IET Research Journals

Submission Template for IET Research Journal Papers

IET Journals
The Institution of Engineering and Technology

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Deep Recurrent-Convolutional Neural Network for Classification of Simultaneous EEG-fNIRS Signals

Hamidreza Ghonchi [1], *Mansoor Fateh* [1*], *Vahid Abolghasemi* [2,3], *Saideh Ferdowsi* [2,3], *Mohsen Rezvani* [1]

[1] *Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran*
[2] *Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran*
[3] *School of Computer Science and Electronic Engineering, University of Essex, UK*
* *E-mail: mansoor_fateh@shahroodut.ac.ir*

Abstract: Brain Computer Interface (BCI) is a powerful system for communicating between the brain and outside world. Traditional BCI systems work based on EEG signals only. Recently, researchers have used combination of EEG signals with other signals to improve the performance of BCI systems. Among these signals, the combination of EEG with fNIRS has achieved favorable results. In most studies, only EEGs or fNIRs have been considered as chain-like sequences, and do not consider complex correlations between adjacent signals, neither in time nor channel location. In this paper, a deep neural network model has been introduced to identify the exact objectives of the human brain by introducing temporal and spatial features. The proposed model incorporates spatial relationship between EEG and fNIRS signals. This could be implemented by transforming the sequences of these chain-like signals into hierarchical three-rank tensors. The tests show that the proposed model has a precision of 99.6%.
keywords: EEG, fNIRS, Hybrid-BCI, Deep Learning,Spatial, Temporal.

1 Introduction

In the real world, some people may lose connectivity or movement due to specific diseases. The human brain examines the content of about 100 billion neurons [1, 2]. The spinal cord, as an interface cable, transmits information from organs and nerves to the brain. This information is used to control different parts of the body and their movements. In people with spinal cord injuries, the brain can generate normal signals, but these are not available in the different parts of the body. In intermediate spinal cord injury, several technologies help the injured people to control a wheelchair or a robotic device [3, 4]. These technologies usually employ head movement, eye gaze, etc. However, these methods rely on the control of one of the muscles of the body by the patient. On the other hand, providing effective technologies for people without voluntary muscle control is more challenging [5, 6]. Such people are incapable of speaking and moving, but they can think and decide. Stroke, severe cerebral palsy, Motor Neuron Disease (MND), Amyotrophic Lateral Sclerosis (ALS) and Encephalitis are diseases that can lead to severe motion palsy [7]. With Brain-Computer interface technology (BCI), some people without voluntary muscle control can be more independent in their daily routines.

A Brain-Computer Interface is a communication system that does not depend on the brain's normal output pathways by nerves and muscles [8]. In other words, BCI is an artificial intelligence system that can detect a special set of patterns generated by the brain [9]. A BCI system can focus on mapping, assisting, enhancing and modifying cognitive and sensory-motor functions [10]. The idea of BCI systems is the extraction of the brain's patterns associated with the mental activity (consciously or unconsciously). This system provides communication for a patient without enough control over the motor system (muscles). The proposed method in this paper is to improve the detection of mental activities.

Non-invasive techniques include all mental activity acquisition methods from outside the body's borders [11]. These methods can measure two groups of signals. The first group includes signals that record neural activities in the brain. These signals referred to as electrophysiological signals [12]. These types of signals have a low spatial resolution and a very high temporal resolution (milliseconds). EEG is one of the important signals in this group. The second group contains signals that record hemodynamic activity. These types of signals measure the level of oxygen in the blood [13] and offer a high temporal resolution [14]. Functional near-infrared spectroscopy (fNIRS) is one of signals in this group. The main advantage of fNIRS is the portability of the device. In this paper, the combination of EEG and fNIRS has been used.

One of the important challenges in designing BCI systems is what signal to choose; electrophysiological or hemodynamic signals. So far, many studies have been done for BCI systems based on electrophysiological signals only [15–19], or based on hemodynamic signals only [20–22]. In addition, there are several recent works addressing, both [9, 23–26]. However, the electrophysiological methods generally suffer from low spatial resolution. On the other hand, the use of hemodynamic methods alone suffers from low temporal resolution [23]. Meanwhile, it is expected that combination of EEG with fNIRS provide appropriate results [23, 27, 28] and can reduce the stand-alone limitations for each modality [29–31]. Another reason for feasibility of using simultaneous EEG and fNIRS signals is that these do not generally interfere. In addition, low cost, low noise, and portability make the use of both EEG and fNIRS signals more functional. But, in this paper, a new approach for combining these two signals is presented.

Furthermore, there are several other challenges. The first challenge is noise interference with signals. Part of the common noise in the signals is frequency interference with other equipment and the inappropriate connection of the electrodes. In these signals, physiological activities cause noise. These activities are associated with reduced signal-to-noise rate, for example, blinking, muscle movement, and heart rate [32]. The preprocessing algorithms are useful to mitigate noise in signals. The second challenge is ambiguity in the relationship between signals and their brain-related intents. In fact, mental intents cannot be directly observed from the signals [32]. Precise analysis and classification of signals can resolve this issue. For this purpose, a deep learning classification approach is proposed in this paper. The third challenge is the extraction of features by the user in order to detect mental intents. Implementation of this step normally requires heavy processings [33]. In most of the processing step in the BCI systems, there are two steps, noise

removal [34] and feature extraction [35]. These steps are impractical for BCI systems due to time-consuming and dependency on the level of professional knowledge. In 2019, She et al. [68] have proposed the new deep learning method of hierarchical semi-supervised extreme learning machine (HSS-ELM) to classify MI-based EEG data. They proved their algorithm can utilize unlabeled data. They also used the deep structure algorithm for extracting features and tried to solve extracting features by user. On the other hand, In 2017, Li et al. [69] tried to classify hybrid EEG-fNIRS using SVM technique. They considered different number of channels but they did not consider the importance of position of all channels in human's head. In 2017, Lu et al. [70] tried to classify MI-based EEG using deep learning. They changed time-series EEG signals to frequency domain. They proved that frequency domain is more suitable than time-series. In this paper, feature extraction is done using the deep learning algorithm. Feature engineering in deep learning algorithms automatically extracts the appropriate features from the signal. Thanks to deep learning method, we mitigate the burden of the feature extraction step.

One of the most important issues tackled in this paper is that accurate classification of the raw signal using the deep learning algorithm is considerably low. In this paper, we consider both spatial (location of electrodes on the scalp) and temporal (samples of EEG and fNIRS) information in order to improve the accuracy. These properties have merely studied in the past [36, 37]. In this paper, we focus on an appropriate combination of these properties and at the same time combining EEG and fNIRS signals. In addition, deep learning algorithm is adopted to classify the signals with high accuracy and detecting the user's mental intents with the least error.

In the proposed method, the raw biological signals are converted to three-rank tensors using the knowledge about the position of electrodes. These three-rank tensors show brain topography in each time instance. Then, a sliding window is used to create data clips. The sliding window aims to exploit the temporal correlations in biological signals. In each clip, there is spatial and temporal feature simultaneously. After these arrangements, deep learning is used to train the model based on the obtained three-rank tensors. In this paper, we could improve the diagnostic accuracy by improving the preprocessing step and adapting deep learning algorithm specifically for EEG-fNIRS combination purpose.

In the remainder of this paper, the basic techniques used in different parts of the BCI system are reviewed in the Section 2. The related works are reviewed in Section 3. In Section 4, the proposed algorithm for data preprocessing and combining and converting signals into three-rank structures are presented. Then, the proposed model based on deep learning is described. In Section 5, the results of the experiments are described. Finally, the conclusion is drawn in Section 6.

2 Background

In this section, the basic techniques used in different parts of a typical BCI system are expressed.

2.1 The structure of BCI systems

The structure of BCI systems is important. In most of BCI systems, it is important to design a comfortable and easy-to-use interface for users. In the following, we describe the structure of BCI systems.

A BCI system requires careful and complex design. Improvement of signal acquisition technique, the development of feature extraction methods and translation algorithms in order to create command for executive in hardware devices, the development of hardware, improvement of operational protocols and user training strategies are design requirements of a BCI system [38–40]. The general steps of a BCI system are shown in Figure 1. Brain signals are measured and converted to control signals for different applications which provide feedback to the user. The first step in this system is data/signal acquisition. In general, the signal processing step includes preprocessing, feature extraction and classification. Preprocessing step includes any classical algorithm with the aim of signals enhancement, for example, noise removal. Feature extraction step includes those techniques to extract features from input signals. Classification step is employed to classify extracted features. The ultimate output is a control signal that is transferred to an application and then returns feedback to the user. Such feedback in general BCI systems can be either Visual or Auditory. These steps are briefly outlined in a block diagram in Figure 1.

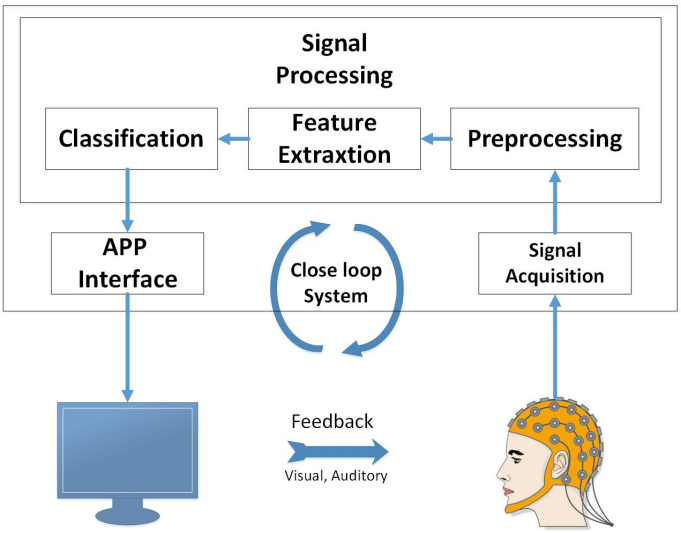

Fig. 1: BCI system steps.

2.2 The Advantages of the combination of EEG and fNIRS

The combination of biological signals has always been a problem. But sometimes, combining these types of signals have good benefits. In this paper, the two signals EEG and fNIRS have been combined. In the following, we will explain the advantages of combining these two signals.

In recent years, BCI studies have shown that combining EEG techniques with other neuroimaging technologies can increase performance [41]. As mentioned in the introduction, simultaneous recording of EEG and fNIRS is associated with challenges. On the other hand, the existence of electromagnetic waves in the other neuroimaging technologies has a destructive effect on the EEG signal during simultaneous recording. In contrast, the fNIRS recording system is well isolated from the EEG part and is suitable for simultaneous recording with electrophysiological signals (due to the use of near-infrared light). In fact, electrophysiological signals and optical signals do not generally interfere. Long-term recordings EEG-fNIRS, interracial epileptic discharges (IED) or epilepsy [26, 42–44] are further investigated using hemodynamic changes locked at the time of IEDs. Combination of EEG and fNIRS can be used for detection of brain conditions around 5 minutes prior occurring some diseases [44]. The combination of event-related potentials (ERP) and hemodynamic responses from fNIRS are also used to study the visual cortex [25, 45].

3 Related Work

In past years, there was much research on EEG or fNIRS individually, and recently, researches have been focused on combining these two signals for using in a hybrid BCI. Different methods and features have been used to classify these kinds of signals.

3.1 EEG-based BCI

Kosmyna et al. [46] presented a multi-dimensional feature space for EEG-based BCI systems. This feature space includes four axes and

nine sub-axes containing 41 points in total. These 4 axes include 4 questions; "when" (temporal aspects), "how" (median aspects), "what" (content aspects) and "where" (spatial aspects). Each of these questions includes nine sub-axes. They have incorporated this feature space under the BCI and HCI systems.

Yang et al. [36] attempted to solve the problem of lack of sufficient information, including spatial and temporal information in EEG signals. In order to solve this problem, they used an LSTM (Long-Short Term Memory) network with a convolutional network. In this method, spatial and temporal information are simultaneously applied to MI-EEG (Motor Imagery EEG) signals. To achieve a higher learning rate, a batch normalization method has been used. In this method, four steps are provided for data preprocessing. Different experiments have been done to evaluate the performance of deep learning network architecture. The results have shown that this method was a powerful and useful model for multiclass classification among many different methods. The deep learning network model is compared with conventional classifier algorithms such as LDA (Linear Discriminant Analysis), KNN (K-Nearest Neighbor), NB (Naïve Bayes), SVM, LSTM and CNN (Convolutional Neural Network) separately. The problems with this article are that they used frequency features. Frequency feature extraction needs a large continuous sampling period, while the motor imagery tasks are periodic short duration brain activities. So extracting the frequency features may damage the temporal information.

Zeng et al. [47] presented a method using the combination of convolutional neural network and deep residual learning in order to predict EEG signals based on the mental states of drivers. Hence, two classification models called EEG-Conv and EEG-Conv-R have been presented. In this method, eye blink has been removed from EEG signals using ICA technique. Then, a bandpass filter has been applied with the cut-off frequencies between 1 to 40 Hz. Z-score normalization has also been used to normalize the data. The results have shown that both models achieve a good performance in mental state classification. EEG-Conv-R is more suitable for prediction of mental state among subjects and EEG-Conv-R converges faster. They used raw signals without any spatial preprocessing except z-score normalization. The raw EEG signal as the chain-like signal is not good for classification.

Zhang et al. [37] classified the humans mental goals using a convolutional-recurrent neural network. In this method, the spatial-temporal features of the EEG signal have been used. In the proposed model of this paper, the spatial correlation between EEG signals has been transformed into two-dimensional tensors. In this model, an LSTM network has been proposed for the extraction of temporal correlations within EEG signals. The corresponding dataset in this paper was based on MI-EEG with 108 subjects and about 3.145.160 samples. This deep learning model achieved an accuracy of about 98.3%.

3.2 fNIRS-based BCI

Peng et al. [20] proposed a multiclass classification on MI-fNIRS signals. Ten healthy subjects have been used to move an object by motion imagery. In this method, prefrontal cortex signals have been used. A combination of Ensemble Empirical Mode Decomposition (EEMD) and ICA (Independent Component Analysis) has been used to remove noise in signals. Signal Average characteristics (SA) have been used for the input of SVM and LDA classifications. The classification accuracy using SVM, and in the HbO2 (oxygenated hemoglobin) mode and 8 to 21 seconds window was higher than LDA. The classification accuracy for all four directions, up-down and left-right motions was 40.55, 73.05 and 70.7 percent, respectively.

The combined motion of the right and left arm was classified using a support vector machine (SVM) in [21]. In this study, fNIRS signals were used. The results of this study showed that the combination of fNIRS-based features achieved a maximum accuracy of 76.67%. These results have shown that distinct patterns in hemodynamic responses from the right and left-hand movements can be used to develop a BCI system.

Ho et al. [22] used deep learning to detect rest and action states. The noise in the recorded fNIRS data has been removed by a bandpass filter with the cut-off frequencies between 0.3 and 3 Hz. In this method, at first, apparent changes in the concentrations of HbO and HbR have been identified for both phases. Then, a method for detecting these parameters is presented. They could increase the classification accuracy using deep learning methods for up to 84.25%.

The main difference between the articles in sections 3.1 and 3.2 and our proposed method is combining two signals. Inadequate information in each signal can be completed with the help of another signal. Therefore, the combination of two signals and the use of both signals in different ways can increase the information and thus increase the classification accuracy. Some of the articles mentioned above did not use deep learning to classify their data or tried to extract various features with the help of feature extraction techniques. It was also the other difference in methods with our proposed method. Extracting the features manually and by a human can cause the error, and it is possible that extracted features are not suitable for the proposed system, and this reduces the classification accuracy.

3.3 hybrid BCI

Khan et al. [48] presented a new classifier using a modified phase vector diagram and power of EEG signals for predicting hemodynamic responses. In this paper, EEG and fNIRS signals have been recorded simultaneously for a motion task (thumb movement). In this paper, the achieved accuracy in 0 to 1.5 seconds window is 86%.

Zhang et al. [49] have designed a technique for combining EEG and fNIRS in a BCI system. This method has resolved some existing issues in EEG- based BCI systems, such as noise and low classification accuracy. This system is based on punching by the participant. In this work, combined features based on EEG wavelet coefficient and fNIRS gradient has been proposed. The classification accuracy has increased by 3% compared to EEG features-only and 9% compared to the fNIRS features-only. The results of this paper have shown that incorporating fNIRS information for feature design could greatly increase the performance.

Hong et al. [24] presented a Brain-Computer Interface (BCI) framework based on EEG and fNIRS to be used by ALS (Amyotrophic lateral sclerosis) patients. In this paper, brain activities, channel selection methods, feature extraction algorithms, and classification algorithms were investigated. Various types of cognitive and motor disorders have been classified to evaluate the BCI system. Also, in this paper, the brain activity of hemodynamic signals was investigated. Mental arithmetic and word formation are workable for patients with ALS. In this method, the spatial area of the brain was interest. In general, proper detection of the activated brain area and features have improved the classification accuracy.

Chiarelli et al. [50] improved performance of EEG- and fNIRS-based hybrid BCI system using deep learning approaches. In this paper, the classification of two imaginary movements of the right and left hand was investigated. They conducted experiments on 15 subjects. Each recording time was one second, and the duration of the experiment was 10 minutes. The recorded signals from EEG device have been divided into one-second windows and then have been filtered with cut-off frequencies between 8 to 30 Hz. Then the ERD/ERS features have been extracted from signal power. The recorded signals from the fNIRS device have been converted to oxy- and deoxy-hemoglobin concentrations using the modified Lambert beer law algorithm. In this paper, the classification has been carried out using a DNN (Deep Neural Network). The classification accuracy was 83.28% for the combination of EEG and fNIRS. The main difference between this paper and our proposed method is the structure of the deep learning model and extracted features. This article uses deep learning of DNN but does not take into account spatial and temporal features of two signals.

Muhammad Jawad et al. [9] used hybrid EEG-fNIRS for decoding eight brain commands. For the fNIRS system and EEG system, they used prefrontal cortex, and around the frontal, parietal and visual cortexes, respectively. In this paper, the related tasks for recording

fNIRS signals were mental arithmetic, mental counter, mental rotation, and word formation. In recorded fNIRS signals, peak signal, minimum value and mean value of HbO in two-seconds windows have been selected as features for classification. The related tasks for recording EEG signals were blinking twice, blinking three times, eye movement in the up-down direction and eye movement in the left-right direction. In recorded EEG signals, peak and mean signals in one-second windows have been selected as features for classification. The eight generated commands have been used to control a quadcopter in the open space. The average accuracy for four commands related to fNIRS signals was 75.% and was 86% for four commands related to EEG signals. In related work, an fNIRS based method was provided for controlling quadcopter [51]. However, it has less efficiency than that of the proposed method in [9].

Shin et al. [52] recorded EEG-fNIRS signals using the tasks of mental arithmetic (MA) and word chain (WC), simultaneously. Each recording time was 5 seconds which was less than previous hBCI studies. In this paper, a shrinkage linear discriminant analysis (sLDA) has been used for the separation of both MA and WC from baselines. The achieved separation accuracy in the offline model was 90% and 85.8%, respectively. The corresponding results in the online model were 85.5% and 79.5%, respectively. In most cases, these accuracy values were significantly higher than EEG- or fNIRS-based BCIs.

There have been also reported some related, EEG-NIRS works in [53–59] where emotion diagnoses were investigated. As these works are out of the scope of the current paper, we omit a detailed description of these papers.

In the works mentioned above, spatial and temporal information was not simultaneously used to combine EEG and fNIRS. In this paper, the effect of using this information for classification of combined EEG-fNIRS signals was investigated and compared with previous works.

4 Proposed method

In this section, our focus is on preprocessing and classification. In the preprocessing step, raw signals are transformed into three-rank tensors, and in the classification step, a deep learning-based model is proposed to diagnose mental intents.

4.1 Preprocessing

Preprocessing is one of the most important steps in BCI systems. This step includes algorithms for changing the signal to prepare it for the classification step and removing noise. In this section, we describe how we convert signals from chain-like to three-rank tensors to prepare for classification.

Mental activities related to motor imagery (MI) appear in 8 to 30 Hz (Alpha and Beta bands) frequency band of EEG signals. On the other hand, the initial sampling rate of fNIRS data is 10 Hz and for EEG data is 200 Hz. These two types of data must become compatible before any processing. Hence, the sampling rate of EEG signals was reduced to 128 Hz, and the sampling rate of fNIRS data was increased to 128 Hz. Also, by adjusting the sampling rates in both signals, the processing burden of EEG signals is significantly reduced. Important information for both EEG and fNIRS are burried in the frequency range up to 64 Hz [66]. Therefore, based on Nyquist theorem, the frequency of 128 Hz chose as the minimum sampling rate to avoid losing information as well as keeping the computational burden low. In this paper, a Fourier transform is used for upsampling fNIRS signals. To downsample, the time-series EEG signals are transformed to the frequency domain using Fourier Transform, as showed in Equation 1. Then, the second and third groups of $N/4$ elements (which correspond to the half with the highest frequency components) are deleted. Then, the frequency domain is transformed back to time domain (Equation 2). To upsample, like downsample, the time-series fNIRS signals are transformed to the frequency domain using Fourier Transform and $N/2$ zeros are added at the end. Then, the frequency domain is transformed back to time domain, which N is length of time-series signal.

$$X(e^{jw}) = \sum_{n=-\infty}^{+\infty} x[n]e^{-jwn} \tag{1}$$

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{+\pi} X(e^{jw})e^{jwn}dw \tag{2}$$

Figure 2 shows how to record the simultaneous EEG-fNIRS signals followed by the proposed method used in this paper. EEG and fNIRS electrodes record brain activities when the subjects are asked to imagine a particular action. Typically, the raw data obtained from the EEG recording system at time t is a one-dimensional vector $e_t = [S_t^1, S_t^2, ..., S_t^n]^T$ where S_t^n is the recorded EEG signal sample from the nth channel and at time t and n is the number of EEG channel. Similarly, the raw data obtained from fNIRS recording system at time t is a vector denoted by $f_t = [(E_t^1, E_t^{'1}), (E_t^2, E_t^{'2}), ..., (E_t^m, E_t^{'m})]$ where E_t^m is the oxyhemoglobin sample in the fNIRS signal from mth channel and at time t, $E_t^{'m}$ is the deoxy-hemoglobin sample in the fNIRS signal from mth channel and at time t, and m is the number of fNIRS channels. A description of all the parameters available throughout this paper is given in Table 1.

For a simpler description of the process, we consider each raw signal from time t to $(N+1)$. By doing this, a number of vectors of length $(N+1)$ is extracted within the range $[t, t+N]$ from raw signals. This procedure is carried out for n EEG channels and m fNIRS signals channels.

Figure 3 is shown the raw EEG signal. As a matrix, every channel is considered as a row in this matrix. Columns in this matrix include the samples within each channel. Each channel is only limited to two upper and lower channels. To clarify this issue, each row in this matrix has two vertical neighbours. One neighbour is an upper row, and another neighbour is a lower row in the matrix. Due to the importance of spatial correlation of electrodes in EEG, spatial correlation of electrodes in fNIRS, and the relationship between different regions of the brain with different activities, Equation 3 is used as a conversion function. In [67], they convert the 62-channel EEG placement to two-dimensional map. In this paper, this conversion method is used to convert the combination of the EEG and fNIRS time series into three-rank tensors. The channel positions in the human head in the system 10-5 [60] are shown in Figure 4. Equation 3 (matrix m_t) is designed according to Figure 4.

Table 1 Description of parameters.

Parameters	Descriptions
e_t	A vector associated with EEG signal at time t
s_t^n	The recorded EEG signal sample from the nth channel and at time t
f_t	A vector associated with fNIRS at time t
E_t^m	The oxyhemoglobin sample in fNIRS signal from mth channel and at time t
$E_t^{'m}$	The deoxy-hemoglobin sample in fNIRS signal from mth channel and at time t
n	The number of EEG channel
m	The number of fNIRS channel
m_t	Three-dimensional meshes
SW_j	The generated clips from three-rank tensors at a temporal window
SF_j	The extracted temporal featured from CNN network
TF_j	The extracted spatial features from RNN network
h_t	Hidden cells of LSTM layer at time t
x_t	Input at time t

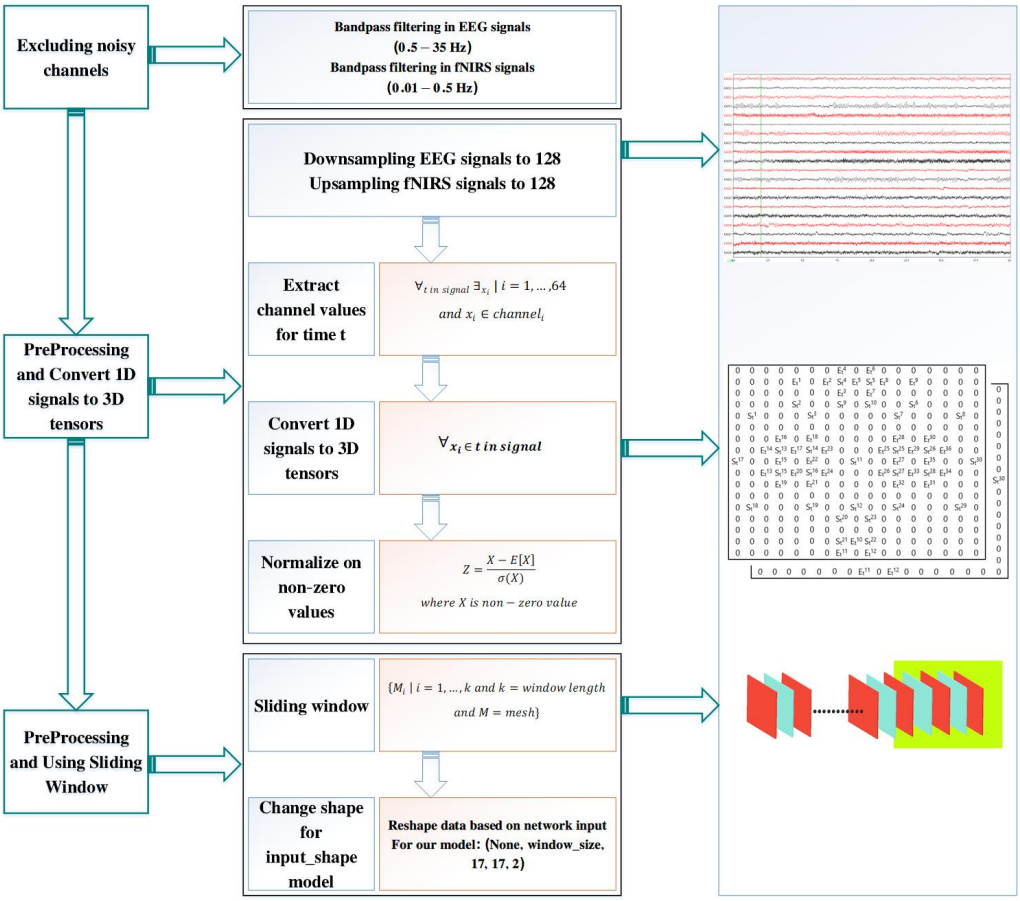

Fig. 2: EEG and fNIRS signal acquisition and preprocessing steps.

$$Tensor\_map(e_t, f_t) =$$

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t^4 & 0 & E_t^6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & E_t^1 & 0 & E_t^2 & S_t^4 & E_t^5 & S_t^5 & E_t^8 & 0 & E_t^9 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t^3 & 0 & E_t^7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & S_t^2 & 0 & 0 & S_t^9 & 0 & S_t^{10} & 0 & 0 & S_t^6 & 0 & 0 & 0 & 0 \\
0 & S_t^1 & 0 & 0 & 0 & S_t^3 & 0 & 0 & 0 & 0 & 0 & S_t^7 & 0 & 0 & 0 & S_t^8 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & E_t^{16} & 0 & E_t^{18} & 0 & 0 & 0 & 0 & 0 & E_t^{28} & 0 & E_t^{30} & 0 & 0 & 0 \\
0 & 0 & E_t^{14} & S_t^{13} & E_t^{17} & S_t^{14} & E_t^{23} & 0 & 0 & 0 & E_t^{25} & S_t^{25} & E_t^{29} & S_t^{26} & E_t^{36} & 0 & 0 \\
S_t^{17} & 0 & 0 & E_t^{15} & 0 & E_t^{22} & 0 & 0 & S_t^{11} & 0 & 0 & E_t^{27} & 0 & E_t^{35} & 0 & 0 & S_t^{30} \\
0 & 0 & E_t^{13} & S_t^{15} & E_t^{20} & S_t^{16} & E_t^{24} & 0 & 0 & 0 & E_t^{26} & S_t^{27} & E_t^{33} & S_t^{28} & E_t^{34} & 0 & 0 \\
0 & 0 & 0 & E_t^{19} & 0 & E_t^{21} & 0 & 0 & 0 & 0 & 0 & E_t^{32} & 0 & E_t^{31} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & S_t^{18} & 0 & 0 & 0 & S_t^{19} & 0 & 0 & S_t^{12} & 0 & 0 & S_t^{24} & 0 & 0 & 0 & S_t^{29} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & S_t^{20} & 0 & S_t^{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & S_t^{21} & E_t^{10} & S_t^{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t^{11} & 0 & E_t^{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & E_t'^{11} & 0 & E_t'^{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ S_t^{30} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} \tag{3}$$

This three-rank tensor denoted by m_t corresponds to the spatial information of the electrodes in the recording system. This three-rank tensor is composed of two matrixes. Each matrix is known as a facet. In three-rank tensor, there are two facets at each time instance t. In each time instance t, the first facet, includes EEG-fNIRS data with low wavelength values, and the second facet contains EEG-fNIRS data with high wavelength values.

As seen in Equation 3, the null positions (not corresponding to any electrode position) were shown by zero. These positions do not affect the training of the neural network because these are similar to the background. Using this information, the raw data vector series $[e_t, e_{t+1}, ..., e_{t+N}]$ and $[f_t, f_{(t+1)}, ..., f_{(t+N)}]$ are converted into three-rank tensor $[m_t, m_{(t+1)}, ..., m_{(t+N)}]$ where e_t is a vector associated with EEG signal at time t, f_t is a vector associated with fNIRS at time t and m_t is a three-rank tenor. After this stage, the non-zero data tensor are normalized using z-score. Each three-rank data tensor includes the spatial information of brain activities. The sliding window method is used to add temporal information to these three-rank tensors.

The sliding window method is used to convert three-rank tensors into data clips. All clips have the same length. Each clip has a 50% overlap with the neighboring clips. Each data clip denoted as SW_j is constructed in the form of Equation 4.

$$SW_j = [m_t, m_{t+1}, ..., m_{t+k-1}] \tag{4}$$

Where k is the window size and $j = 1, 2, ..., q$. Each SW_j includes both spatial and temporal feature and now it's ready to train the proposed model.

The purpose of this paper is to develop an effective model for organizing a set of human intents using data clips. This model predicts human intents over the data clips.

4.2 *The proposed Recurrent Convolutional Neural Network*

The RCNN architecture involves using Convolutional Neural Network (CNN) [61] layers for feature extraction of input data combined with LSTMs [62] to support sequence prediction. RCNNs have been developed for visual time series prediction problems and the application of generating textual descriptions from sequences of images (e.g. videos). RCNNs are a class of models that is both spatially and temporally deep and has the flexibility to be applied to a variety of vision tasks involving sequential inputs and outputs [63]. Overall, these models can read a sequence of image inputs, such as a video, and generate a prediction.

Figure 5 shows our proposed structure of the RCNN-based EEG-fNIRS classifier. This structure is used to give spatial and temporal features from EEG-fNIRS time series. This structure includes

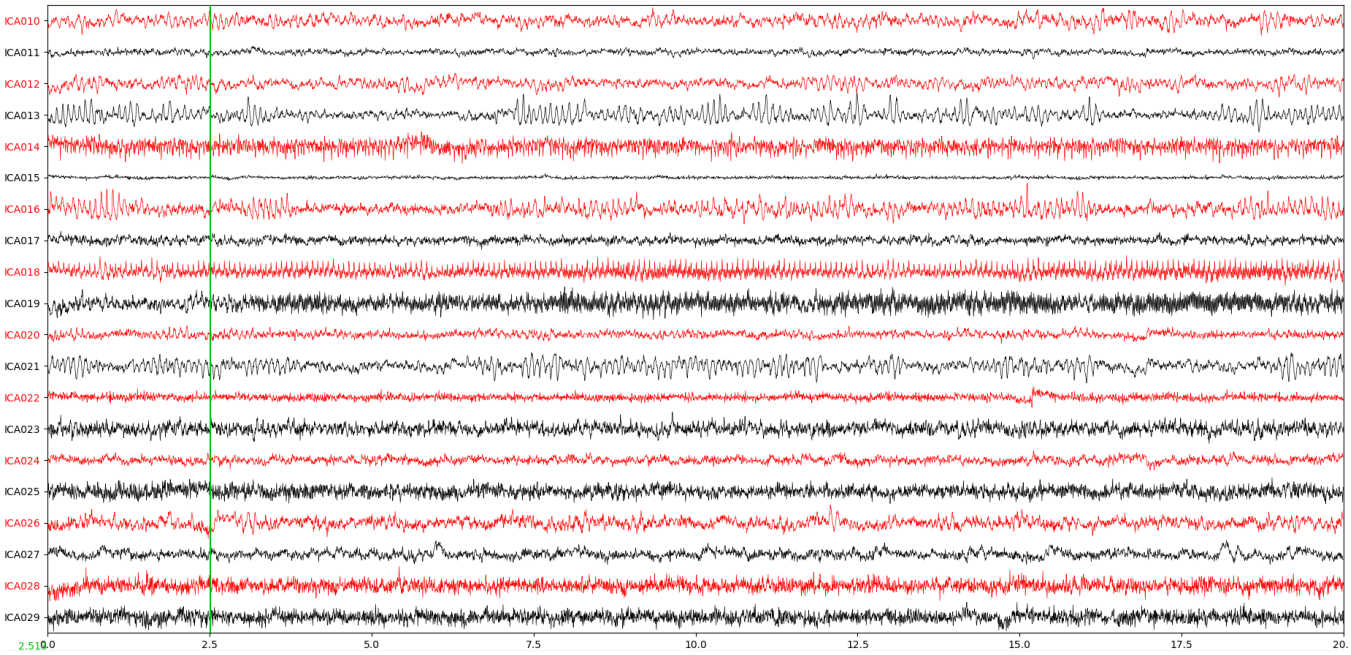

Fig. 3: Raw EEG signal.

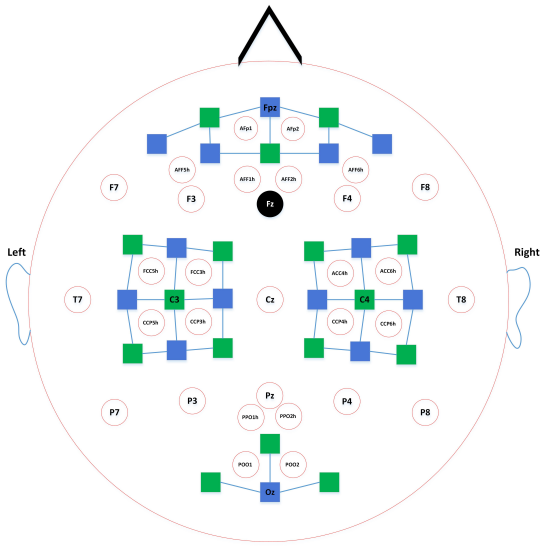

Fig. 4: Placement of EEG electrodes in 10-5 system (red and black (ground) circles) and NIRS sources (blue squares) and detectors (green squares). blue solid lines denote NIRS channels.

eight layers. These layers include three convolutional layers, two fully connected (dense) layers, two LSTM layers, and one output layer. The input of this structure is the preprocessed clips of three-rank tensors (SW_j). Finally, these clips are converted to four-dimensional structures. These four-dimensional structures convey both spatial and temporal information. First, the spatial features (SF_j) are extracted using CNN layers from each three-rank tensor. Then, a fully-connected (dense) layer receives the extracted spatial features sequence SF_j and feed them to LSTM layers to extract temporal features TF_j. Finally, a fully-connected (dense) layer receives the temporal features from LSTM last output layer, and then feeds the Softmax layer to predict.

In the proposed model, there are three two-dimensional convolutional layers with the same 3×3 kernel size. As a result, these three layers extract spatial features from the input three-rank tensors. The size of each data tensor is $h \times w$. The input to the proposed model, SW_j, contains $s-1$ elements where s is the length of the sliding window. The first convolutional layer includes 32 feature maps. This layer generates 32 feature maps of size $h \times w$. This number has doubled for the next convolutional layer. These values are achieved with numerous examinations. Finally, there are 128 feature maps in the last layer. After three convolutional layers, a fully-connected (dense) layer with 1024 neurons are obtained. The purpose of this layer is to extract high-level signal features. This layer extracts the final spatial features SF_j by receiving 128 feature maps of last convolutional layer. The fully-connected (dense) layer has been used to feed the RNN.

The spatial features sequence (SF_j) is the input to RNN network to extract temporal features (TF_j). In this paper, two LSTM layers are used to extract temporal features. The second LSTM layer input is the extracted temporal features from the previous layer. This layer extracts temporal features using x_t and h_{t-1} value. Then, this layer transfers the obtained information to the current state h_t and affects the final output. Finally, the second LSTM layer is a sequence of hidden states in its previous layer $[h_t, h_{t+1}, ..., h_{t+s-1}]$. The purpose of this research is to find the brains commands during the sliding window period. Therefore, the extracted temporal features in the last LSTM layer are used for analysis. These features have been extracted from the windowed samples SW_j. In fact, these features are brains commands the during the windowing period. The last LSTM layer output is given to fully-connected (a dense) layer to extract the final temporal features sequence TF_j. Finally, the Softmax classification predicts the final probability for each data class. Proper architecture design increases the performance of deep learning network. In this paper, we have tried to provide an appropriate architecture for finding the brains commands.

In addition to designing the proper architecture, choosing the optimal parameters has a great influence on the correct identification of the brains command. There are a lot of optimizer functions for parameter optimization. However, based on our experiments Adams [64] optimizer function with a learning rate of 0.0001 demonstrates the best performance. Also, for all layers except the last one, the Rectified Linear Units (ReLU) is used. Softmax function was used for the last layer. These choices along with other parameters have been achieved after numerous examinations and studies on various activation functions.

Since various papers have examined different structures of the deep neural network, we have used the best parameters in these models. The best accuracy in most of the methods used in this field is related to [37]. In this paper, the RCNN hybrid model is used to train

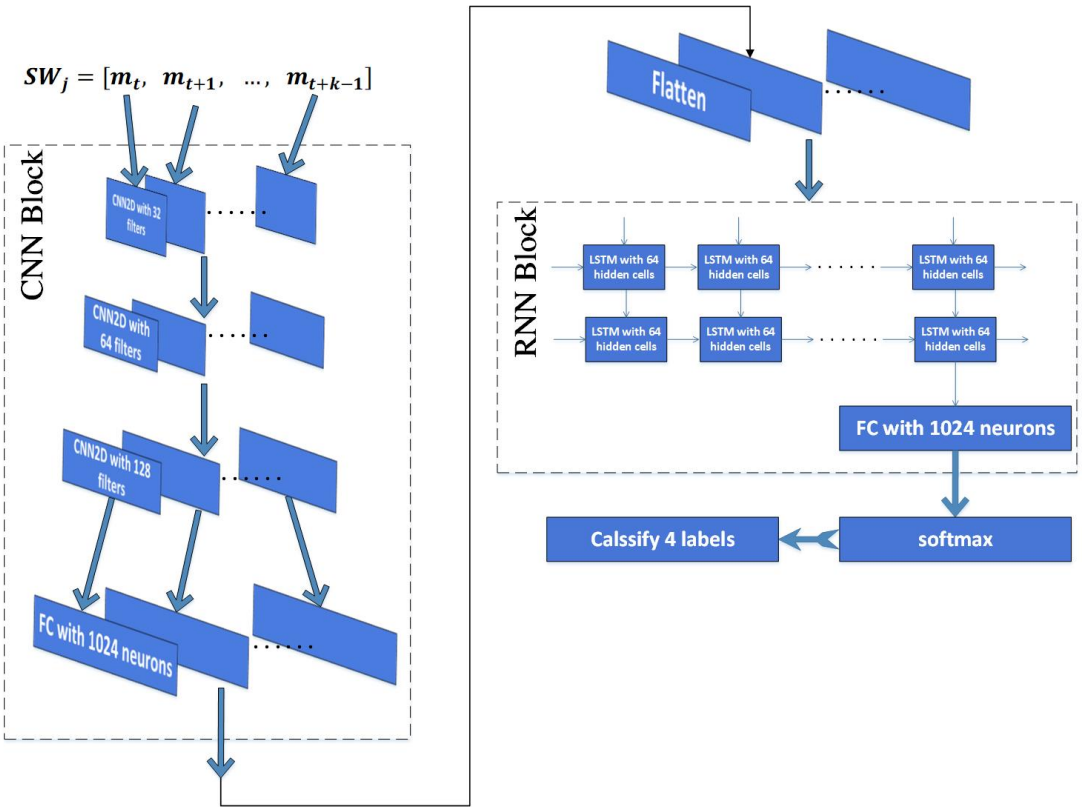

Fig. 5: The proposed RCNN structure.

the model. The model and parameters used in this paper are based on the model presented in [37]. Table 2 summarizes the internal details of proposed RCNN structures in all layers.

Table 2 Details of the proposed RCNN model structure used in this paper.

Layer	Layer Name	Kernel Size	Number of Output	Stride
1	Conv2D	3Œ3	32	1
2	Conv2D	3Œ3	64	1
3	Conv2D	3Œ3	128	1
4	Dense	-	1024	-
5	Flatten	-	-	-
6	LSTM	-	64	-
7	LSTM	-	64	-
8	Dense	-	1024	-
9	Dense	-	4	-

5 Experiments and Results

In this paper, the EEG-fNIRS dataset from Technische Universität Berlin [27] has been used to evaluate RCNN network. Several experiments have been conducted to evaluate the performance of the proposed method and compare it with related models.

5.1 Dataset Description

In this section, we will examine the used datasets in this paper. This dataset is publically available and combines two EEG and fNIRS signals.

Twenty-eight right-handed and one healthy left-handed subjects participated in this study (fourteen males and fifteen females, average age (years) 28.5 ś 3.7). None of them reported neurological, psychiatric or other brain-related diseases. All volunteers were informed about the experimental procedure, and written consent was obtained from all participants [27].

EEG and NIRS data were collected in an ordinary bright room. The EEG signals were recorded by a multichannel BrainAmp EEG amplifier with thirty active electrodes (Brain Products GmbH, Gilching, Germany). EEG signal sampling rate was 1000 Hz. Thirty EEG electrodes were placed on a custom-made stretchy fabric cap (EASYCAP GmbH, Herrsching am Ammersee, Germany) and placed according to the international 10-5 system (AFp1, AFp2, AFF1h, AFF2h, AFF5h, AFF6h, F3, F4, F7, F8, FCC3h, FCC4h, FCC5h, FCC6h, T7, T8, Cz, CCP3h, CCP4h, CCP5h, CCP6h, Pz, P3, P4, P7, P8, PPO1h, PPO2h, POO1, POO2 and Fz for ground electrode). The fNIRS signals were recorded by NIRScout (NIRx GmbH, Berlin, Germany). The fNIRS signal sampling rate was 12.5 Hz. Fourteen sources and sixteen detectors create thirty-six physiological channels. These channels were placed at frontal (nine channels around Fp1, Fp2, and Fpz), motor (twelve channels around C3 and C4, respectively) and visual areas (three channels around Oz). The inter-optode distance was 30 mm. NIRS optodes were fixed on the same cap as the EEG electrodes. All signals were recorded simultaneously (Figure 4).

There are two kinds of datasets, motor imagery, and mental arithmetic. For motor imagery, subjects were instructed to perform haptic motor imagery (i.e. to imagine the feeling of opening and closing their hands as they were grabbing a ball) to ensure that actual motor imagery, not visual imagery, was performed. For mental arithmetic, there are two tasks, MA task, and baseline task. The MA task, an initial subtraction such as 'three-digit number minus one-digit number'(e.g., 384-8) appeared at the center of the screen for 2 s. For the baseline task, no specific sign but the black fixation cross was displayed on the screen, and the subjects were instructed to take a rest. For more details, see [27]. The raw data were downsampled to 200 Hz (EEG) and 10 Hz (fNIRS) rates. No pre-processing was applied to the raw data.

5.2 Summary of experiments

To evaluate the effectiveness of combined EEG and fNIRS, we investigated the accuracy of our model with EEG-only, fNIRS-only, and EEG-fNIRS. In our experiments, the length of the sliding window was 10 samples. We randomly considered, 80% of the data as training data, 10% as evaluation data for tuning parameters, and 10% as test data. Our evaluation setting is inter-subject. All data has been shuffled and the training set and test set separated. Our model is person-independent and all data are used to train and test the proposed model. In this paper, we also investigated the impact of spatio-temporal information on the accuracy of classification. All models are implemented with Keras framework in Python programming language with the TensorFlow backend. The Adam optimizer function with binary_crossentropy loss function is used. Network parameters are taught with a learning rate of 10e-4.

5.3 Results of preprocessing method

As mentioned in Section 4.1, we use the spatial and temporal features of the two signals EEG and fNIRS. In this section, the results of the preprocessing have been presented.

Figures 3 and 6 show the time-series signal, and the result of the proposed preprocessing method, respectively. Figure 6 includes spatial information of the combined EEG and fNIRS signals.Figure 6, presents the topographic maps generated by the proposed model. On the other hand, Figure 6 demonstrates three-rank tensor (m_t). This figure shows brain activity during different tasks and is similar to scalp topography. In our method, when raw EEG and fNIRS data at time t (e_t and f_t) turns into a three-rank tensor (m_t), this tensor is corresponded to brain activity at time t. According to Equation 3, each m_t has 17 rows and 17 columns, so the vertical and horizontal axis are shown these values. By comparing Figures 3 and 6, we can conclude that the converting time-series signal to three-rank tensors provides more detailed information. In the proposed method, extracted spatial features include information about the active regions of the brain at the desired time, as shown in Figure 6. After applying the sliding window, the temporal information is added to spatial information (Figure 7). These maps display changes in the brains active regions during the subject 7 movements. Figure 7 presents input of deep network.

In Figure 6, brain topography is shown for four tasks. The changes in the activity levels in the various brain regions can be seen in these four tasks. In this figure, each color trace represents the spectrum of the activity in one data channel. In the MI right hand, the activity on the left side of the brain is more. This activity can be seen on left side channels. These channels include C1, C3 and so on. But in MI right hand, the activity is deferent on various frequencies. This activity is seen throughout the brain. Also, little activity is seen on the right side of the brain [65]. In MA, the highest level of activity is seen in the frontal lobe and in the Baseline. little activity is seen throughout the brain because the brain is resting during the baseline. The clip includes several three-rank tensors is shown in figure Figure 7 which is the input of proposed RCNN model. Each clip includes spatial and temporal information. In fact, The impact of time during movement by the user is shown in Figure 7. During movement, the brain records activities that can consist of noise. These activities can be similar to another activity for another movement. During record activities, user can move head or tongue and makes noise. So, In each moment, there are similar activities. When time is not considered, the network could not learn activities very well. Because there are the same activities with different labels. Therefore, considering a few samples in tandem will help the network understand the activity in the brain well, and noisy areas of activity will not be considered.

The first step, the CNN model extracts spatial features from this information. Then the RNN model extracts temporal features from the sequence of spatial features in each clip. The proposed RCNN model can read and understand these sequence of figures and this model can extract appropriate features from them.

5.4 Results of proposed method

As mentioned in Section 4.2, we presented a new model based on deep learning that includes the combination of CNN and RNN algorithms. In this section, first, we compare the accuracy of the proposed model with other models in different paper. Second, we investigate the accuracy of different implementations. Finally, we analyze the diagram of accuracy based on number of epochs.

To investigate the importance of simultaneous temporal and spatial analysis, we used two models: CNN, to extract spatial features, and RNN, to extract temporal features, separately. Then, we compared them with our proposed model. Performance of these models compared to other works and the results are shown in Table 3.

Table 3 Comparison models.

Reference	EEG or fNIRS or Hybrid	Classifier	Accuracy
[37]	EEG	Recurrent Convolutional Neural Network	98.3%
[47]	EEG	Deep Residual Learning	92.682%
[36]	EEG	Recurrent Convolutional Neural Network with DWT	86.7%
[9]	EEG or fNIRS	Linear Discriminant Analysis	fNIRS=75.6%, EEG=86%
[50]	EEG-fNIRS	Deep Neural Network	83.28%
[22]	EEG-fNIRS	Deep Belief Network	84.26%
[52]	EEG-fNIRS	Shrinkage Linear Discriminant Analysis	90%
Proposed Method	EEG-fNIRS	Convolutional Neural Network	98.22%
Proposed Method	EEG-fNIRS	Recurrent Neural Network	95.81%
Proposed Method	EEG-fNIRS	Recurrent Convolutional Neural Network	99.63%

It is observed from this table that the proposed model achieves high accuracy of 99.63%. In our study, choosing a larger sliding window also drops the model performance dramatically. This method has the flexibility to adopt different kinds of EEG-fNIRS based intention recognition by varying sliding window size. Compared to previous studies, our model requires less preprocess on raw data making it more suitable for real-time applications, such as BCI.

In [9] and [52], they have not been used deep learning algorithm. The deep learning algorithm is new technology and more accurately than other classifiers. In [47], [36], [50] and [22], a deep learning algorithm for classification for a hybrid BCI and EEG-based BCI system has been used. They changed deep learning model architecture and achieved higher accuracy compared to other methods. Because the deep learning algorithm automatically extracts features. Feature extraction by a human may cause a lot of errors and reduce the data size. The preprocessing step is an important step in BCI systems. In all of the research presented in Table 3, the preprocessing has been used alone to remove noise. In [37], the preprocessing step has not been used to remove noise and they used a preprocessing step to change EEG signals. They changed signals to use special hidden information. Also, they used RCNN model to classify data. The highest achieved accuracy in Table 3 is 98.63%. In the proposed method, by adding the fNIRS signal to EEG and use temporal and spatial information simultaneously and use RCNN model as classifier the accuracy increases for 1% and the accuracy is 99.63%.

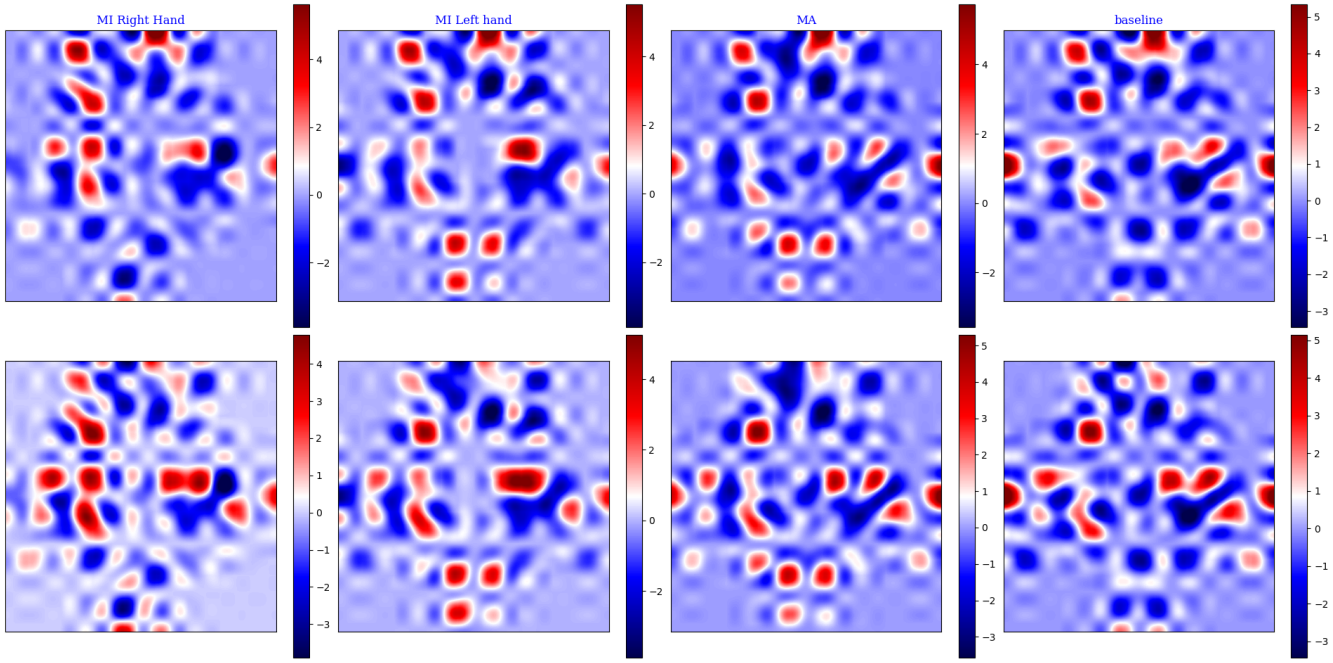

Fig. 6: Result of proposed preprocessing method. Brain activity during different task like brain topography. First row is showing the combining EEG and low-length fNIRS (facet one in three-rank tensor) and the second row is showing the combining EEG and high-length fNIRS (facet two). First column to forth column are showing one of the tasks. These columns are showing brain activity for MI right hand, MI left hand, MA, and Baseline, respectively. In these figures, the colors are showing low or high activity in specific area in brain. Dark red color indicates highest activity and dark blue color indicates lowest activity in brain.

On the other hand, as already mentioned, fNIRS data has two wavelengths. We investigated the effect of both wavelengths along with EEG signal. The performance of the proposed model for EEG and each wavelength of fNIRS is shown in Figure 8.

According to Figure 8, we observed that achieved accuracy for the combination of EEG with low wavelength, and high wavelength in fNIRS signals are almost equal, and both wavelengths provide equal information. As a result, we can use any of these two. However using fNIRS alone for classification greatly reduces the accuracy. The advantage of such selection is that the computational burden is reduced. According to Figure 8, the accuracy of the combined EEG-fNIRS is higher than those of the two. The accuracy of 98% for EEG signal, accuracy of 86.03% for fNIRS signals and 99.63% for EEG-fNIRS indicates that adding fNIRS data to EEG data increases the accuracy of the prediction of intentions. As already mentioned, recording fNIRS data simultaneously does not cause any problem in recording EEG data. Therefore, the simultaneous use of these two signals in our proposed method can be a great help in signal analysis and prediction of intentions and use in systems requiring high accuracy, such as BCI. One of the most important parameters in deep neural networks is the number of epochs. The epoch is a complete implementation of model training. In most cases, choosing a higher number of epochs leads the network to be trained well. But in higher epochs, changes in accuracy may be negligible. To find an optimal number of epochs, we must attend to the changes in accuracy. The accuracy diagram based on the number of epochs clearly shows these changes.

Figure 9 shows the accuracy diagram based on the number of epochs. As can be seen, with increasing number of epochs, accuracy has increased. However, with increasing number of epochs from 60 to 80, the changes in accuracy were insignificant and the overall system accuracy was not increasing. Hence, in order to reduce the computational burden of the system. We consider 60 as the number of epochs in all experiments.

One of the most important hyper-parameters is the size of the sliding window. This number has a huge impact on network training and its accuracy. The impact of different size of the sliding window on the accuracy of the network is presented in Table 4. It is observed from this table that the accuracy of network decreases by increasing this number. The reason is that the large window has more information. It is possible when we consider the large size for sliding window, the information consist of noise and the network cannot be trained well. On the other hand, the large size of the sliding window decreases the number of training data. Therefore, in this paper, we chose 10 for this size.

Table 4 The accuracy of network due to the different size of sliding window.

Size of sliding window (sample)	Accuracy
10	99.63%
64	94.95%
100	90.91%
200	86.87%

5.5 Discussion

As shown in Table 3, the use of both spatial and temporal information and extracting the relationship between adjacent signals can help to detect human intention. The accuracy of the classification is 95.81% using the temporal information alone. Therefore, the temporal chain-like signals do not have much effect on the accuracy of the classification. In the most recent methods, these signals are investigated as time series. In this paper, we have not used any feature extraction algorithm, but we have used feature engineering in deep learning. On the other hand, considering the spatial information, the accuracy of the classification up to 98.22% has been upgraded. But, as shown in Table 3, the proposed model, RCNN, is achieved the accuracy of 99.63%. Therefore, the combination of temporal and spatial information greatly impacts the better analysis of signals. The conversion of EEG and fNIRS signals to three-rank tensors, and the use of the location of active brain regions increase the accuracy significantly. In proposed method, unlike other methods, preprocessing has not been used to remove noise in signals. Using raw signals are good but the raw biological signals always have noise. Therefore, it

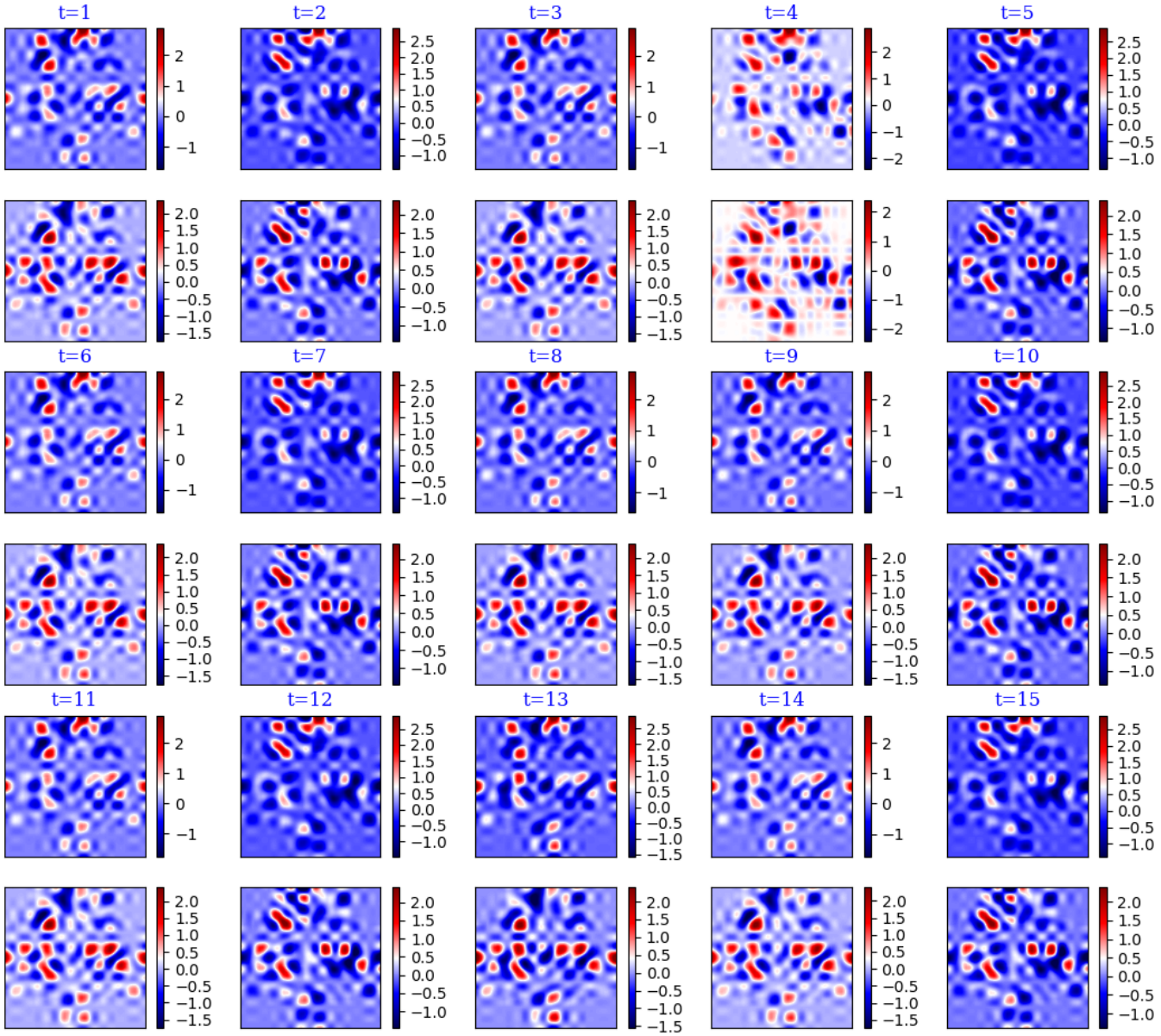

Fig. 7: After use sliding window. Spatial and temporal feature is used to train proposed model. For each time, first row is showing the combining EEG and low-length fNIRS (facet one in three-rank tensor) and the second row is showing the combining EEG and high-length fNIRS (facet two). All of these figures are showing brain activity for MI right hand at time t=1 to 15. In these figures, the colors are shown low or high activity in specific area in brain. Dark red color indicates highest activity and dark blue color indicates lowest activity in brain.

is better to use preprocessing algorithms to remove noises. This may improve detection.

Using more channels in particular areas of the head can also help to increase the model classes and thus more mental activity detection. Hence, in future work, we can increase the number of translated commands into the system by acquisition signal from more channels.

6 Conclusion and Future Work

In this paper, we classified biological signals using spatial and temporal information contained in them and using deep learning. In the proposed method, the combination of two data EEG and fNIRS has been used. We continued to extract two spatial and temporal characteristics of the signal. In the proposed method, signals have been classified by converting chain-like signals into three-rank tensors. Then the deep learning of RCNN has been used to classify this data. Therefore, we observed that the accuracy of the method for combining these two signals was significantly higher than other methods.

Adding preprocessing step to remove noise from the signal, as described in Section 4.1, can be done in the future. This may improve the classification. Finding the optimal parameters for a neural network is always a grand challenge. To result this issue, a consistent algorithm to find and set the optimal parameters for the deep network can be presented in the future. Another interesting future path of research could be acquisition and combining different datasets with more number of classes for the real-world BCI systems.

7 References

1 Hofman, Michel A and Falk, Dean: Evolution of the primate brain: from neuron to behavior. Elsevier. 195, (2012).
2 Illes, J. and Sahakian, B.J. eds.: Oxford handbook of neuroethics. Oxford University Press. (2013).
3 Guo, F., Hong, B., Gao, X. and Gao, S.: A braincomputer interface using motion-onset visual evoked potential. Journal of neural engineering. 5(4), p.477 (2008).
4 Ware, C. and Mikaelian, H.H.: An evaluation of an eye tracker as a device for computer input2. Acm sigchi bulletin. 18(4), 183-188 (1987).
5 Bauer, G., Gerstenbrand, F. and Rumpl, E.: Varieties of the locked-in syndrome. Journal of neurology, 221(2), 77-91 (1979).

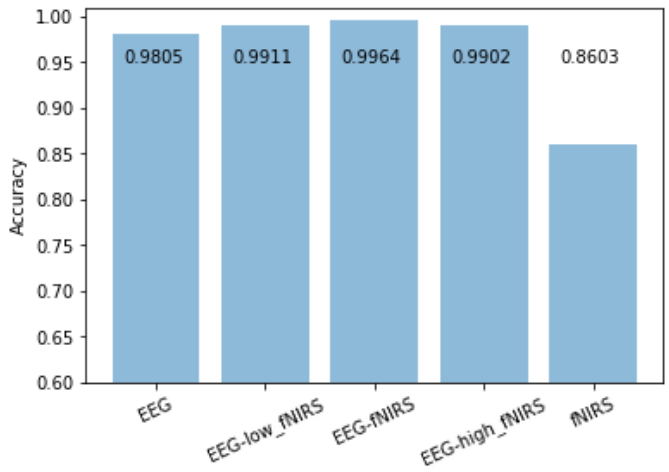

Fig. 8: Accuracy of various implementations.

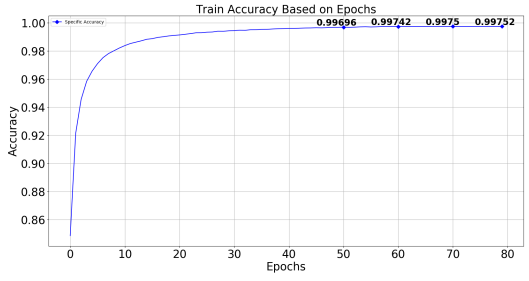

Fig. 9: Accuracy diagram based on epochs.

6 Patterson, J.R. and Grabois, M.: Locked-in syndrome: a review of 139 cases. Stroke, 17(4), 758-764 (1986).

7 Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J.R. and Birbaumer, N.: Braincomputer communication: Unlocking the locked in. Psychological bulletin, 127(3), 358 (2001).

8 Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J. and Vaughan, T.M.: Brain-computer interface technology: a review of the first international meeting. IEEE transactions on rehabilitation engineering, 8(2), 164-173 (2000).

9 Khan, M.J. and Hong, K.S.: Hybrid EEGfNIRS-based eight-command decoding for BCI: application to quadcopter control. Frontiers in neurorobotics, 11, 6 (2017).

10 Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. and Arnaldi, B.: A review of classification algorithms for EEG-based braincomputer interfaces. Journal of neural engineering, 4(2), R1 (2007).

11 Tabar, Y.R. and Halici, U.: Brain Computer Interfaces for Silent Speech. European Review, 25(2), 208-230 (2017).

12 Wolpaw, J.R., Loeb, G.E., Allison, B.Z., Donchin, E., do Nascimento, O.F., Heetderks, W.J., Nijboer, F., Shain, W.G. and Turner, J.N.: BCI meeting 2005-workshop on signals and recording methods. IEEE Transactions on neural systems and rehabilitation engineering, 14(2), 138-141 (2006).

13 Gelenbe, E., Feng, Y. and Krishnan, K.R.R.: Neural network methods for volumetric magnetic resonance imaging of the human brain. Proceedings of the IEEE, 84(10), 1488-1496 (1996).

14 Bauernfeind, G., Leeb, R., Wriessnegger, S.C. and Pfurtscheller, G.: Development, set-up and first results for a one-channel near-infrared spectroscopy system/Entwicklung, Aufbau und vorläufige Ergebnisse eines Einkanal-Nahinfrarot-Spektroskopie-Systems. Biomedizinische Technik, 53(1), 36-43 (2008).

15 Wang, H., Zhang, Y., Waytowich, N.R., Krusienski, D.J., Zhou, G., Jin, J., Wang, X. and Cichocki, A.: Discriminative feature extraction via multivariate linear regression for SSVEP-based BCI. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(5), 532-541 (2016).

16 Gruzelier, J.H.: EEG-neurofeedback for optimising performance. III: a review of methodological and theoretical considerations. Neuroscience & Biobehavioral Reviews, 44, 159-182 (2014).

17 Ahn, M. and Jun, S.C.: Performance variation in motor imagery braincomputer interface: a brief review. Journal of neuroscience methods, 243, 103-110 (2015).

18 Ramli, R., Arof, H., Ibrahim, F., Mokhtar, N. and Idris, M.Y.I.: Using finite state machine and a hybrid of EEG signal and EOG artifacts for an asynchronous wheelchair navigation. Expert Systems with Applications, 42(5), 2451-2463 (2015).

19 Zhang, R., Li, Y., Yan, Y., Zhang, H., Wu, S., Yu, T. and Gu, Z.: Control of a wheelchair in an indoor environment based on a braincomputer interface and automated navigation. IEEE transactions on neural systems and rehabilitation engineering, 24(1), 128-139 (2015).

20 Peng, H., Chao, J., Wang, S., Dang, J., Jiang, F., Hu, B. and Majoe, D.: Single-trial classification of fNIRS signals in four directions motor imagery tasks measured from prefrontal cortex. IEEE transactions on nanobioscience, 17(3), 181-190 (2018).

21 Janani, A. and Sasikala, M.: Classification of fNIRS Signals for Decoding Right-and Left-Arm Movement Execution Using SVM for BCI Applications. In Computational Signal Processing and Analysis, Springer, Singapore. 315-323 (2018).

22 Ho, T.K.K., Gwak, J., Park, C.M., Khare, A. and Song, J.I.: Deep Leaning-Based Approach for Mental Workload Discrimination from Multi-channel fNIRS. In Recent Trends in Communication, Computing, and Electronics, Springer, Singapore. 431-440 (2019).

23 Fazli, S., Mehnert, J., Steinbrink, J., Curio, G., Villringer, A., Müller, K.R. and Blankertz, B.: Enhanced performance by a hybrid NIRSEEG brain computer interface. Neuroimage, 59(1), 519-529 (2012).

24 Hong, K.S., Khan, M.J. and Hong, M.J.: Feature extraction and classification methods for hybrid fNIRS-EEG brain-computer interfaces. Frontiers in human neuroscience, 12 (2018).

25 Herrmann, M.J., Huter, T., Plichta, M.M., Ehlis, A.C., Alpers, G.W., Mühlberger, A. and Fallgatter, A.J.: Enhancement of activity of the primary visual cortex during processing of emotional stimuli as measured with eventrelated functional near-infrared spectroscopy and eventrelated potentials. Human Brain Mapping, 29(1), 28-35 (2008).

26 Wallois, F., Patil, A., Héberlé, C. and Grebe, R.: EEG-NIRS in epilepsy in children and neonates. Neurophysiologie Clinique/Clinical Neurophysiology, 40(5-6), 281-292 (2010).

27 Shin, J., von Lühmann, A., Blankertz, B., Kim, D.W., Jeong, J., Hwang, H.J. and Müller, K.R.: Open access dataset for EEG+ NIRS single-trial classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(10), 1735-1745 (2016).

28 Shin, J., Müller, K.R., Schmitz, C.H., Kim, D.W. and Hwang, H.J.: Evaluation of a compact hybrid brain-computer interface system. BioMed research international. (2017).

29 Biessmann, F., Plis, S., Meinecke, F.C., Eichele, T. and Muller, K.R.: Analysis of multimodal neuroimaging data. IEEE Reviews in Biomedical Engineering, 4, 26-58 (2011).

30 Friston, K.J.: Modalities, modes, and models in functional neuroimaging. Science, 326(5951), 399-403 (2009).

31 Shibasaki, H.: Human brain mapping: hemodynamic response and electrophysiology. Clinical Neurophysiology, 119(4), 731-743 (2008).

32 Elsayed, N., Zaghloul, Z.S. and Bayoumi, M.: Brain computer interface: EEG signal preprocessing issues and solutions. International Journal of Computer Applications, 169(3), 975-8887 (2017).

33 Sun, S. and Zhou, J.: A review of adaptive feature extraction and classification methods for EEG-based brain-computer interfaces. In 2014 International Joint Conference on Neural Networks (IJCNN), IEEE. 1746-1753 (2014).

34 Heydari, E. and Shahbakhti, M.: Adaptive wavelet technique for eeg de-noising. In 2015 8th Biomedical Engineering International Conference (BMEiCON), IEEE. 1-4(2015).

35 Yin, Z., Wang, Y., Liu, L., Zhang, W. and Zhang, J.: Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination. Frontiers in neurorobotics, 11, 19 (2017).

36 Yang, J., Yao, S. and Wang, J.: Deep Fusion Feature Learning Network for MI-EEG Classification. IEEE Access, 6, 79050-79059 (2018).

37 Zhang, D., Yao, L., Zhang, X., Wang, S., Chen, W. and Boots, R.: EEG-based intention recognition from spatio-temporal representations via cascade and parallel convolutional recurrent neural networks. arXiv preprint arXiv:1708.06578 (2017).

38 Vaadia, E. and Birbaumer, N.: Grand challenges of brain computer interfaces in the years to come. Frontiers in neuroscience, 3, 15 (2009).

39 Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L.E., Sturm, I., Curio, G. and Mueller, K.R.: The Berlin brain-computer interface: non-medical uses of BCI technology. Frontiers in neuroscience, 4, 198 (2010).

40 Van Erp, J., Lotte, F. and Tangermann, M.: Brain-computer interfaces: beyond medical applications. Computer, 45(4), 26-34 (2012).

41 Pfurtscheller, G., Allison, B.Z., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T.O., Mueller-Putz, G., Neuper, C. and Birbaumer, N.: The hybrid BCI. Frontiers in neuroscience, 4, 3 (2010).

42 Nguyen, D.K., Tremblay, J., Pouliot, P., Vannasing, P., Florea, O., Carmant, L., Lepore, F., Sawan, M., Lesage, F. and Lassonde, M.: Non-invasive continuous EEG-fNIRS recording of temporal lobe seizures. Epilepsy research, 99(1-2), 112-126 (2012).

43 Peng, K., Nguyen, D.K., Tayah, T., Vannasing, P., Tremblay, J., Sawan, M., Lassonde, M., Lesage, F. and Pouliot, P.: fNIRS-EEG study of focal interictal epileptiform discharges. Epilepsy research, 108(3), 491-505 (2014).

44 Seyal, M.: Frontal hemodynamic changes precede EEG onset of temporal lobe seizures. Clinical Neurophysiology, 125(3), 442-448 (2014).

45 Näsi, T., Kotilahti, K., Noponen, T., Nissilä, I., Lipiäinen, L. and Meriläinen, P.: Correlation of visual-evoked hemodynamic responses and potentials in human brain. Experimental brain research, 202(3), 561-570 (2010).

46 Kosmyna, N. and Lécuyer, A.: A conceptual space for EEG-based brain-computer interfaces. PloS one, 14(1), e0210145 (2019).

47 Zeng, H., Yang, C., Dai, G., Qin, F., Zhang, J. and Kong, W.: EEG classification of driver mental states by deep learning. Cognitive neurodynamics, 12(6), 597-606 (2018).

48 Khan, M.J., Ghafoor, U. and Hong, K.S.: Early detection of hemodynamic responses using EEG: a hybrid EEG-fNIRS study. Frontiers in Human Neuroscience, 12, 479 (2018).

49 Zhang, M., Hua, Q., Jia, W., Chen, R., Su, H. and Wang, B.: Feature extraction and classification algorithm of brain-computer interface based on human brain central nervous system. NeuroQuantology, 16(5) (2018).

50 Chiarelli, A.M., Croce, P., Merla, A. and Zappasodi, F.: Deep learning for hybrid EEG-fNIRS braincomputer interface: application to motor imagery classification. Journal of neural engineering, 15(3), 036028 (2018).
51 Zafar, A., Khan, M.J., Park, J. and Hong, K.S.: Initial-dip Based Quadcopter Control: Application to fNIRS-BCI. IFAC-PapersOnLine, 51(15), 945-950 (2018).
52 Shin, J., Kim, D.W., Müller, K.R. and Hwang, H.J.: Improvement of information transfer rates using a hybrid EEG-NIRS brain-computer interface with a short trial length: Offline and pseudo-online analyses. Sensors, 18(6), 1827 (2018).
53 Yohanandan, S.A., Kiral-Kornek, I., Tang, J., Mshford, B.S., Asif, U. and Harrer, S.: A Robust Low-Cost EEG Motor Imagery-Based Brain-Computer Interface. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. 5089-5092 (2018).
54 Shin, J., Kwon, J. and Im, C.H.: A ternary hybrid EEG-NIRS brain-computer interface for the classification of brain activation patterns during mental arithmetic, motor imagery, and idle state. Frontiers in neuroinformatics, 12, 5 (2018).
55 Liu, T., Liu, X., Yi, L., Zhu, C., Markey, P.S. and Pelowski, M.: Assessing autism at its social and developmental roots: A review of Autism Spectrum Disorder studies using functional near-infrared spectroscopy. Neuroimage, 185, 955-967 (2019).
56 Yoo, S.H., Woo, S.W. and Amad, Z.: Classification of three categories from prefrontal cortex using LSTM networks: fNIRS study. In 2018 18th International Conference on Control, Automation and Systems (ICCAS), IEEE. 1141-1146 (2018).
57 Mustafa, I. and Mustafa, I.: Smart thoughts: BCI based system implementation to detect motor imagery movements. In 2018 15th International Bhurban Conference on Applied Sciences and Technology (IBCAST), IEEE. 365-371 (2018).
58 Pandey, P. and Seeja, K.R.: Subject-Independent Emotion Detection from EEG Signals Using Deep Neural Network. In International Conference on Innovative Computing and Communications, Springer, Singapore. 41-46 (2019).
59 Lin, X., Sai, L. and Yuan, Z.: Detecting concealed information with fused electroencephalography and functional near-infrared spectroscopy. Neuroscience, 386, 284-294 (2018).
60 Oostenveld, R. and Praamstra, P.: The five percent electrode system for high-resolution EEG and ERP measurements. Clinical neurophysiology, 112(4), 713-719 (2001).
61 LeCun, Y., Kavukcuoglu, K. and Farabet, C.: Convolutional networks and applications in vision. In Proceedings of 2010 IEEE International Symposium on Circuits and Systems, IEEE. 253-256 (2010).
62 Gers, F.A., Schmidhuber, J. and Cummins, F.: Learning to forget: Continual prediction with LSTM (1999).
63 Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. and Darrell, T.: Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2625-2634 (2015).
64 Kingma, D.P. and Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).
65 Rashed-Al-Mahfuz, M., Islam, M.R., Hirose, K. and Molla, M.K.I.: Artifact suppression and analysis of brain activities with electroencephalography signals. Neural regeneration research, 8(16), 1500 (2013).
66 Chambon, S., Galtier, M.N., Arnal, P.J., Wainrib, G. and Gramfort, A.: A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), 758-769 (2018)
67 Li, J., Zhang, Z. and He, H.: Implementation of eeg emotion recognition system based on hierarchical convolutional neural networks. In International Conference on Brain Inspired Cognitive Systems, Springer, Cham, 22-33 (2016)
68 She, Q., Hu, B., Luo, Z., Nguyen, T. and Zhang, Y.: A hierarchical semi-supervised extreme learning machine method for EEG recognition. Medical & biological engineering & computing, 57(1), 147-157 (2019)
69 Li, R., Potter, T., Huang, W. and Zhang, Y.: Enhancing performance of a hybrid EEG-fNIRS system using channel selection and early temporal features. Frontiers in human neuroscience, 11, 462 (2017)
70 Lu, N., Li, T., Ren, X. and Miao, H.: A deep learning scheme for motor imagery classification based on restricted boltzmann machines. IEEE transactions on neural systems and rehabilitation engineering, 25(6), 566-576 (2016)