. For example, emerging deep multimodal machine learning approaches might be adopted to explore the possibility of capturing a data-driven, in-depth joint representation of the multimodal data that does not require feature extraction and/or fused feature-level selection processes (Venugopalan et al. 2021). In addition, future works need to be conducted to support our interpretations of the proposed nonlinear feature extraction methods further as a new informative dimension to understand better the neural characteristics that underlie motor imagery and enhance BCI performance by validating the statistical nonlinearity in the dataset through a surrogate data procedure. Analyzing the complex neural responses across different temporal scales encoded by EEG and fNIRS hold a promising fertile ground for BCI and neuroscience research in general. The amount of information that can be decoded from the latent dynamics of instantaneous EEG and the synergy that exist between these dynamics and the relatively slower metabolic responses encoded by fNIRS are relatively unexplored. Future work should consider studying the

relationship between these decoded fused features with fine and gross motor skills to further advance our understanding of the underlying neural dynamics of complex motor functions.

## 5. Conclusion

The proposed computational data fusion framework decoded a discriminative electrical-vascular multimodal neural response representation of MI by complementing the EEG spectral features with graph-based RQA features that quantify the nonlinear recurrence dynamics that underlie MI. These new features showed an increased synergy and complementarity between EEG and fNIRS modalities that enhanced the contribution of EEG features to the total number of multimodal features as well as the MI-BCI performance. The most discriminative task-informative features were identified through a multimodal fused feature selection scheme to address the computational challenges of multimodal characterization of the underlying discriminative neural dynamics. The performance evaluation revealed an average improvement of approximately 15%, 10%, and 4% by fusing EEG (linear), EEG (nonlinear) and fNIRS features when compared to unimodal EEG (linear), EEG (nonlinear), and fNIRS features respectively. On a unimodal level, the graph-based RQA features and the fNIRS features, respectively, improved the performance by  $\sim 5\%$  and  $\sim 11\%$  over the EEG spectral features, and demonstrated their discriminative strength. On a multimodal level, the optimum hybrid performance of 96.1% was achieved using discriminative fused feature selection from all feature types. This improved the conventional hybrid classification accuracy by  $\sim 2\%$ , with a considerable increase in the contribution of EEG features to the total number of fused features selected. This indicates an increased synergy between the nonlinear graph-based and fNIRS features. These findings highlight the importance of characterizing the underlying neural dynamics of the MI neural response across modalities, and suggest that nonlinear graph-based and fNIRS features are valuable information dimensions that can be exploited to improve hybrid MI-BCI performance. Further, our results highlight the importance of fused feature selection schemes and optimized fused discriminative feature sets for a systematic computational data fusion framework.

## Information Sharing Statement

Sharing the data used in this study is bound by the ethics of the institutional review boards of University of Rhode Island which approved the study. The custom code used in this study contained MATLAB<sup>®</sup> routines for preprocessing, analyzing and classifying the data based on the MATLAB<sup>®</sup> Signal Processing and Machine Learning toolboxes. The ICA was performed using the EEGLAB toolbox publicly available from the website: <https://sccn.ucsd.edu/eeglab/index.php>. The custom code for computing the recurrence plots and extracting the nonlinear features is available upon request and was adapted from CRPtoolbox, publicly available upon request from the website: <https://tocsy.pik-potsdam.de/CRPtoolbox/>.

## **Funding**

This study was supported by the National Science Foundation (NSF-1913492, NSF-2006012) and the Institutional Development Award (IDeA) Network for Biomedical Research Excellence (P20GM103430).

## **Ethics Approval**

The data recording was performed in the NeuralPC Lab, University of Rhode Island (URI) with Institutional Review Board (IRB) approval. Consent forms were obtained from all the subjects participating in the study.

## **Financial Interests**

The authors declare they have no financial interests



## TABLES

**TABLE 1:** Optimized averaged 5-fold cross-validation accuracy for unimodal classification and median of the optimized number of selected features across folds (maximum accuracy corresponding to each subject indicated in bold).

<i>Participant No.</i>	<i>fNIRS</i>	<i>Optimized #features (median)</i>	<i>EEG (Linear)</i>	<i>Optimized #features (median)</i>	<i>EEG (Nonlinear)</i>	<i>Optimized #features (median)</i>
S-1	<b>90.6</b>	19	75.3	11	82.4	17
S-2	84.7	9	85.9	11	<b>91.8</b>	7
S-3	<b>100.0</b>	9	74.7	7	74.7	9
S-4	84.3	7	77.1	11	<b>87.1</b>	13
S-5	<b>95.0</b>	9	75.0	9	80.0	19
S-6	<b>95.8</b>	11	83.2	9	88.4	7
S-7	<b>100.0</b>	5	86.3	7	90.0	9
S-8	86.7	5	<b>91.7</b>	9	<b>91.7</b>	5
Mean±SD	92.1±6.5	9.0±4.5	81.2±6.5	9.0±1.7	85.8±6.2	11.0±5.1

**Table 2:** Optimized Averaged 5-fold Cross-validation Accuracy for Multimodal EEG-fNIRS Fusion Classification and Median of the Optimized Number of Selected Features Across Folds (maximum accuracy corresponding to each subject and each EEG-fNIRS fusion type indicated in bold).

<i>Participant No.</i>	<i>EEG (linear)-fNIRS Fusion</i>	<i>Optimized #features (median)</i>	<i>%of EEG (linear) features</i>	<i>EEG (nonlinear)-fNIRS Fusion</i>	<i>Optimized #features (median)</i>	<i>%of EEG (nonlinear) features</i>	<i>EEG (linear + nonlinear)-fNIRS Fusion</i>	<i>Optimized #features (median)</i>	<i>%of EEG features</i>
S-1	<b>92.9</b>	15	45.2	91.8	11	47.3	90.6	15	57.1
S-2	88.2	19	32.6	<b>96.5</b>	13	50.2	<b>96.5</b>	15	49.6
S-3	<b>100.0</b>	9	12.1	98.7	17	40.7	98.7	13	38.7
S-4	84.3	9	34.7	90.0	15	61.5	<b>91.4</b>	17	67.4
S-5	95.0	19	21.7	<b>98.3</b>	19	46.9	<b>98.3</b>	15	43.6
S-6	<b>98.9</b>	17	25.1	96.8	15	61.1	97.9	15	57.8
S-7	97.5	13	20.9	<b>100.0</b>	13	35.8	<b>100.0</b>	13	48.5
S-8	93.3	7	33.0	<b>96.7</b>	17	58.2	95.0	15	60.5
Mean±SD	93.8±5.4	13.5±4.8	28.16±10.3	96.1±3.5	15±2.6	50.2±10.4	96.1±3.5	14.8±1.3	52.9±9.5

## REFERENCES

- Acharya, U Rajendra, S Vinitha Sree, et al. 2011. "Application of Recurrence Quantification Analysis for the Automated Identification of Epileptic EEG Signals." *International journal of neural systems* 21(3): 199–211.
- Acharya, U Rajendra, Eric Chern-Pin Chua, et al. 2011. "Automated Detection of Sleep Apnea from Electrocardiogram Signals Using Nonlinear Parameters." *Physiological measurement* 32(3): 287–303.
- Ahn, Sangtae, and Sung C Jun. 2017. "Multi-Modal Integration of EEG-FNIRS for Brain-Computer Interfaces – Current Limitations and Future Directions." *Frontiers in Human Neuroscience* 11: 503.
- Al-Shargie, Fares et al. 2016. "Mental Stress Assessment Using Simultaneous Measurement of EEG and FNIRS." *Biomedical Optics Express* 8 2583–98.
- Al-Shargie, Fares, Tong Boon Tang, and Masashi Kiguchi. 2017. "Assessment of Mental Stress Effects on Prefrontal Cortical Activities Using Canonical Correlation Analysis: An FNIRS-EEG Study." *Biomedical optics express* 8(5): 2583–98.
- Ayaz, Hasan et al. 2013. "Continuous Monitoring of Brain Dynamics with Functional near Infrared Spectroscopy as a Tool for Neuroergonomic Research: Empirical Examples and a Technological Development." *Frontiers in Human Neuroscience*.
- Baghdadi, Golnaz, Mahmood Amiri, Egidio Falotico, and Cecilia Laschi. 2021. "Recurrence Quantification Analysis of EEG Signals for Tactile Roughness Discrimination." *International Journal of Machine Learning and Cybernetics* 12(4): 1115–36.
- Bauer, C M et al. 2017. "The Effect of Muscle Fatigue and Low Back Pain on Lumbar Movement Variability and Complexity." *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology* 33: 94–102.
- Brunner, Clemens, Arnaud Delorme, and Scott Makeig. 2013. "Eeglab – an Open Source Matlab Toolbox for Electrophysiological Research." *Biomedical Engineering / Biomedizinische Technik*.

- Buccino, Alessio Paolo, Hasan Onur Keles, and Ahmet Omurtag. 2016. “Hybrid EEG-FNIRS Asynchronous Brain-Computer Interface for Multiple Motor Tasks.” *PLOS ONE* 11(1): 1–16.
- Chiarelli, Antonio Maria, Pierpaolo Croce, Arcangelo Merla, and Filippo Zappasodi. 2018. “Deep Learning for Hybrid EEG-FNIRS Brain-Computer Interface: Application to Motor Imagery Classification.” *Journal of neural engineering* 15(3): 36028.
- Cui, Xu, Signe Bray, and Allan L. Reiss. 2010. “Functional near Infrared Spectroscopy (NIRS) Signal Improvement Based on Negative Correlation between Oxygenated and Deoxygenated Hemoglobin Dynamics.” *NeuroImage* **49** 3039-3046.
- Deligani, Roohollah Jafari, Seyyed Bahram Borgheai, John McLinden, and Yalda Shahriari. 2021. “Multimodal Fusion of EEG-FNIRS: A Mutual Information-Based Hybrid Classification Framework.” *Biomed. Opt. Express* 12(3): 1635–50.
- Donner, Reik V. et al. 2010. “Recurrence Networks-a Novel Paradigm for Nonlinear Time Series Analysis.” *New Journal of Physics* **12** 033025.
- Donner, Reik V., Michael Small, Jonathan F. Donges, Norbert Marwan, Yong Zou, Ruoxi Xiang, and Jürgen Kurths. 2011. “Recurrence-Based Time Series Analysis by Means of Complex Network Methods.” *International Journal of Bifurcation and Chaos* **21** 1019-1046.
- Eckmann, J. P., O. Oliffson Kamphorst, and D. Ruelle. 1987. “Recurrence Plots of Dynamical Systems.” *World Scientific Series on Nonlinear Science Series A* **16** 441-446.
- Fazli, Siamac et al. 2012. “Enhanced Performance by a Hybrid NIRS-EEG Brain Computer Interface.” *NeuroImage* 59(1): 519–29.
- Feldhoff, J. H. et al. 2013. “Geometric Signature of Complex Synchronisation Scenarios.” *EPL* **102** 30007.
- Gao, J B. 1999. “Recurrence Time Statistics for Chaotic Systems and Their Applications.” *Physical Review Letters* 83(16): 3178.
- Holper, Lisa, Diego E. Shalóm, Martin Wolf, and Mariano Sigman. 2011. “Understanding

- Inverse Oxygenation Responses during Motor Imagery: A Functional near-Infrared Spectroscopy Study.” *European Journal of Neuroscience* **33** 2318-2328.
- Hong, Keum Shik, M. Raheel Bhutta, Xiaolong Liu, and Yong Il Shin. 2017. “Classification of Somatosensory Cortex Activities Using FNIRS.” *Behavioural Brain Research* **333** 225-234.
- Hong, Keum Shik, M. Jawad Khan, and Melissa J. Hong. 2018. “Feature Extraction and Classification Methods for Hybrid FNIRS-EEG Brain-Computer Interfaces.” *Frontiers in Human Neuroscience* **12** 246..
- Hosni, S M, S B Borgheai, J McLinden, and Y Shahriari. 2020. “An FNIRS-Based Motor Imagery BCI for ALS: A Subject-Specific Data-Driven Approach.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28(12): 3063–73.
- Hosni, S. M. et al. 2019. “An Exploration of Neural Dynamics of Motor Imagery for People with Amyotrophic Lateral Sclerosis.” In *Journal of Neural Engineering* **17** 16005.
- Hu, Xiao Su, Keum Shik Hong, Shuzhi S. Ge, and Myung Yung Jeong. 2010. “Kalman Estimator- and General Linear Model-Based on-Line Brain Activation Mapping by near-Infrared Spectroscopy.” *BioMedical Engineering Online* **9** 1-15.
- Ikegawa, Shigeki et al. 2000. “Nonlinear Time-Course of Lumbar Muscle Fatigue Using Recurrence Quantifications.” *Biological Cybernetics* **82** 373-382.
- Ismail Hosni, Sarah et al. 2021. “Graph-Based Recurrence Quantification Analysis of EEG Spectral Dynamics for Motor Imagery-Based BCIs.” In *43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, accepted.
- Javorka, M. et al. 2009. “The Effect of Orthostasis on Recurrence Quantification Analysis of Heart Rate and Blood Pressure Dynamics.” *Physiological Measurement* **30** 29.
- Jiang, Jing et al. 2020. “Temporal Combination Pattern Optimization Based on Feature Selection Method for Motor Imagery BCIs.” *Frontiers in Human Neuroscience* 14: 231.
- Kasahara, Takashi et al. 2012. “The Correlation between Motor Impairments and Event-Related Desynchronization during Motor Imagery in ALS Patients.” *BMC Neuroscience* 13 1-10.
- Khan, M Jawad, Melissa Jiyoun Hong, and Keum-Shik Hong. 2014. “Decoding of Four

- Movement Directions Using Hybrid NIRS-EEG Brain-Computer Interface.” *Frontiers in human neuroscience* 8: 244.
- Kübler, A. et al. 2005. “Patients with ALS Can Use Sensorimotor Rhythms to Operate a Brain-Computer Interface.” *Neurology* **64** 1775-1777.
- Li, Rihui, Thomas Potter, Weitian Huang, and Yingchun Zhang. 2017. “Enhancing Performance of a Hybrid EEG-FNIRS System Using Channel Selection and Early Temporal Features.” *Frontiers in Human Neuroscience* **11** 462.
- Lotte, F et al. 2007. “A Review of Classification Algorithms for EEG-Based Brain-Computer Interfaces.” *Journal of neural engineering* 4(2): R1–13.
- von Lüthmann, Alexander, Antonio Ortega-Martinez, David A Boas, and Meryem Ayşe Yücel. 2020. “Using the General Linear Model to Improve Performance in FNIRS Single Trial Analysis and Classification: A Perspective.” *Frontiers in Human Neuroscience* 14: 30.
- Marwan, N. 2013. “Cross Recurrence Plot Toolbox for MATLAB®.” <http://tocsy.pik-potsdam.de/CRPtoolbox/>.
- Marwan, Norbert et al. 2002. “Recurrence-Plot-Based Measures of Complexity and Their Application to Heart-Rate-Variability Data.” *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **66** 026702.
- Marwan, Norbert, Donges, J. F., Zou, Y., Donner, R. V., & Kurths, J. 2009. “Complex Network Approach for Recurrence Analysis of Time Series.” *Physics Letters, Section A: General, Atomic and Solid State Physics*.
- Marwan, Norbert, M. Carmen Romano, Marco Thiel, and Jürgen Kurths. 2007. “Recurrence Plots for the Analysis of Complex Systems.” *Physics Reports* **438** 237-329.
- Marwan, Norbert, and Anja Meinke. 2004. “Extended Recurrence Plot Analysis and Its Application to ERP Data.” In *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering* **14** 761-771.
- McFarland, Dennis J., Lynn M. McCane, Stephen V. David, and Jonathan R. Wolpaw. 1997. “Spatial Filter Selection for EEG-Based Communication.” *Electroencephalography and*

- Clinical Neurophysiology* **103** 386-394.
- McKenna, T M, T A McMullen, and M F Shlesinger. 1994. "The Brain as a Dynamic Physical System." *Neuroscience* 60(3): 587–605.
- Naseer, Noman, and Keum Shik Hong. 2013. "Classification of Functional Near-Infrared Spectroscopy Signals Corresponding to the Right- and Left-Wrist Motor Imagery for Development of a Brain-Computer Interface." *Neuroscience Letters* **553** 84-89.
- Naseer, Noman, and Keum-Shik Hong 2015. "FNIRS-Based Brain-Computer Interfaces: A Review." *Frontiers in Human Neuroscience* **9** 3.
- Naseer, Noman, Farzan M. Noori, Nauman K. Qureshi, and Keum Shik Hong. 2016. "Determining Optimal Feature-Combination for LDA Classification of Functional near-Infrared Spectroscopy Signals in Brain-Computer Interface Application." *Frontiers in Human Neuroscience* **10** 237.
- Ngamga, Eulalie Joelle et al. 2016. "Evaluation of Selected Recurrence Measures in Discriminating Pre-Ictal and Inter-Ictal Periods from Epileptic EEG Data." *Physics Letters, Section A: General, Atomic and Solid State Physics* **380** 1419-1425.
- Nguyen, Thien et al. 2017. "Utilization of a Combined EEG/NIRS System to Predict Driver Drowsiness." *Scientific Reports* 7(1): 43933.
- Pfurtscheller, G., and F. H. Lopes Da Silva. 1999. "Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles." *Clinical Neurophysiology* 110(11): 1842–57.
- Pitsik, Elena et al. 2020. "Motor Execution Reduces EEG Signals Complexity: Recurrence Quantification Analysis Study." *Chaos* **30** 023111.
- Qureshi, Nauman Khalid et al. 2017. "Enhancing Classification Performance of Functional Near-Infrared Spectroscopy-Brain-Computer Interface Using Adaptive Estimation of General Linear Model Coefficients." *Frontiers in Neurorobotics* **11** 33.
- Saadati, Marjan, Jill Nelson, and Hasan Ayaz. 2020a. "Convolutional Neural Network for Hybrid FNIRS-EEG Mental Workload Classification" In *International Conference on Applied Human Factors and Ergonomics*, ed. Hasan Ayaz. Cham: Springer International

- Publishing, 221–32.
- Saadati, Marjan, Jill Nelson, and Hasan Ayaz 2020b. “Multimodal FNIRS-EEG Classification Using Deep Learning Algorithms for Brain-Computer Interfaces Purposes.” In *Advances in Neuroergonomics and Cognitive Engineering*, ed. Hasan Ayaz. Cham: Springer International Publishing, 209–20.
- Santosa, Hendrik, Melissa Jiyoun Hong, and Keum Shik Hong. 2014. “Lateralization of Music Processing with Noises in the Auditory Cortex: An FNIRS Study.” *Frontiers in Behavioral Neuroscience* **8** 418.
- Sassaroli, Angelo, and Sergio Fantini. 2004. “Comment on the Modified Beer-Lambert Law for Scattering Media.” *Physics in Medicine and Biology* **49** N255.
- Schalk, Gerwin et al. 2004. “BCI2000: A General-Purpose Brain-Computer Interface (BCI) System.” *IEEE Transactions on Biomedical Engineering* **51** 1034-1043.
- Shin, Jaeyoung et al. 2018. “Simultaneous Acquisition of EEG and NIRS during Cognitive Tasks for an Open Access Dataset.” *Scientific Data* 5(1): 180003.
- Takens, Floris. 1981. “Detecting Strange Attractors in Turbulence.” *Dynamical Systems and Turbulence, Warwick 1980. Lecture Notes in Mathematics* 898. Springer, Berlin, Heidelberg.
- Venugopalan, Janani, Li Tong, Hamid Reza Hassanzadeh, and May D Wang. 2021. “Multimodal Deep Learning Models for Early Detection of Alzheimer’s Disease Stage.” *Scientific reports* 11(1): 1–13.
- Webber, Jr., C L, and N Marwan. 2015. “Recurrence Quantification Analysis -- Theory and Best Practices.” In *Springer Series: Understanding Complex Systems. Springer International Publishing, Cham Switzerland*.
- Wu, Changwei W et al. 2019. “Indication of Dynamic Neurovascular Coupling from Inconsistency between EEG and fMRI Indices across Sleep–Wake States.” *Sleep and Biological Rhythms* 17(4): 423–31.
- Yin, Xuxian et al. 2015. “A Hybrid BCI Based on EEG and FNIRS Signals Improves the



Performance of Decoding Motor Imagery of Both Force and Speed of Hand Clenching.”  
*Journal of neural engineering* 12(3): 36004.

Zbilut, Joseph P, Nitza Thomasson, and Charles L Webber. 2002. “Recurrence Quantification Analysis as a Tool for Nonlinear Exploration of Nonstationary Cardiac Signals.” *Medical Engineering & Physics* 24(1): 53–60.

## FIGURE LEGENDS

**Fig.1:** Flow diagram of the proposed computational multimodal framework. The multimodal EEG and fNIRS data are simultaneously recorded and independently preprocessed, then the multimodal data is split into training (sub-training and validation) and test sets through a nested cross-validation procedure for hyper-parameter optimization and performance evaluation.

**Fig. 2:** Left: The schematic head montage mode of the EEG-fNIRS sensor-layout for simultaneous electrical-hemodynamic data acquisition. Right: The motor imagery (MI) task experimental protocol.

**Fig. 3:** Visualization of the steps of nonlinear EEG analysis for a representative Rest-MI-Rest (30-sec) task. (a) Representation of the one-dimensional EEG time series from a single channel band-pass filtered in  $\mu$  (8-13 Hz) or  $\beta$  (13-25 Hz) bands used for the 3-dimensional reconstruction of trajectory in phase space illustrated in (b). (c) Recurrence Plots ( $RP_{N \times N}$ ) used to visualize and quantify the recurrence patterns of the 3-dimensional phase space trajectory matrix  $X_{N \times 3}$  calculated based on an  $\varepsilon$ -neighborhood threshold of 5% of the maximum phase space diameter.

**Fig. 4:** Grand-averaged classification accuracy based on the optimized nested inner-loop 5-fold cross validation accuracies across the 5 outer-folds to compare the overall trends of hyper parameter optimization of the 3 types of unimodal features for each subject. The nested inner-loop optimization was performed to optimize the classification parameters for each of the 5 outer-folds (outer-loop) independently (see Table S.1, Table S.2, and Table S.3 in the supplementary section for the details of the optimized parameters corresponding to each type of features for unimodal classification). Each plot in this figure shows the average of the optimized nested 5-fold accuracies averaged across the 5 outer folds for each unimodal (EEG/fNIRS) features for each subject.

**Fig. 5:** Grand-averaged classification accuracy based on the optimized nested inner-loop 5-fold cross validation accuracies across the 5 outer-folds to compare the overall trends of hyper parameter optimization of the 3 types of EEG-fNIRS data fusion for each subject. The nested inner-loop optimization was performed to optimize the classification parameters for each of the 5 outer-folds (outer loop) independently (see Table S.4, Table S.5, and Table S.6 in the supplementary section for the details of the optimized parameters corresponding to each fusion type). Each plot in this figure shows the average of the optimized nested 5-fold accuracies averaged across the 5 outer folds for each multimodal EEG-fNIRS type of data fusion for each subject.

**Fig. 6:** Bar plot showing the defined “Fusion Level” for each type of EEG-fNIRS data fusion (i.e., EEG (linear) - fNIRS, EEG (nonlinear) – fNIRS, and EEG (linear+nonlinear) – fNIRS) for each subject. The figure shows the averaged 5-fold cross-validation classification accuracy” ACC” for each fusion type on top of each bar.

**Fig. 7:** Comparison between the optimum unimodal (i.e., EEG-linear/EEG-nonlinear/fNIRS) and the optimum multimodal (i.e., EEG (linear) - fNIRS/ EEG (nonlinear)-fNIRS/ EEG (linear+nonlinear) - fNIRS) averaged 5-fold cross-validation classification accuracy for each subject.

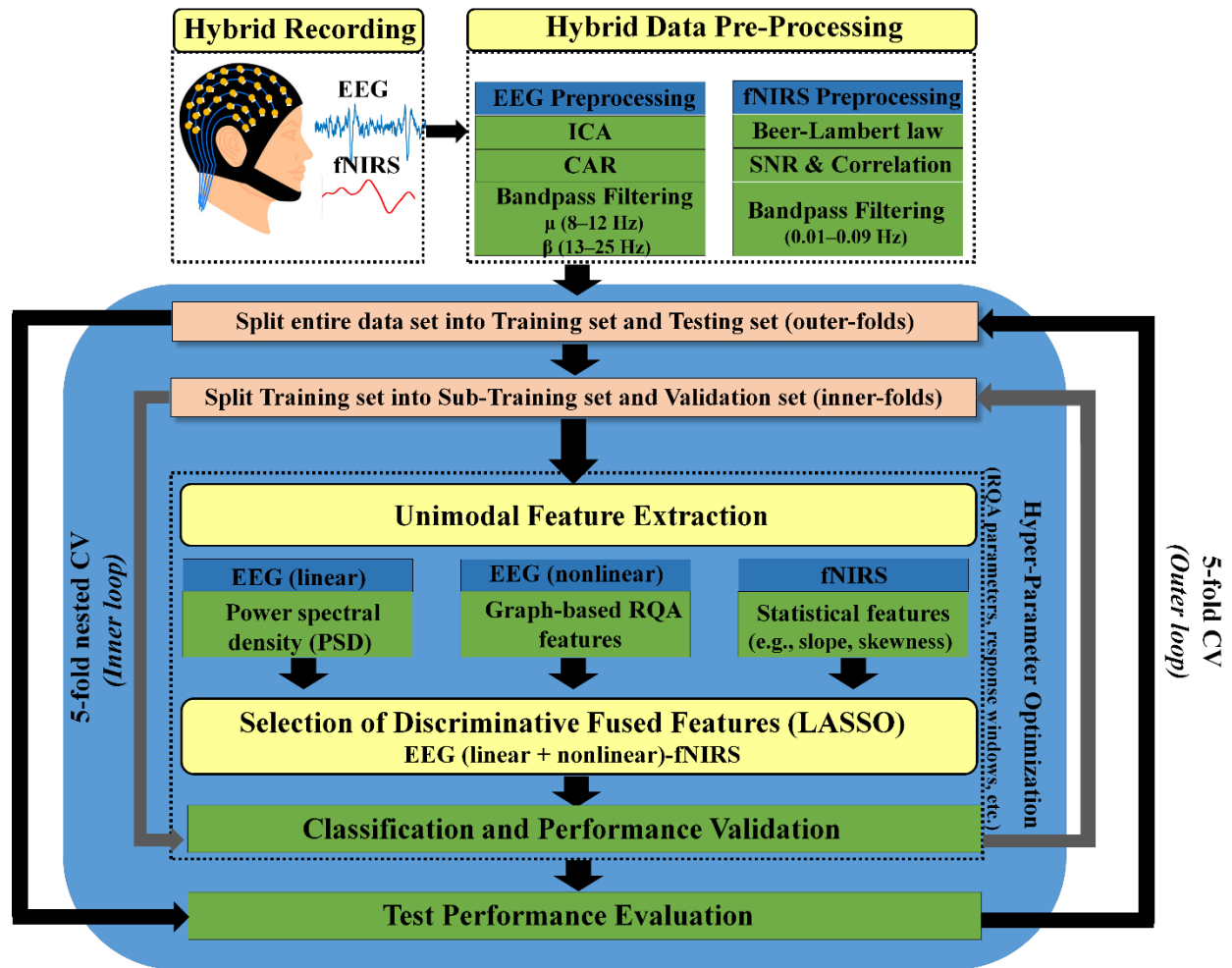
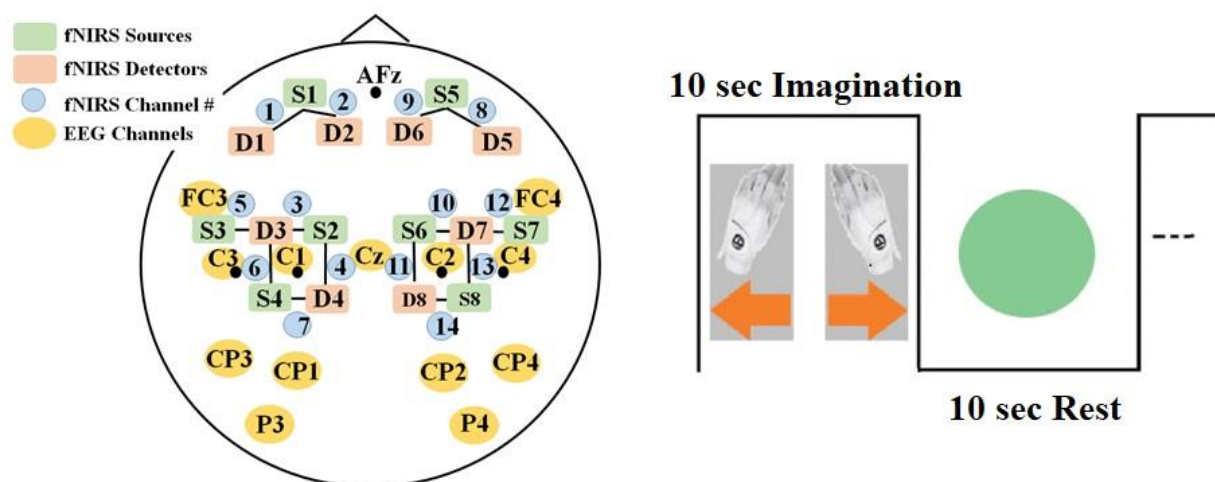


FIGURE 1



**FIGURE 2**

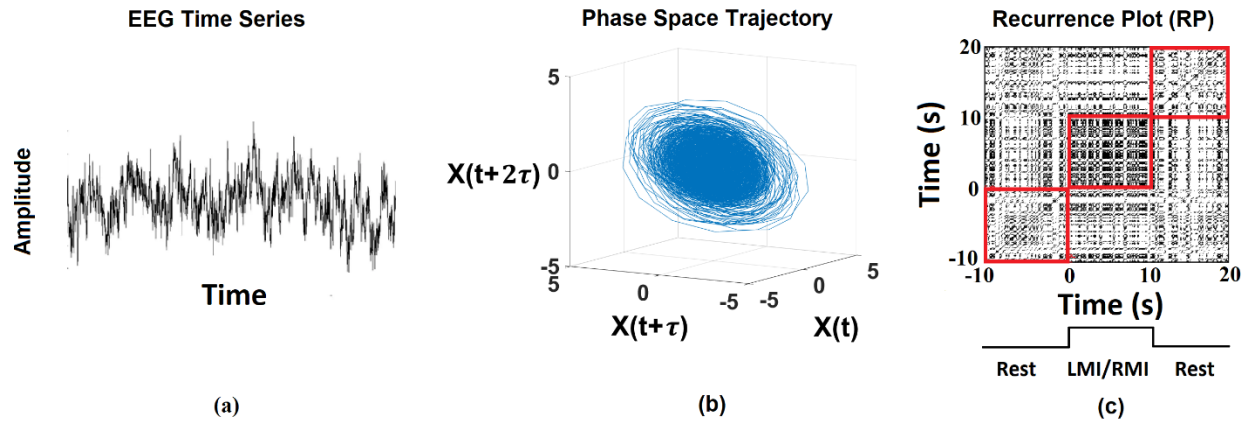
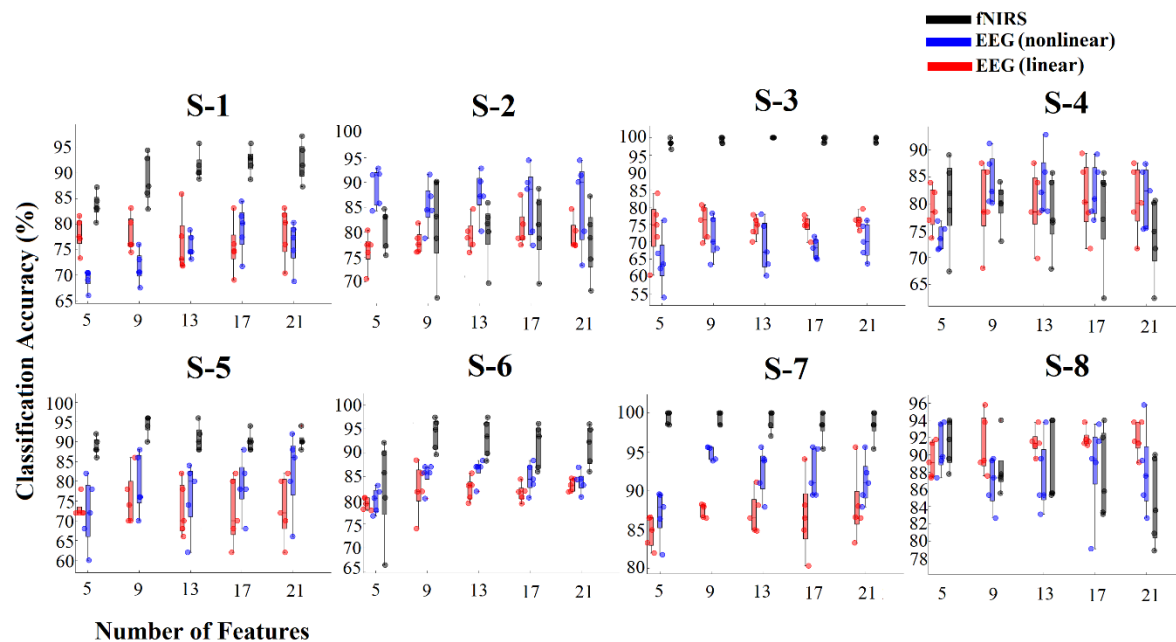


FIGURE 3



**FIGURE 4**

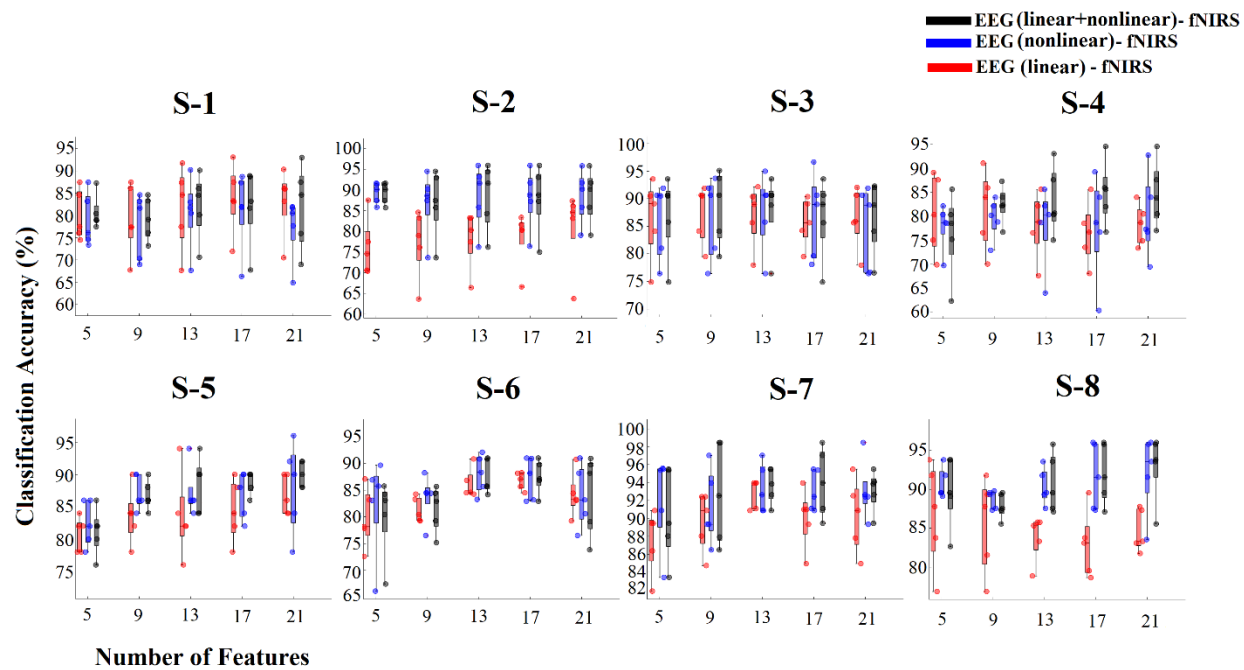
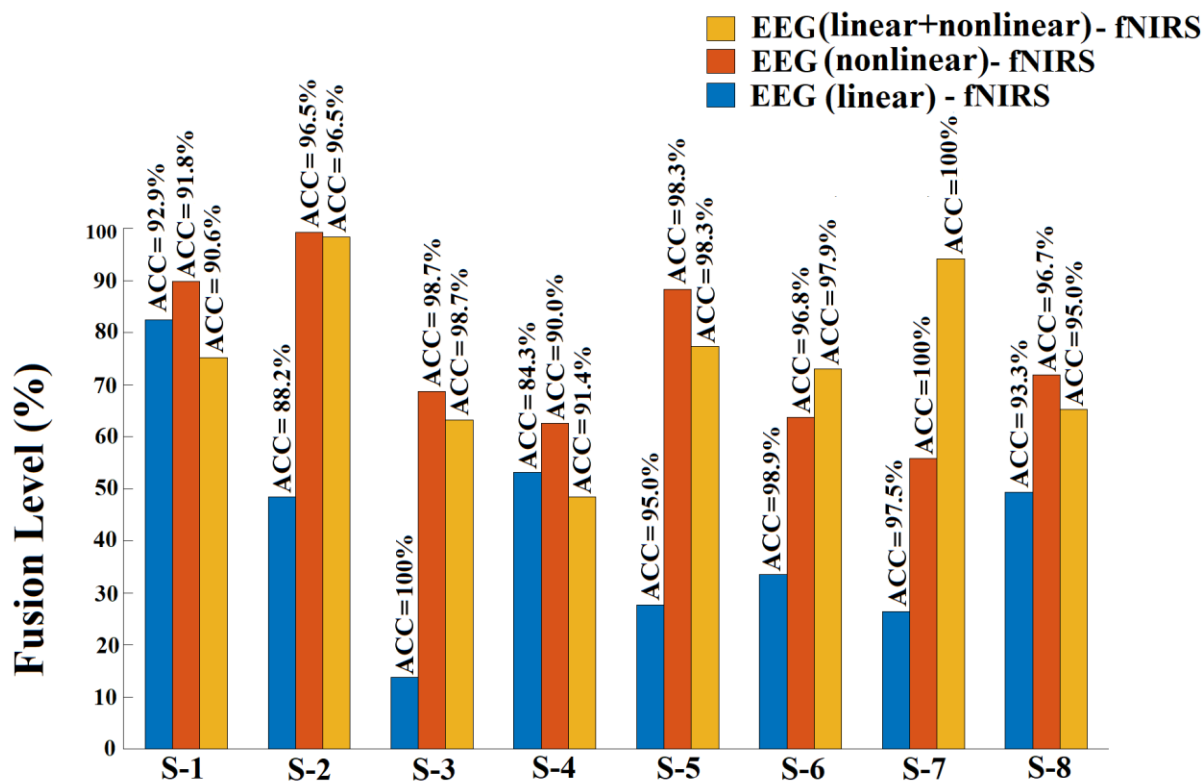
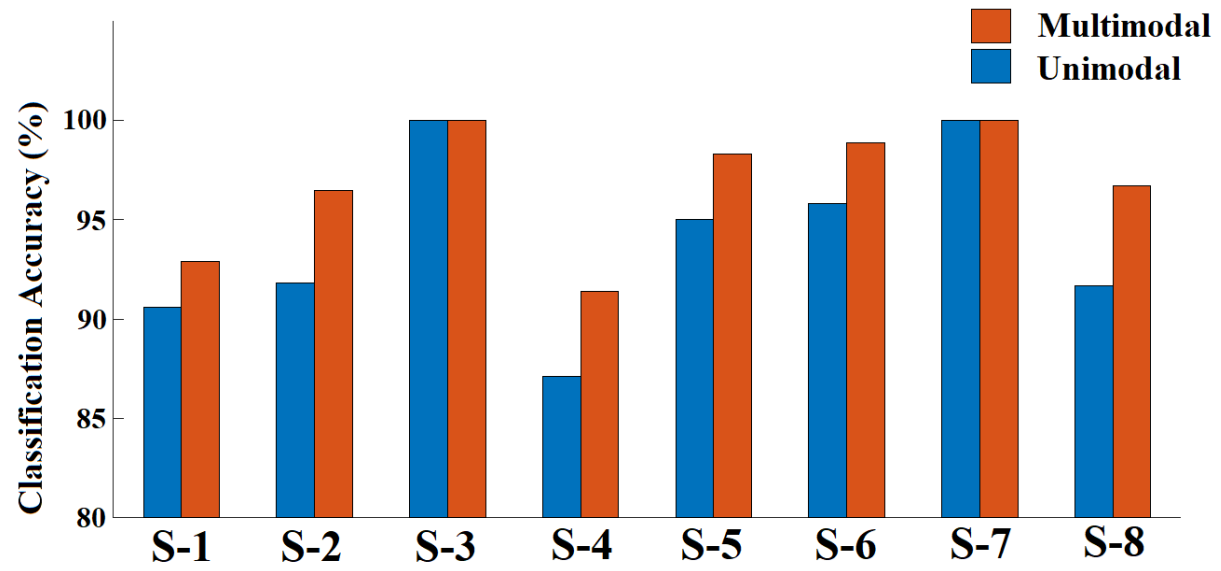


FIGURE 5



**FIGURE 6**





**FIGURE 7**