Neural Computing and Applications
https://doi.org/10.1007/s00521-020-04804-y

S.I. : APPLYING ARTIFICIAL INTELLIGENCE TO THE INTERNET OF THINGS

A hybrid classifier combination for home automation using EEG signals

Partha Pratim Roy[1] · Pradeep Kumar[1] · Victor Chang[2]

Received: 13 November 2019 / Accepted: 19 February 2020

Abstract

Over the years, the usage of artificial intelligence (AI) algorithms is increased to develop various smart applications using Internet-of-Things. Home automation is a fast emerging area that involves monitoring and controlling of household appliances for user comfort and efficient management. Using mental commands to control different electrical appliances and objects in house is a very interesting application. Brain–Computer Interface is used to relay the information from the subject's brain to an Electronic device, and such devices can be used for this purpose. The information from the subject's brain is collected in form of Electroencephalogram (EEG) signals. In this paper, we analyze the use of EEG signals for applications related to home automation. We present a hybrid model which makes use of Long Short-Term Memory which is considered as a robust temporal classification model in AI and classical Random Forest Classifier for EEG classification. We also discuss how our proposed hybrid model overcomes the limitation presented by the individual models. To arrive at the best model, we have analyzed various parameters such as sampling rate and combination of different brain rhythms which we finally use in our hybrid model. Based on experiments conducted on a custom-built dataset, we also discuss the spatial significance of different electrodes of the EEG device and get insight in signals generated from different areas of the brain.

Keywords EEG · IoT · LSTMs · Hybrid machine learning

1 Introduction

The increasing popularity of artificial intelligence (AI) algorithms leads to the development of various smart applications using Internet-of-Things (IoT) devices. These IoT devices are capable of transferring information over the internet by integrating data from various devices to provide different solutions. Home automation is remote monitoring and controlling of electrical appliances. Currently, remote monitoring is done manually using remote controller, voice commands, gestures, etc. What if we could do remote controlling simply through mental commands, i.e., merely through our thoughts? This will help people do mundane tasks easily in their hectic life. In this paper, we present a prototype to test the feasibility of such a model. This prototype has only one appliance, bulb, which we are trying to switch on/off by mental commands.

The rise of modest off-the-shelf wireless EEG gadgets drove analysts to investigate novel standards in the field of Human–Computer Interaction. Indeed, the consistence of these gadgets with the IoT standards toward inescapable EEG motioning in savvy home situations enables new models of connection and an alternate point of view from traditional affective figuring. Distinctive cerebrum states are the consequence of various examples of neural connection. These examples prompt waves described by various amplitudes and frequencies. This neural collaboration is finished with various neurons. Each collaboration between neurons makes an infinitesimal electrical release. Diverse mind states are the after effect of various examples of neural collaboration. These examples prompt waves portrayed by various amplitudes and frequencies.

✉ Partha Pratim Roy
proy.fcs@iitr.ac.in

Pradeep Kumar
pra14.dcs2014@iitr.ac.in

Victor Chang
v.chang@tees.ac.uk

1 Department of Computer Science and Engineering, IIT Roorkee, Roorkee 247667, India

2 School of Computing, Engineering and Digital Technologies, Teesside University, Middlesbrough, UK

Brain–Computer Interface (BCI) is a way to achieve the aforementioned objective. It can provide a better alternative to the traditional remote controllers, in which we have to manually give the command. This would also be advantageous over other voice recognition systems and gesture-controlled devices as it directly translates the thought of the person into the desired action [1–3]. This makes it robust to the accent and other nuances of voice as well as to noise in loud settings. The current remote controlling devices are inconvenient for people with severe motor disabilities like advanced Amyotrophic Lateral Sclerosis (ALS), brainstem stroke, and serious cerebral palsy. Thus, BCI also helps patients suffering with such diseases. BCI will also increase safety and economize energy consumption. It is an extension of existing home automation where instead of using gadgets we translate the intentions of the user into actions automatically [4, 5]. Apart from these, in medical applications where the surgeons or nurses may not be have their hands free but still need to control the system, our application can come to the rescue.

BCI uses brain signals recorded in form of Electroencephalogram (EEG), Magnetoencephalography (MEG), Functional magnetic resonance imaging (fMRI), etc. fMRI measures brain activity by detecting changes associated with blood flow in the brain which is noninvasive and safe. fMRI has been used successfully by Berns and Moore [6] to predict music popularity. It describes brain function with spatial and temporal resolutions of one millimeter and per second change, respectively [7]. However, fMRI devices are highly expensive amounting up to several millions of dollars [8] which limits the use of fMRI devices for applications such as ours. Moreover, EEG's can detect changes within a millisecond time-frame as compared to fMRI which has a greater time resolution. EEG measures the brain's electrical activity directly, while fMRI records change in blood flow which are indirect measure of brain electrical activity. Similarly, MEG is more expensive than EEG although both measure same type of brain activity. MEG is considered superior for localization. Since localization is not the main focus of our study, EEG should be sufficient for our purpose. EEG being mobile and relatively inexpensive enables to extend the use of the device from laboratory to general practice. This is the major motivation to use EEG signals for the proposed work. Researchers have worked on prototypes of EEG-based BCI [9]. Motor imagery and action observation have been proposed in the literature by Nueper et al. [10]. EEG in combination with Electromyogram (EMG) signals have been used for rehabilitation of patients with difficulty in motor functions [11]. BCI have been successfully used in virtual reality games [12].

Currently, latest generation embedded systems are used for smart applications. In future, we can build embedded systems which can make use of EEG signals for smart applications. Embedded systems have been successful in data acquisition [13]. BCI are an excellent example where embedded systems can be successfully used [14]. Embedded devices can be used for monitoring diseases like epilepsy [15]. Latest pattern recognition techniques are of utmost importance in embedded devices [16].

Researchers have explored sparse Bayesian approaches for EEG signal classification [17]. Sparse models have been used for the task of mental imagery classification [18] successfully. Sparse models are complex, and hence it has certain limitations. EEG signals have been used in a variety of applications [19–21]. One of the important applications is neuromarketing [22]. EEG data was acquired from human subjects while showing different products, and their brain activity were recorded. The EEG data was used to provide insight into consumer behavior. EEG signals can be used to get insight into understanding human behavior, and one such important study is analyzing behavior during listening to music [23]. Understanding EEG signals is still an unexplored territory, and still it is not known in which applications they can be applied and still researchers have not been able to use the EEG signals up to their full potential. The main contributions of this paper are listed below:

1. First, we present a hybrid model that captures the structural and temporal information contained in the EEG signals. For achieving this, we make use of advanced AI techniques for IoT, i.e., Long Short-Term Memory (LSTM) and classical Random Forest Classifier (RFC). We have also tuned different parameters such as sampling rate and combination of brain rhythms.
2. The next main contribution is the analysis of spatial significance of the electrodes in classifying EEG signals for our task.

The rest of the paper is organized as follows. Section 2 provides an overview of the existing research. Next, we discuss pre-processing, classification methodology and spatial characteristics of channels in Sect. 3. Results obtained using a custom-built dataset in our laboratory are presented in Sect. 4, and finally the conclusion of the paper with possible future extensions is discussed in Sect. 5.

2 Related work

There has been some work proposed in the area of home automation using vocal commands [24]. Researchers have been fascinated about controlling devices by mental commands for a long time. Soon after Hans Berger and Moore [6] in 1929 recorded the first human EEG, researchers have proposed its use for communication and control which will eliminate the use of any intermediaries like peripheral nerves and muscles. Our task involves brain commands and mental imagery. There has been a lot of work related to mental imagery and its classification. In [25], Costa and Cabral have investigated the use of Adaptive Gaussian Representation of EEG Signals as feature and trained a Multilayer perceptron (MLP) using the classical back-propagation algorithm for prediction of imaging right- and left-hand movement. Hashimoto and Ushiba [26] have done work on classification of imaginary left- and right-foot movements using EEG signals. EEG signals have been used to understand word familiarity [27].

Zhou et al. [28] added feature based on higher-order statistics of the bi-spectrum of the signals to the feature set. This was done to exploit the information of nonlinearity and non-Gaussian nature in the EEG signals. These approaches talk about various ways in which we could harness the data hidden in the EEG signals. In [29], the authors attempt to classify mental tasks using spatio-spectral statistics. However, these works do not exploit the state-of-the-art machine learning techniques for EEG signal for mental imagery classification.

Most of the work done in the direction of classifying the EEG signals involves a two-phase approach.

1. Extraction of features from EEG Signals which can be used for classification.
2. Classify the signals using the extracted features.

There has been a lot of work about extracting features from EEG signals, or getting a meaningful representation of the EEG signals. Siuly et al. [30] present a two-level approach for classification of EEG signals. In their hybrid model, they first extract the features from EEG signals using clustering technique (CT). They then classify the signal using Least Square SVM (LS-SVM) into two classes [31]. In another paper, the same author has explored sampling in the detection of multi-category signals [32].

The authors in [33] described a two-phase classification of EEG signals into seizure and seizure free class. They make use of empirical mode decomposition (EMD) to decompose the signals into components known as intrinsic mode functions (IMFs). They use area parameter and mean frequency estimation of IMFs as features to LS-SVM. Aslan et al. [34] compared Multilayer Perceptron Neural Network (MLPNN) and Radial Basis Function Neural Network (RBFNN) for classification of EEG signals into epilepsy groups. They found RBFNN surpasses MLPNN for this task. RBFNN is faster than MLPNN and performs rather well for a simplified architecture.

We could only find very little work involving the state-of-the-art deep neural networks for classification of the EEG signals. The work done by Bashivan et al. in [35] inspired us to use deep neural networks for the task of EEG signal classification. The authors transform the raw signals into a more robust representation of EEG signals which preserves the spatial, temporal and spectral information of the signals. After transforming the signals, they pass them through deep convolutional neural network to extract feature in the spatial and spectral domain. Then, the features from the previous layer are passed to an LSTM to extract any information in the time domain. Finally, they are passed through a fully connected layer for classification.

Inspired by the above work in the area of classification of the EEG signals, we also propose a two-phase approach for our task. We develop a hybrid model which uses stacking, where the first layer of stack extracts the feature from the pre-processed EEG signals with the help of Long Short-Term Memory (LSTM). The second layer is Random Forest Classifier (RFC) which is used to classify the signals into three classes. We describe the entire process in detail in the next section.

3 Proposed framework

This section describes the total work-flow from system setup to data analysis. System setup and data collection are described in Sect. 3.1. Then, we describe our pre-processing techniques in Sect. 3.2. In Sect. 3.3, we describe the proposed hybrid training model. Figure 1 shows the overview of the work-flow. Finally, in Sect. 3.4, we describe the second contribution of this paper, i.e., spatial significance of the electrodes.

3.1 System setup and data collection

Emotiv Epoc+ (Fig. 2a) is used to record EEG signals. It is a wireless device with 14 channels aligned, as shown in Fig. 2a. Electrodes placed above the ears (CMS & DRL) act as reference electrodes although the device internally samples at a frequency of 2048, which is down-sampled to 128 Hz at the time of collection. The setup was fully automated using custom-written programs, and the collected data was continuously transmitted to a computer via Bluetooth technology. To improve the signal quality, the felt pads (sensors) are dampened by saline solution.

The experiment was conducted on 18 healthy male participants between ages 20 and 25 where all the volunteers gave informed consent. The subjects were made to sit in a dimly lit quiet room, and they were provided instructions through audio and visual stimuli from a

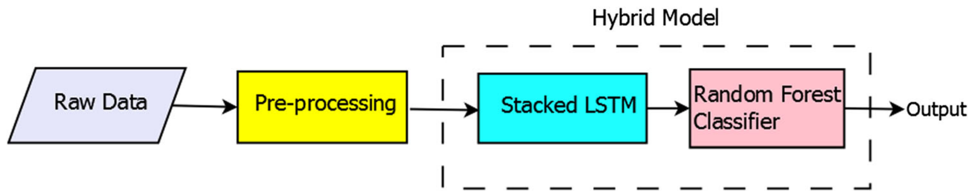


Fig. 1 Overview of the proposed BCI framework

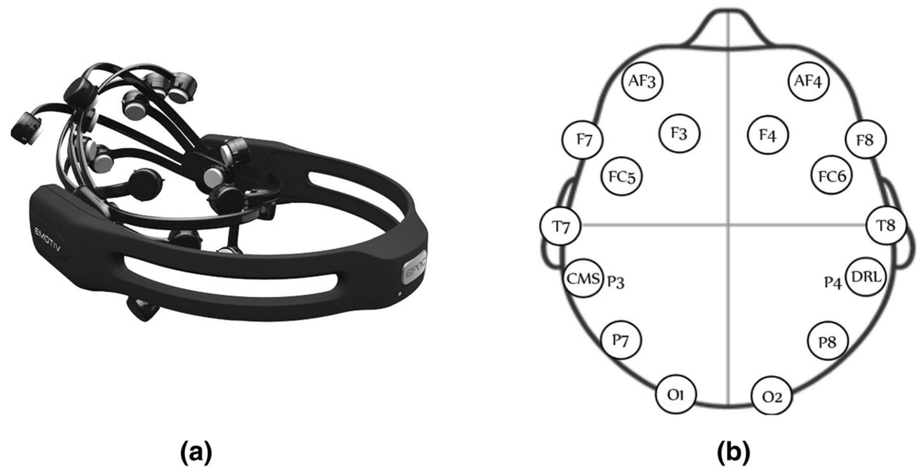


Fig. 2 EEG brain sensor details: **a** Emotiv EPOC+ EEG headset, **b** EEG electrode's position over the scalp

Fig. 3 Trial split-up. 's' stands for time in seconds. Subjects were given stimulus for 14 s for performing one task followed by 15 s of relaxation time

computer. EMOTIV was installed on their scalps, as shown in Fig. 2a, b.

The experiment was conducted into two phases. In the first phase, subjects were shown images of glowing bulb and they were asked to give mental commands for turning the bulb-off. The subjects were supposed to keep their eyes closed and limit other body movements to reduce noise in the collection of data. This was done for 14 s, and after completion of this time an prerecorded audio signal informed the subject that the command time was complete and they can relax. They were allowed to relax for 15 s with the Emotiv on. This was done to give some rest and make the human subject ready for next part. Also, this data is used to represent the state of 'no-command' or equivalently a 'do-nothing' command. In the next phase, the subjects were shown an unlit bulb and they were required to give mental commands to turn it on. So, the dataset has data belonging to three different classes, namely on-to-off, off-to-on, do-nothing (no command). The entire trial is described in Fig. 3, where X and Y correspond to data belonging to unwanted data collected during start and end of each trial. Data corresponding to 'Lights-On' phase is labeled as 'off-to-on' data, 'Lights-Off' phase is labeled as 'on-to-off' data and 'relax' phase is labeled as 'do-nothing' data. Each subject provided 5 trials for the experiment.

The captured signal goes through certain signal pre-processing and digital filtering steps. The first step involves applying Savitzky–Golay (S–Golay) noise filter on the captured signal. Then digital filtering is applied to extract corresponding brain rhythms (alpha, beta, theta, gamma, delta). Next, classification models are built according to the collected ground truth corresponding to various commands. In the testing phase, a recorded test sample is pre-processed and decomposed into frequency bands for feature extraction and tested against the trained model. The results were classified in three classes, namely 'off-to-on,' 'on-to-off' and 'do-nothing,' corresponding to desired actions of turning-on, turning-off and letting it remain as it is.

3.2 Pre-processing

This section presents the details of the applied signal processing methods used to pre-process the signal for feature extraction. Pre-processing comprises the following four stages:

1. Denoising
2. Digital Filter
3. Cleaning
4. Parameter-based pre-processing

First of all, raw data is denoised by applying Savitzky–Golay filter, then the denoised signal is used to extract signals from five different frequency bands, namely alpha, beta, theta, delta and gamma. Data corresponding to these five filters and the original denoised signals are cleaned by discarding irrelevant data. Irrelevant data corresponds to the data being collected during some portion of relax time and some portion of the data during start and end of each trial. Finally, data is subjected to parameter-based processing, parameters being time segment and sampling rate. The work-flow for pre-processing is shown in Fig. 4. Detailed description of each of these steps is given in the following subsections.

3.2.1 Denoising

Savitzky–Golay (S–Golay) filter is used to denoise the signals. S–Golay filter has been widely used successfully by researchers in signal denoising. It smooths the signal by convoluting subsets of adjacent data points with a low-degree polynomial using linear least squares. Hence, from a signal $S_j = f(t_j), (j = 1, 2, \ldots, n)$ having length n, S–Golay filter can be computed using Eq. 1.

$$Q_j = \sum_{i=-\frac{m-1}{2}}^{\frac{m-1}{2}} c_i S_{j+i}, \quad \frac{m+1}{2} \le j \le n - \frac{m-1}{2} \tag{1}$$

where m is the length of the subset called frame span, C_i's are the convolution coefficients and Q_j the denoised signal. For our data, we have used frame size (m) to be 5 and convoluted it with a quadratic polynomial.

3.2.2 Digital filter

In this experimental study, the goal was to analyze the contribution of different brain wave frequency bandwidths and apply it in discriminating different instructions that the subject thought during the trials. There are five frequency bands of brain waves, and we need to filter these individual brain waves from the principal brain wave.

There are two ways by which we can accomplish this task as described by Le et al. [36]

- Extraction of different wave rhythms by digital filtering
- Extraction of different wave rhythms by Wavelet transform

As described in [36], digital filtering seems to be more apt for the task as the wavelet transform is less effective in rejecting frequencies outside the desired range for each type of rhythm. For this reason, the filtering process was completed using digital filters. A band-pass filter according to the frequency bands for each category of wave rhythm is applied to individually extract the corresponding waveforms. As higher ranges of frequencies are filtered, the filtered signals are 'cleaner,' i.e., less undesired frequencies found in the signal [36].

3.2.3 Cleaning

The data acquired from EMOTIV test-bench needed further cleaning. Cleaning was done by trimming the initial 2 s of data of each part as this denoted the transition time because keeping this transition time could have potentially caused irregularities in classification. To handle the unbalanced class problem, we also need to discard the initial 3 s and final 4 s of no command (relax time) data. This ensures that we have equal size dataset for each class to combat the problem of imbalanced dataset.

3.2.4 Parameter-based pre-processing

This is the last stage of pre-processing, and in this stage cleaned data is subjected to parameter-based processing. The parameters considered for this work are (a) time segment, in which we tune the length of optimum time for a subject to think for a command, and (b) down-sampling, in which the optimum number of samples required per second of data is tuned. The following subsections describe these parameters in detail.

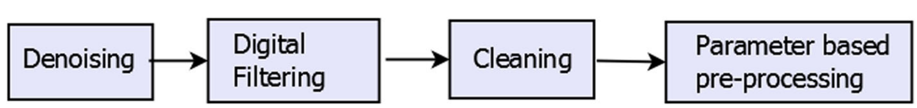


Fig. 4 Work-flow for pre-processing

Time Segment In this experiment, subjects were instructed to close their eyes and think to either turn on the bulb or turn off the bulb for continuous 14 s. Of these 14 s, first 2 s are discarded because those 2 s are utilized in closing their eyes and change in state of mind. In the remaining 12 s, the subject is continuously thinking to turn on/off the bulb. This continuous data of 12 s is divided into several small segments, each of which may fully determine whether the user is thinking to turn on or turn off the bulb.

It is intuitive that 4-s segment should have enough information, such that it easily gets classified into one of the possible classes (on_to_off, off_to_on, do-nothing), but we want an optimum value of time segment, such that it maximizes the number of data samples as well as hold enough information to result in good accuracy. Several possible time segments (4, 2, 1, 0.5) were tried to get a good trade-off between number of data samples and performance.

Down-sampling The data is collected on a default frequency of 128 Hz, which means that in 1 s, the device produces 128 values from all the channels. In this work, we have trained our models on various sampling rates 128, 64, 32, 16, 8, 4, 2 and 1 Hz by down-sampling the original data. However, we achieved optimum performance on 16 Hz and 1 Hz frequency. Down-sampling proved to be the most important step in our experiment. We have tried three different approaches in down-sampling, first approach was taking average, second involved taking maximum and third was by calculating Root Mean Square (RMS) values. We consider 1 s data from a set of data sampled at M Hz (in our case, $M = 128$), Original Dataset:

$$\begin{aligned} S = \{&X_0^0, X_1^0, \ldots, X_{N-1}^0, \\ &X_0^1, X_1^1, \ldots, X_{N-1}^1, \\ &\vdots \\ &X_0^{M-1}, X_1^{M-1}, \ldots, X_{N-1}^{M-1}\} \end{aligned} \tag{2}$$

where N = number of channels (in our case, $N = 14$).

Assume that the data is down-sampled from M to D, where M is divisible by D. In our study, we tried several values of $D = 1, 2, 4, 8, 16, 32, 64$. Let us represent down-sampled dataset as follows:

$$\begin{aligned} S = \{&Y_0^0, Y_1^0, \ldots, Y_{N-1}^0, \\ &Y_0^1, Y_1^1, \ldots, Y_{N-1}^1, \\ &\vdots \\ &Y_0^{D-1}, Y_1^{D-1}, \ldots, Y_{N-1}^{D-1}\} \end{aligned} \tag{3}$$

Down-sampling using maximization is better than averaging or root mean square.

3.3 Proposed hybrid model

We propose a hybrid model comprising of Long Short-Term Memory (LSTM) and Random Forest Classifier (RFC). Before moving to the details of the proposed model, we first describe LSTM fundamentals in Sect. 3.3.1 and Random Forest Classifier in Sect. 3.3.2. Independent LSTM architectures are described in Sect. 3.3.3. Finally, the proposed hybrid model is described in Sect. 3.3.4.

3.3.1 LSTM fundamentals

LSTM is a form of recurrent neural network (RNN) which is more efficient that the traditional RNNs. They are gated recurrent unit which works better than their traditional counter parts by maintaining a cell state and remembering relevant context which represents long dependencies. For more references on LSTM, please refer to these studies [37–39]. An LSTM unit is composed of a cell and three gates, namely input, output, and forget gate. The cell remembers values over arbitrary time intervals and the gate are responsible for controlling the flow of information into the network.

3.3.2 Random forest classifier

Random Forests [40] is an ensemble machine learning technique which uses multiple decision trees to achieve the task of classification or regression. During training phase, it builds many decision trees from the available training data and during testing phase it outputs the result using max-vote strategy from the different decision trees for performing the classification task.

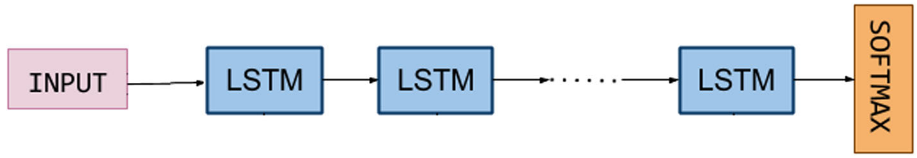
3.3.3 LSTM networks

For this study, we have considered two types of LSTM networks. The first LSTM model (Model A) which is a standard LSTM model that was trained is shown in Fig. 5. This model is a variable sequence length model depending on the input sequence and can be varied dynamically, and it is a parameter which was heuristically optimized to get the optimum length. This model was trained for different possible sequence lengths to determine the best one. Sequence length is determined by the following equation:

$$\text{Sequence\_length} = \text{Time\_segment} \times \text{Sampling\_rate} \tag{4}$$

$$\begin{aligned} \text{Input\_Dimension} = {} & (\text{Time\_segment} \times \text{Sampling\_rate}) \\ & \times \text{Number\_of\_channels} \end{aligned} \tag{5}$$

Fig. 5 Model A: variable sequence length LSTM architecture



Sampling rate is determined empirically and is treated as a parameter for the LSTM model. In the optimal case, Time segment value is found to be 4 and Sampling rate value to be 16. In Sect. 4.1, it is described how these parameters are obtained. So, the sequence_length is 64 (4×16 and Inputdimension $= 64 \times 14$. Each LSTM cell takes an input of size 14 (number of channels).

A drawback of this model is that, it was difficult to take input from various combination of filters such that it could learn time domain information from all the filters. To overcome this drawback and to improve the accuracy, a different LSTM architecture was proposed which could learn from different filters without losing the time domain information; let us call it Model B which is a stacked LSTM model (Fig. 6).

This model is a variable sequence length LSTM and also a variable layer LSTM. Variable layer LSTM means that the architecture is composed of more than one layer, depending upon the types of filters of data we want to train. Layers are organized as shown in Fig. 6. Each layer takes input from one of the filters (none, alpha, beta, theta, delta, gamma). The input to the filters to be fed as input is decided dynamically and is treated as parameters. Sampling rate is treated as another parameter in this model as well.

Sequence-length is determined by Eq. 4, and in the optimal case, sequence-length was found to be 64 and number of layers 6. Input from each layer has same dimension as that of LSTM Model A and is determined by Eq. 5. Number of units in hidden layer of LSTM was kept 32. This results in a mapping from $32 \times 6 = 192$ hidden units to 3 units of Softmax layer at the end of network.

3.3.4 Proposed hybrid LSTM–RFC model

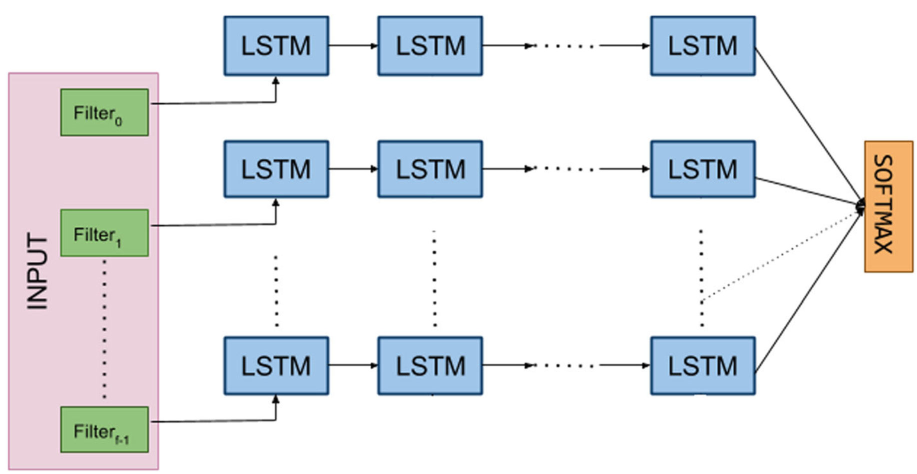
LSTM works well when the data contains more information in its time domain, and the Random Forest Classifier on the other hand learns information represented by the overall data. In this application, we wanted to find to whether the important information is contained in time domain or in the data sample as a whole or both. To deal with this, we built a hybrid model of LSTM and Random Forest Classifier. First, the LSTM model B as discussed in Sect. 3.3.3 is trained, and then in the hybrid model (Fig. 7), features from last LSTM cells are concatenated with original training input. Input data for LSTM is sampled at sampling rate $= 16$, whereas input data for random forest classifier is sampled at sampling rate $= 1$. This was done because the individual models provide the best accuracies on these sampling rates.

3.4 Spatial significance of the electrodes

In this section, we discuss the spatial characteristics of data collected from various channels. From this study, we found the spatial significance of the electrodes is important and needs more attention. We discuss the following major characteristics:

1. Individual channel importance.
2. Channel combination importance.
3. Symmetric correlation.

Fig. 6 Model B: variable sequence length and variable layer LSTM architecture



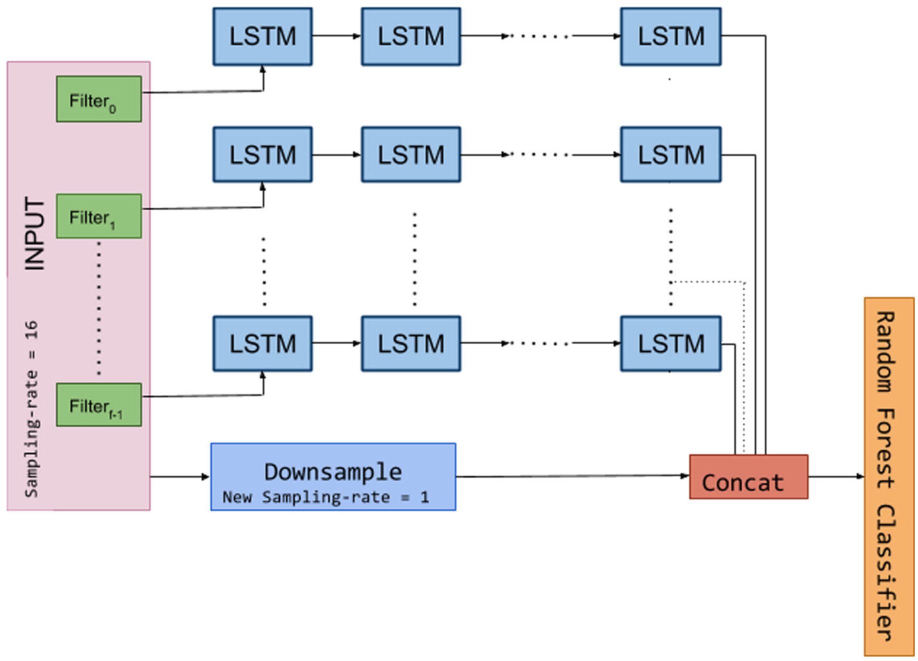


Fig. 7 Proposed LSTM-Random Forest Classifier hybrid model. The proposed hybrid model was developed by combining the stacked LSTM model with random forest classifiers

3.4.1 Individual channel importance

Here, we describe the individual importance of each channel. For this, the performance across each channel was tested to find which channel would be best for discriminating mental imagining of instructions. The main purpose of the channel selection as mentioned in [41] is threefold:

1. To decrease the computational unpredictability of any handling undertaking performed on EEG motions by choosing the pertinent channels and henceforth extricating the highlights of real significance
2. To decrease the measure of overfitting that may emerge because of the use of superfluous channels, to improve the execution, and
3. To decrease the setup time in a few applications.

3.4.2 Channel combination importance

The significance of different electrode spread all through the head varies as per task. The cerebrum is the largest part of the human brain, and is divided into symmetrical left and right hemispheres by a profound groove. Each hemisphere is partitioned into four lobes: the frontal lobe, parietal lobe, temporal lobe, and occipital lobe. Each lobe is related with maybe a couple specialized function; however, there is some functional overlap between them.

So, we aggregate together different EEG signals obtained from frontal left hemisphere and right hemisphere along with temporal left and right hemisphere. We additionally gather together the signals we get from parietal left and right hemisphere with occipital left and right hemisphere. This was done after we found that specific arrangement of channels situated in various regions is more apt for the task of motor imagery and others. Many studies demonstrate that for the mental imagination task, alpha intensities on the right frontal lobe are noteworthy while beta wave intensities on the frontal lobe are connected with concentration. So, in this effort to discover a relationship between our task and the position of the electrode we led our study to get a clear understanding between the two.

3.4.3 Symmetrical correlation

As discussed earlier, we had an intuition that there is some correlation between the task and the channels individually and channels collectively from a specific spatial region of the brain. We found that it was helpful in various studies to find out the correlation between symmetrically placed channels on our scalp. For this, we pick out the symmetric pairs of electrodes and apply three modalities:

1. Ratio of the signals obtained from the symmetric pair.
2. Difference of the signals obtained from the symmetric pair.
3. Product of the signals obtained from the symmetric pair.

The intuition behind this step was that maybe the symmetric pair of electrodes during the task of mental imagining would either both be receiving high amplitudes of signals or low or would be antisymmetric; in either case, the correlation between them could act as a distinguishing feature for our task.

4 Results and discussion

This section firstly discusses the effect of various pre-processing parameters in performance of the model in Sect. 4.1. Next, we compare the performance of the proposed hybrid model with various other models in Sect. 4.2. Finally, we describe the spatial significance of electrodes in Sect. 4.3.

4.1 Effect of pre-processing parameters

In this section, we describe the effect of the following pre-processing parameters:

A. Combined effect of sampling rate and time segments
B. Effect of individual frequency bands
C. Effect of individual sampling types
D. Combined effect of various combinations of sampling types and frequency bands
E. Effect of sequence length in LSTM

4.1.1 Combined effect of sampling rate and time segments

First of all, we calculated accuracies of the model for all possible values of pair (Sampling Rate, Time_Segment) keeping sampling type = 'max' and Filter = None on Random Forest Classifier. Results are shown in Table 1. Filter value None denotes that the data without any applied filter was used as input to the model.

From Table 1, we conclude that we obtain the best results when Time Segment is set to 4 s and Sampling Rate is set to 1.

4.1.2 Effect of individual frequency bands

Next, we compared the performance among all filters using the optimal parameters sampling rate = 1, time segment = 4, sampling-type = max. Result is shown in Fig. 8. Vertical lines show uncertainty in accuracy. In general, it was observed that random forest has uncertainty of up to $\pm 3\%$ in accuracy, so we report average accuracies over 10 different values of number of estimators in random forest.

From Fig. 8, it can be observed that the model performed best when original data (data without applying any filter) was used.

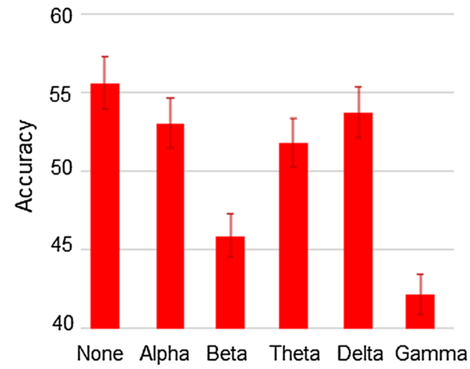


Fig. 8 Accuracies over all the filters using parameters: sampling type = 'max,' time segment = 4, sampling rate = 1, Model = Random Forest Classifier

4.1.3 Effect of individual sampling types

We have also compared different sampling types, namely the average, max and RMS sampling types as discussed in Sect. 3.2.4 over different values of sampling rate keeping Time segment as 4. Results obtained are shown in Fig. 9.

From observing Fig. 9, we infer that Random Forest performs best when sampling type is 'max,' sampling rate is 1.

4.1.4 Combined effect of various combinations of sampling types and frequency bands

We also compared different combinations of sampling types with different combinations of filters. In these combinations, features extracted for one type of sampling type are concatenated with the corresponding filters feature into a single feature vector. Results are shown in Table 2. In Table 2, we only present the most important combinations of filters because other combinations produced poor results.

From Table 2, it can be concluded that we get the best accuracy with random forest classifier when all the filters are used and data is sampled using sampling type = 'max.' In all the subsequent results we use, the parameters sampling-type is set to 'max' and time segment = 4.

4.1.5 Effect of sequence length in LSTM

Here, we discuss the LSTM parameters are tuned and present the effect of sequence length. LSTM Model A gave

Table 1 Accuracies over (Sampling rate, Time_segment) pair using parameters: sampling type = 'max,' Model = Random Forest Classifier

Sampling rate/time segment	128 (%)	64 (%)	32 (%)	16 (%)	8 (%)	4 (%)	2 (%)	1 (%)
4	47.99	48.68	49.48	52.97	53.51	54.41	54.58	55.62
2	48.19	48.58	48.98	50.05	50.87	50.92	52.59	52.46
1	48.36	48.98	49.17	47.42	48.45	49.93	50.32	49.06
0.5	48.49	48.30	47.93	46.70	46.68	47.23	46.76	–

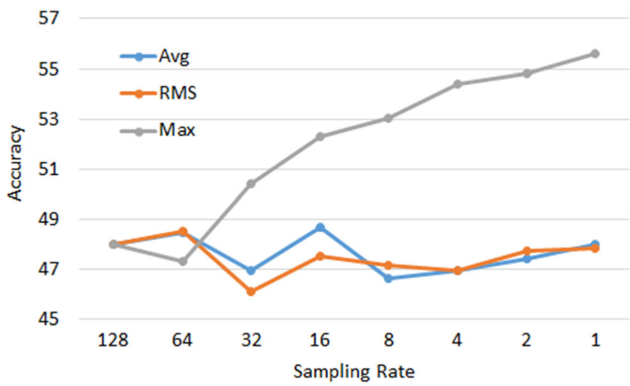


Fig. 9 Accuracies over (sampling-type,sampling rate) pairs using parameters: time segment = 4, Model = Random Forest Classifier, Filter = None

poor results; however, LSTM Model B provided better results when sequence length was set to 64 (Refer Table 3). This suggests that there exists time domain information inherent in the data; however, low accuracy at sequence length = 128 may suggest that the sequence length is too long to learn temporal variations in the dataset. For lower sequence lengths, the data was down-sampled to much lower sampling rates which resulted in decrease in accuracy, suggesting that the time domain information is lost when we down-sample to sampling rates lower than a threshold.

4.2 Performance of hybrid model

Best result on LSTM was obtained when LSTM Model B was trained on all the filters (none, alpha, beta, theta, delta, gamma) at sampling rate = 16. Next, the hybrid model of LSTM and random forest classifier was trained, with sampling-type = 'max,' sampling rate = 16 for input to LSTM and sampling rate = 1 for input to random forest classifier. The total acquired data was split into three subsets in the ratio 60–20–20, where 60% of the total data was used for training the model, 20% for validation and 20% for testing the model. The validation set was used for tuning the parameters of the model. For comparison, we have also trained SVM on the dataset having all the filters concatenated at sampling rate = 1 but the SVM model gave very poor accuracy. We also observe that the proposed hybrid model has the best performance. Results are shown in Table 4.

4.3 Spatial significance of electrodes

In Sect. 3.4.2, we discussed about how various spatial regions of our brain are specifically more apt for the task of cognition or language comprehension, etc. This is an important part of our study, and here we present our findings. We therefore grouped together the processed signals obtained from different regions, namely

A. Entire Brain
B. Parietal and Occipital Left Hemisphere
C. Parietal and Occipital Right Hemisphere
D. Parietal and Occipital Lobe
E. Frontal and Temporal Right Hemisphere
F. Frontal and Temporal Lobe
G. Frontal and Temporal Left Hemisphere

Results of this study suggest that the model is definitely getting information from the entire scalp and hence achieves better performance when we use the signal obtained from entire scalp. What is interesting is to see that the signals obtained from the frontal and temporal lobe are more discriminatory than its counterpart parietal and occipital lobe.

Table 2 Accuracies over (sampling types, filter combination) pairs using parameters: time segment = 4, sampling rate = 1, Model = Random Forest Classifier

Sampling types/filters	Avg (%)	RMS (%)	Max (%)	Avg, RMS (%)	Avg, max (%)	RMS, max (%)	Avg, RMS, max (%)
None, alpha, theta, delta	52.60	63.09	66.52	61.77	65.76	65.69	64.96
None, alpha, beta, theta, delta	53.61	63.02	66.25	62.36	65.38	65.17	65.45
None, alpha, theta, delta, gamma	53.99	63.95	66.80	62.67	66.87	66.49	65.76
None, alpha, beta, theta, delta, gamma	53.95	63.99	67.18	62.84	65.86	66.35	65.97

Table 3 Accuracies over sequence length using parameters: model = LSTM B, filters = (none, alpha, theta, delta)

Sequence length	128 (%)	64 (%)	32 (%)	16 (%)	8 (%)	4 (%)
Accuracy	48.61	62.84	56.94	56.94	50.69	54.86

Table 4 Accuracies over various models with parameters: filters = (none, alpha, beta, theta, delta, gamma), time segment = 4, sampling type = max

Models	Random Forest	SVM (%)	LSTM model A (%)	LSTM model B (%)	Hybrid model (%)
Accuracy	67.18	33.38	57.45	67.01	**68.36**

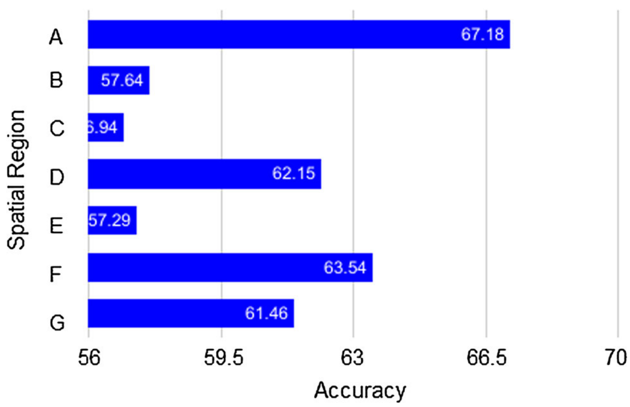


Fig. 10 Channel Position comparison using parameters: time segment = 4, Model = Random Forest Classifier, Filter = None

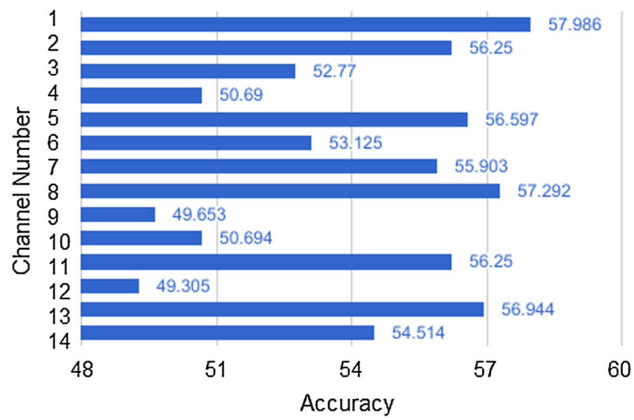


Fig. 11 Individual channel-wise comparison using parameters: time segment = 4, Model = Random Forest Classifier, Filter = None

The result shown in Fig. 10 is not surprising given that frontal lobe is responsible for concentration and focusing. It is also known as the 'control panel' of our body as it is used to control various parts of our entire body and bring our thoughts to action. Though the human subjects are stationary, they do have to think about the three different instructions where they continuously convert their imagination to thoughts of switching on or off the light bulb.

It is also observed that more discriminatory signals come from the left hemisphere of frontal and temporal lobe than the right hemisphere. The reason behind is beyond our knowledge, and more investigation is beyond the scope of this study. What we can claim is that since the frontal lobe is responsible for concentration and translating our thoughts to action, it is more apt for our experiment.

As discussed in Sect. 3.4.1, experiments were conducted to see which channel was more apt for this task of mental imagery. We observe from Fig. 11 that AF3, O2 electrode both of them crossing the 57% threshold seem to be getting the signals from the brain more apt for distinguishing the three signals. These electrodes are present in the pre-frontal left hemisphere and occipital right hemisphere, respectively. This result agrees with the general theory that we need to make use of our frontal lobe as well as pre-frontal lobe for cognitive tasks.

The interesting point observed here is that the occipital right hemisphere electrode 'O2' seems to be receiving enough discriminatory information. After studying through the different literature in this field, we did find out that occipital lobe is responsible for processing visual information. In our experiment, we show our subject an image of bulb and ask them to imagine the same while thinking to switch it on or off; the occipital lobe must have been used for processing that image and imagining of switching that bulb on or off, hence we acquired those discriminatory signals.

4.4 Symmetric correlation

Tests were conducted to find out the correlation between the symmetric pairs of electrodes and whether their correlation could provide us with some new information which could be used for our experiment. There were three modalities that we employed which were previously described. So correlation factors were calculated between each pair of symmetric electrodes, namely:

$$\begin{aligned} SC = \{ & (AF3, AF4), (F7, F8), (F3, F4), (FC5, FC6), \\ & (T7, T8), (P7, P8), (O1, O2)\} \end{aligned} \tag{6}$$

These correlations were also compared by using the best method. The best method for achieving this is the MAX method which is simply taking the maximum of the two values. We had earlier seen that during down-sampling, taking maximum values provided us with better result so this prompted us to use the same to compare our result against it. We have also normalized our features before finding out the correlation.

Multiplication of the values obtained from the symmetric electrodes seems to perform equally well as the

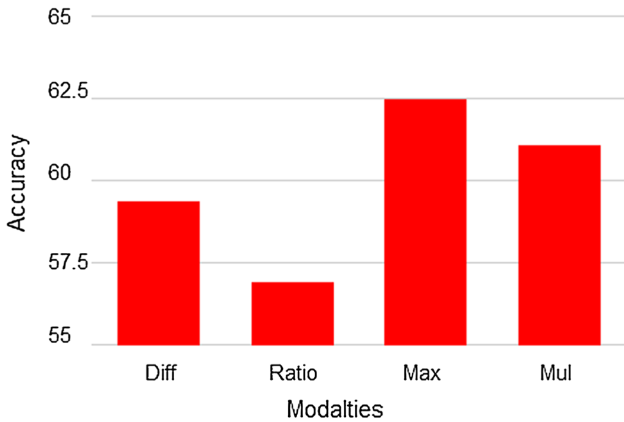


Fig. 12 Symmetric correlation between channels with parameters: time segment = 4, Model = Random Forest Classifier, Filter = None

MAX method, while the ratio between the two seems to be the worse. The reason behind this seems to be the fact that the signal strength for each of the electrode was nearly the same; their difference and ratio would have turned out to be nearly zero or one, respectively, in all cases. But since in the case of multiplication we multiply the intensity of the signal obtained from the symmetric electrode, the net intensity is relatively high and hence does not lose the information. The results of using different modalities are shown in Fig. 12.

5 Conclusion

In this paper, we have explored the feasibility of AI algorithms for home automation using IoT devices. We have explored the behavior of EEG signals and tried to extract information from the signals. We compared different machine learning models for this task and observed that RFC- and LSTM-based AI techniques perform superior in comparison with SVM which performs poorly. Since the acquired data is nonlinear, support vector machines are ill-fitted for this task and even the SVM with quadratic kernel performs poorly which may be because of the high dimensionality of the data. Our study suggests there is a need for future work where an optimal subset of feature from the high-dimensional EEG data could be extracted.

We also observe that the sampling rate plays an important role in the classification task. Both Random Forest Classifier and LSTM converge to an optimal solution when we set the sampling rate to 1 Hz and 16 Hz, respectively. This suggests that there is important information stored in the time domain of the acquired data. In the proposed hybrid model, we stack the LSTM and Random Forest, wherein we send the features extracted from LSTM along with original data to RFC; there seems to be a small improvement over the previous methods.

We were able to get some insight into functioning of the human brain and how brain waves can be acquired and information can be extracted. We have also got some insight into the human decision-making process and how different parts of the brain are responsible for controlling different parts and performing different tasks. However, there are some unanswered questions like why the signals from the left hemisphere of the frontal and temporal lobe are more discriminatory than the right hemisphere? These sort of issues can be taken up by researchers in future, and that knowledge will help us to build more smarter and intelligent interfaces for interacting with the environment.

The issues listed in this paper are only the tip of the iceberg. This is vastly an unexplored field of research, and there is lot of scope for future improvement. As future work, one can try to devise better methods for feature extraction and collect more data to create better models for classification using temporal CNNs or other advance time series models. The signals can also be transformed into other domains, and more robust features can be extracted. Another possible extension can be EEG devices can be used simultaneously with other data-acquiring devices, and input from multiple sensors can be fused to achieve full home automation. We have proposed a hybrid classifier where we have combined LSTM and RFC to achieve better accuracy. In future, this proposed hybrid classifier can be applied to control multiple devices and to other applications belonging to different domains.

Acknowledgements The authors would like to acknowledge Mr. Tauqueer Ahmad, Mr. Shubham Kumar Pandey and Mr. Prakhar Pandey for their help in designing the protocol of the experiments.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. Konstantinidis E, Conci N, Bamparopoulos G, Sidiropoulos E, De Natale F, Bamidis P (2015) Introducing neuroberry, a platform for pervasive EEG signaling in the IoT domain. In: Proceedings of the 5th EAI international conference on wireless mobile communication and healthcare, pp 166–169. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering)
2. Ghodake AA, Shelke S (2016) Brain controlled home automation system. In: 2016 10th international conference on intelligent systems and control (ISCO). IEEE, pp 1–4
3. Shivappa VKK, Luu B, Solis M, George K (2018) Home automation system using brain computer interface paradigm based on auditory selection attention. In: 2018 IEEE international instrumentation and measurement technology conference (I2MTC). IEEE, pp 1–6

4. Holzner C, Guger C, Edlinger G, Gronegress C, Slater M (2009) Virtual smart home controlled by thoughts. In: 2009 18th IEEE international workshops on enabling technologies: infrastructures for collaborative enterprises. IEEE, pp 236–239
5. Lee WT, Nisar H, Malik AS, Yeap KH (2013) A brain computer interface for smart home control. In: 2013 IEEE international symposium on consumer electronics (ISCE). IEEE, pp 35–36
6. Berns GS, Moore SE (2012) A neural predictor of cultural popularity. J Consum Psychol 22(1):154–160
7. Stanton SJ, Sinnott-Armstrong W, Huettel SA (2016) Neuromarketing: ethical implications of its use and potential misuse. J Bus Ethics 144(4):1–13
8. Telpaz A, Webb R, Levy DJ (2015) Using EEG to predict consumers' future choices. J Market Res 52(4):511–529
9. Guger C, Schlogl A, Neuper C, Walterspacher D, Strein T, Pfurtscheller G (2001) Rapid prototyping of an EEG-based brain–computer interface (BCI). IEEE Trans Neural Syst Rehabil Eng 9(1):49–58
10. Neuper C, Scherer R, Wriessnegger S, Pfurtscheller G (2009) Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain–computer interface. Clin Neurophysiol 120(2):239–247
11. Kim B, Kim L, Kim YH, Yoo SK (2017) Cross-association analysis of EEG and emg signals according to movement intention state. Cogn Syst Res 44:1–9
12. Yeh SC, Hou CL, Peng WH, Wei ZZ, Huang S, Kung EYC, Lin L, Liu YH (2018) A multiplayer online car racing virtual-reality game based on internet of brains. J Syst Arch 89:30–40
13. Losonczi L, Márton LF, Brassai TS, Farkas L (2014) Embedded EEG signal acquisition systems. Procedia Technol 12:141–147
14. Bueno L, Pons JL, Bastos Filho TF (2013) An embedded system for an EEG based BCI. In: Biosignals and biorobotics conference (BRC), 2013 ISSNIP. IEEE, pp 1–5
15. Pinho F, Cerqueira J, Correia J, Sousa N, Dias N (2017) Mybrain: a novel EEG embedded system for epilepsy monitoring. J Med Eng Technol 41(7):564–585
16. Kawala-Janik A, Pelc M, Podpora M (2015) Method for EEG signals pattern recognition in embedded systems. Elektron Elektrotech 21(3):3–9
17. Zhang Y, Zhou G, Jin J, Zhao Q, Wang X, Cichocki A (2016) Sparse bayesian classification of EEG for brain–computer interface. IEEE Trans Neural Netw Learn Syst 27(11):2256–2267
18. Zhang Y, Wang Y, Jin J, Wang X (2017) Sparse bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. Int J Neural Syst 27(02):1650,032
19. Gupta A, Sahu H, Nanecha N, Kumar P, Roy PP, Chang V (2019) Enhancing text using emotion detected from EEG signals. J Grid Comput 17(2):325–340
20. Das BB, Kumar P, Kar D, Ram SK, Babu KS, Mohapatra RK (2019) A spatio-temporal model for EEG-based person identification. Multimed Tools Appl 78(19):28157–28177
21. Gauba H, Kumar P, Roy PP, Singh P, Dogra DP, Raman B (2017) Prediction of advertisement preference by fusing EEG response and sentiment analysis. Neural Netw 92:77–88
22. Yadava M, Kumar P, Saini R, Roy PP, Dogra DP (2017) Analysis of EEG signals and its application to neuromarketing. Multimed Tools Appl 76(18):1–25
23. Kaur B, Singh D, Roy PP (2016) A novel framework of EEG-based user identification by analyzing music-listening behavior. Multimed Tools Appl 76(24):1–22
24. AlShu'eili H, Gupta GS, Mukhopadhyay S (2011) Voice recognition based wireless home automation system. In: 2011 4th international conference on mechatronics (ICOM), pp 1–6. https://doi.org/10.1109/ICOM.2011.5937116
25. Costa EJ, Cabral EF (2000) Eeg-based discrimination between imagination of left and right hand movements using adaptive gaussian representation. Med Eng Phys 22(5):345–348
26. Hashimoto Y, Ushiba J (2013) EEG-based classification of imaginary left and right foot movements using beta rebound. Clin Neurophysiol 124(11):2153–2160
27. Khurana V, Kumar P, Saini R, Roy PP (2018) EEG based word familiarity using features and frequency bands combination. Cogn Syst Res 49:33–48
28. Zhou SM, Gan JQ, Sepulveda F (2008) Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface. Inf Sci 178(6):1629–1640
29. Mahanta MS, Aghaei AS, Plataniotis KN, Pasupathy S (2010) Spatio-spectral sufficient statistic for mental imagery EEG signals. In: The 2010 international joint conference on neural networks (IJCNN), pp 1–7. https://doi.org/10.1109/IJCNN.2010.5596467
30. Siuly, Li Y Wen P (2011) EEG signal classification based on simple random sampling technique with least square support vector machine. Int J Biomed Eng Technol 7(4):390–409
31. Li Y, Wen PP et al (2011) Clustering technique-based least square support vector machine for EEG signal classification. Comput Methods Programs Biomed 104(3):358–372
32. Siuly S, Kabir E, Wang H, Zhang Y (2015) Exploring sampling in the detection of multicategory EEG signals. Comput Math Methods Med. https://doi.org/10.1155/2015/576437
33. Bajaj V, Pachori RB (2012) Eeg signal classification using empirical mode decomposition and support vector machine. In: Proceedings of the international conference on soft computing for problem solving (SocProS 2011) December 20–22, 2011. Springer, pp 623–635
34. Aslan K, Bozdemir H, Şahin C, Oğulata SN, Erol R (2008) A radial basis function neural network model for classification of epilepsy using EEG signals. J Med Syst 32(5):403–408
35. Bashivan P, Rish I, Yeasin M, Codella N (2015) Learning representations from EEG with deep recurrent-convolutional neural networks. arXiv preprint arXiv:1511.06448
36. Le HTS, Kamaruzaman MHB, Shukri NBA, Asiri SAB (2015) Lab report signal processing electroencephalogram (EEG). Technical Report KED 130002, University of Malaya. http://www.academia.edu/11959194/EEG_Signal_Processing. Accessed July 2018
37. Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: continual prediction with lstm. Neural Comput 12(10):2451–2471
38. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
39. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780
40. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
41. Alotaiby T, El-Samie FEA, Alshebeili SA, Ahmad I (2015) A review of channel selection algorithms for EEG signal processing. EURASIP J Adv Signal Process 2015(1):66