Feature Extraction Methods for Electroencephalography based Brain-Computer Interface: A Review

Dipti Pawar, *Member, IAENG,* and Sudhir Dhage

Abstract—Introduction: **A brain-computer interface (BCI) is a rapidly growing cutting-edge technology in which a communication pathway is built between the human brain and computer. The BCI is also known as a direct neural interface where user can control external devices with the help of the brain signals. Neural signals are typically measured using electroencephalography (EEG).**
Objective: **Feature extraction from EEG data performs a significant role in the wearable BCI computing field. Since a large amount of EEG data, the major challenge is the effective feature extraction and reduce the computation burden. The objective of this paper is to review such different feature extraction techniques for the development of effective and robust BCI systems.**
Approach: **We reviewed feature extraction techniques employed in EEG based BCI studies. We synthesize these studies in order to present the taxonomy and report their usage with pros and cons.**
Significance: **This paper provides a comprehensive review of feature extraction techniques for EEG based BCI with their properties. Furthermore, open challenges are also discussed for further advancement in BCI studies.**

Index Terms—**Brain-Computer Interface, Electroencephalography (EEG), Feature Extraction, Brain-Machine Interface.**

I. Introduction

Brain-Computer Interface (BCI) is a system which provides a direct communication pathway between the human brain and computers/machines by translating brain activity patterns into commands or messages for an interactive application. In general, brain activity patterns are measured using electroencephalography (EEG) which is noninvasive, easy to use and low cost a method. For instance, with the help of BCI it is possible to control computer without any physical activity and thus fundamentally useful in various applications e.g notably for motor impaired users to control assistive technologies by imagining the motor movement such as controlling wheelchairs [1], or controlling cursors [2], games [3], rehabilitation devices for stroke patients [4].

Figure 1 shows BCIs are as Pattern Recognition system. The system consists of two phases: 1) Training phase 2) Testing phase. The system is calibrated in the training phase while in the testing phase system recognize a neural pattern and translate them into useful commands for a device. Initially, EEG patterns are captured using the signal acquisition system, which may also perform artefact processing and noise removal. Then these EEG signals are preprocessed using spectral and spatial features, followed by feature extraction technique which is compact representation of a signal. Lastly, these features are classified and classified signals translated into meaningful commands to control device.

The most difficult step in pattern recognition is still as it was stated some 65 years ago by [5], 'The extraction of significant features from a background of irrelevant details'. To date, very interesting BCI reviews are published including classification algorithm as in [6], [7]. However, none has been particularly dedicated to feature extraction techniques used for BCI and their properties. This study aims at filling this lack. Therefore one of the major objectives of this review is to study various feature extraction techniques utilized in BCI research and to analyse their critical properties. Another objective is to discuss some challenges which are important to solve considering future progress in BCI research.

The remainder of this review paper proceeds as follows. Section II describes the taxonomy of feature extraction methods used in EEG-based BCI. Section III reviews a wide range of feature extraction techniques for EEG based BCI. Section IV deals with discussion along with open challenges which are important for advancement in the BCI system. Finally, Section V concludes the paper.

II. Taxonomy of Feature Extraction Methods

EEG based BCI systems involve the extraction of useful information from the highly complex EEG data. In general, it is achieved by applying suitable feature extraction on EEG signals which are acquired by subjects during performing a specific mental activity. Subsequently, these features are fed to classifier for the training, where it gets learning to identify the pattern class. The EEG signals are highly nonstationary and dynamically changes due to technical and biological effects such as subject attention, the sessions variability, mental state, anatomical differences among subjects, amplifier and ambient noise [8]. Moreover, from the cognitive neuroscience perspective, the oscillatory (frequency) components of EEG have non-stationary and distinct characteristics. Due to these reasons, it is difficult to classify BCI patterns accurately. Therefore to enhance the performance of EEG based BCI systems, the selection of appropriate feature extraction technique is a very important issue.

The taxonomy of well established feature extraction methods that are applied to EEG is outlined in Figure 2. While there are many ways the two most commonly used feature extraction methods are time and frequency domain features.

Dipti Pawar is with the Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, 400058 India. e-mail:dipti.rohan.pawar@gmail.com.
Sudhir Dhage is with the Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, 400058 India. e-mail:sudhir_dhage@spit.ac.in.

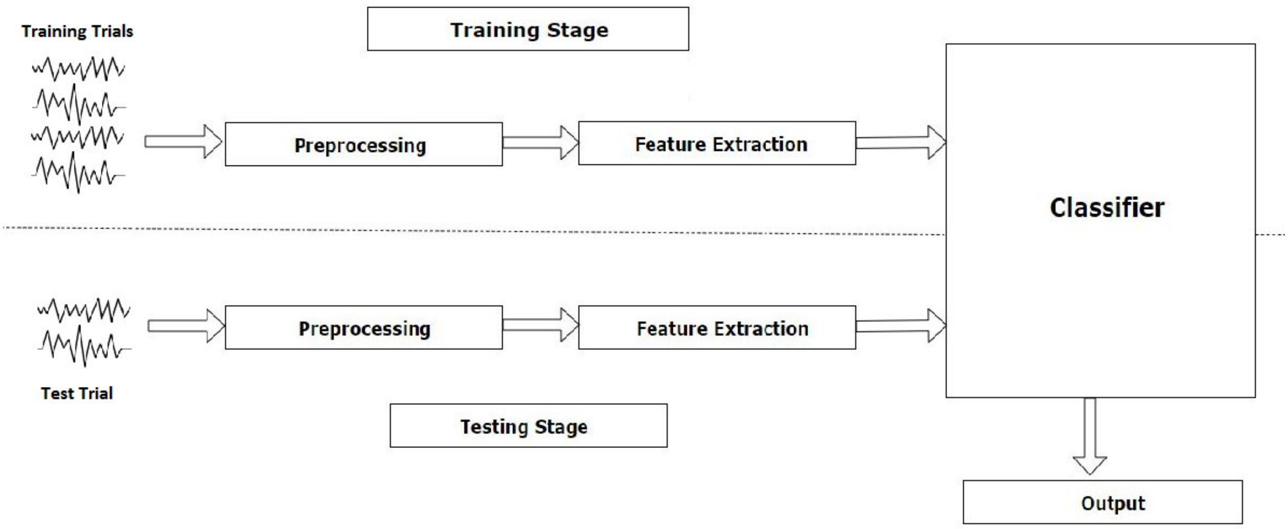

Fig. 1: BCI-Pattern Recognition System

Temporal features i.e time domain features represents employing the EEG signal values at distinct time windows or at distinct time points. Frequency domain features are also called spectral features which represent the signal power in the specific frequency band. However, for the EEG signals having nonstationary in nature, time-frequency methods are useful, which can provide useful information by taking into consideration the dynamic changes. Spatial features deal with the spatial representation of the signal i.e the selection of most appropriate channels selection for the specific task.

III. Materials and Methods

This section details a review of different feature extraction techniques applied for EEG based BCI in the past. In general, these feature extraction techniques are with different domains i.e time, frequency, time-frequency and spatial features. The EEG signals are recorded of subjects with a specific mental activity using the signal acquisition system of multiple channels. Moreover, dimension reduction techniques are also presented since the information provided by all the channels may not be significant for the underlying phenomena of interest.

A. Time Domain Features

The preprocessing such as low-pass filtering or bandpass filtering and downsampling followed by the extraction of time domain features. These features measure the temporal variations within time-locked EEG signal amplitudes. Table I listed the time domain features along with their properties.

1) Hjorth Parameters: Hjorth parameters provide a fast way of computing three important characteristics of a time-varying signal, namely, Activity, Mobility, and Complexity. Activity parameter of the signal z(t) is computed as in equation 1, which is the variance of the time function, which designates the power spectra surface in the frequency domain.

$$Activity = var\left(z(t)\right) \tag{1}$$

Mobility parameter represents standard deviation proportion or mean frequency of power spectrum. Mobility is computed as in equation 2, where $z^{'}(t)$ represents the first derivative of the signal z(t).

$$Mobility = \sqrt{\frac{var\left(z'\left(t\right)\right)}{var\left(z\left(t\right)\right)}} \tag{2}$$

Complexity parameter computed as in equation 3, which indicates the signal shape similarity to a pure sine wave, where value converges to 1 as the similarity index is high.

$$Complexity = \frac{Mobility\left(z^{'}\left(t\right)\right)}{Mobility\left(z\left(t\right)\right)} \tag{3}$$

Activity, Mobility, and Complexity are also called as ‘mean power’, ‘the root mean square frequency’, and the ‘root mean square frequency spread’ respectively.

2) Statistical Features: Various statistical measures characterize the EEG time series. The following six features are widely used in BCI studies and computed by considering N i.e the number of data samples and the signal x(i).

1) Energy: The energy of signal is calculated as in equation 4.

$$\sum_{i=1}^{N} x_i^2 \tag{4}$$

2) Entropy: Entropy computes randomness in the signal and computed as in equation 5.

$$-\sum_{i=1}^{N} x_i log_2\left(x_i\right) \tag{5}$$

3) Mean: The mean (μ) which is average of the signal calculated as shown in equation 6.

$$\frac{1}{N}\sum_{i=1}^{N} x_i \tag{6}$$

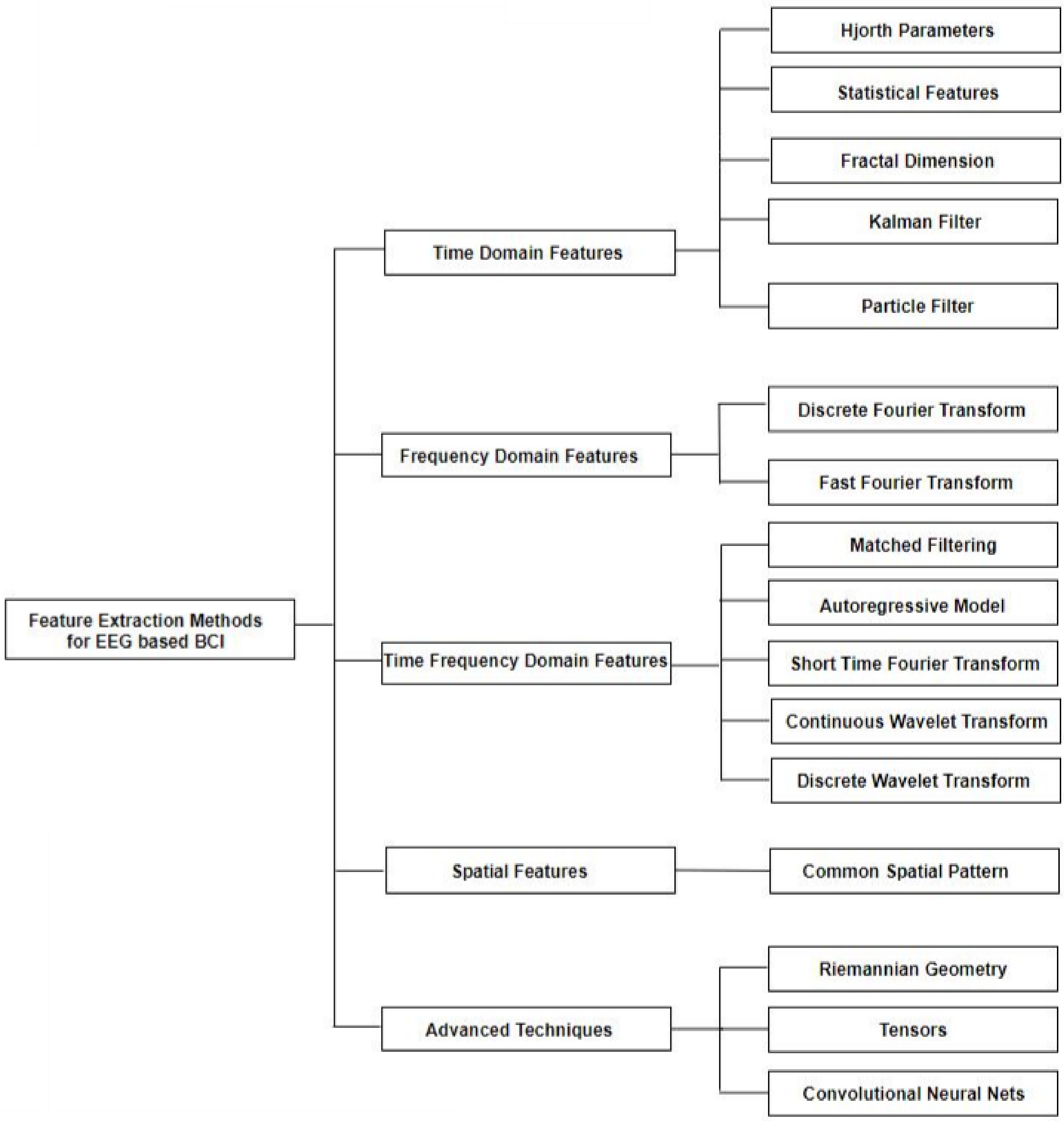

Fig. 2: Taxonomy of Feature Extraction Methods for EEG based BCI

4) Std Deviation: The standard deviation measures how the data $\{x_1, x_2, ..., x_N\}$ are spread out. The standard deviation is computed as in equation 7.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2} \tag{7}$$

5) Skewness: The asymmetry of the data samples around mean is calculated with skewness as shown in equation 8.

$$skewness = \frac{\mu_3}{\sigma^3} \tag{8}$$

where μ_3 is the 3^{rd} order moment which is calculated as shown in equation 9.

$$\mu_3 = \sum_{i=1}^{N}(x_i-\mu)^3 \tag{9}$$

6) Kurtosis: Kurtosis measure the 'tailedness' and 'outlier' characteristic of the distribution of data. Kurtosis is computed as shown in equation 10.

$$kurtosis = \frac{\mu_4}{\sigma^4} \tag{10}$$

where μ_4 is the 4^{th} order moment which is calculated as hown in euation 11.

$$\mu_4 = \sum_{i=1}^{N}(x_i-\mu)^4 \tag{11}$$

3) Fractal Dimension: The Fractal Dimension (FD) is a statistical index which represents the self-similarity measure of signal over some space or time interval. EEG has a fractal nature, hence fractal pieces can be used to obtain features. Since the complexity and limited predictability of EEG signals, Fractal dimension measure provides a 'complexity' of the time-varying brain signal. Several classical methods have been proposed to compute the FD value such as

TABLE I: Time Domain Features

Method	Properties
Hjorth Features	•Low computational complexity •Suitable for Stationary signal •Fast to compute •Easy to apply
Statistical Features	•Easy to apply •Suitable for Stationary signal •Can apply to non-stationary signal in conjunction with freuency features
Fractal Dimension (FD)	•Measures the self similarity of signal •Provides complexity index •Evaluation is time consuming
Kalman Filter	•Suitable foe EEG resource localization problem •Best known of Bays filter which measures the uncertainty of a signal
Particle Filter	•Particle Filter scales well •Computationally expensive •Non-deterministic

apparent entropy, Kolmogorov-Sinai entropy and Correlation dimension. However, the evaluation of these methods is time-consuming.

4) Kalman Filter: Uncertainty representation is important in BCI because potentially disastrous actions based on poor estimates can be avoided if the amount of uncertainty associated with an estimate is taken into account before committing to a decision. Bayesian filtering techniques provide a statistically sound methodology for estimating signal properties and their uncertainty. The Kalman filter is perhaps the best known of Bayesian filtering algorithms. The time-varying coefficients can be updated online using a recursive least-square optimization procedure such as Kalman filtering. The coefficients capture the local statistical structure of the signal as it evolves over time and can be used as features in further processing e.g classification in a BCI. Kalman filters assume that the dynamics and measurement process are linear and Gaussian. This simplifying hypothesis may not dominance true in many real-world examples.

5) Particle Filter: EEG neural signals acquired from human scalp are nonlinear in nature. However, the various linear regression model are unable to reflect the nonlinear component of EEG. To overcome this drawback nonlinear decoding model i.e particle filter can be used. A particle filter is a method of estimating a posterior distribution over the hidden state for non-linear non-Gaussian processes. It is achieved by applying Monte Carlo simulations based recursive Bayesian filter. The particle filtering is easy to implement and scales very well i.e embarrassingly parallel. Nevertheless, the shortcoming of particle filtering is computationally expensive since a good particle filter requires a lot of particles. If the distribution is unimodal, it is good practice to use a Kalman filter. Particle filters are nondeterministic i.e they can produce different outputs for the same input, which make them difficult to predict and debug.

6) State-of-the-art: Though the time domain features are not dominant in EEG BCI, the aforementioned methods are useful to identify the time series analysis of EEG signals. In general the Horth parameters were widely used in motor imagery EEG studies [9], [10]. Hu et al. [11] proposed the investigation of the learners affect during the learning process by applying the horth parameters in conjunction with Autoregressive models (AR) and nonlinear features including Singular-value Decomposition Entropy (SVDen), Approximate Entropy (ApEn), the correlation dimension (D2), the largest Lyapunov exponent, the Kolmogorov entropy, the spectral entropy and C0-complex. The statistical features are widely used in conjunction with other feature extraction techniques like dwt [12]–[14]. Several Fractal dimension computation methods such as Fractal Brownian Motion [15], Sevciks method [16], Higuchi algorithm [17] or Box-counting [18] were employed. In [19] the fractal dimension coupled with dwt is used for EEG signal classification. Fractal dimensions are also employed in seizure detection [104] and classifying depression patients [105]. Luke and Wouters [20] presented Auditory steady-state responses (ASSR) using Kalman filter analysis and illustrate several benefits over DFT methods. Moreover, the Kalman filters were used in Epilepsy patients EEG studies [21], [22]. However, the particle filters were employed in P300-BCI paradigm [23], [24].

B. Frequency Domain Features

The frequency domain features are popular in EEG based BCI which compute the amplitude/power changes from different frequency band. Table II details the frequency domain features along with their properties.

1) Discrete Fourier Transform: The basic idea behind Fourier analysis is the decomposition a signal into a weighted sum of sinusoidal and cosine waves of distinct frequencies. The Fourier decomposition of a signal into its amplitudes represents the signal in terms of frequency content rather than time. The Inverse Fourier Transform (IFT) is used to recover the original signal. For BCI applications, the brain signals are typically sampled at discrete time intervals. In Discrete Fourier Transform (DFT) the Fourier series is modified and apply on discretely sampled signals. The DFT takes as input a time series and sampled at time points $t = 0, 1, \ldots, T-1$ and transforms it to corresponding complex Fourier coefficients which can capture signal information such as amplitude and phase.

2) Fast Fourier Transform: The Fast Fourier Transform (FFT) computes the DFT efficiently. FFT reduced the no. of computations and, thus made processing more affordable and efficient. Many BCI systems rely on features extracted from the power spectrum of a brain signal such as EEG or Electrocorticography (ECoG) over a time interval. The widely used method for power spectrum estimation is welch's method

TABLE II: Frequency Domain Features

Method	Properties
Fast Fourier Transform (FFT)	•Efficiently compute Discrete Fourier Trnasform •Suitable for Stationary signal

(based on FFT), and the power of a specific frequency band is used as a spectral feature in further analysis such as classification. For example, to determine subject-specific frequency bands using motor imagery, the subjects perform different movements, and frequency bands that exhibit robust changes in power during movement are utilized. A more primitive approach is to utilize a bank of spectral features and allow machine learning algorithm to automatically select features that enhance classification accuracy of test data.

The FFT is widely used in emotion recognition EEG studies [25]–[27]. Polat and Gunes [28] proposed epileptic seizure detection system based on FFT. However, Wen and Zhang [29] proposed an EEG analysis of epilepsy patients using frequency domain features in conjunction with non-linear features. The Tinnitus EEG classification proposed by Wang et al. [30] were FFT based power values are used as features. Nevertheless, FFT is also used in BCI game controlling [31]. The commonly used alternative to FFT is the Welch method based power spectral density (PSD) estimation which is applied by Reuderink et al. [33], Kroupi et al. [32]. Nowadays connectivity features are dominantly used in conjunction with frequency domain features. These features compute the synchronization or correlation between signals and/or sensors. Several features are available for this measurement in particular spectral coherence or direct transfer functions or phase locking values, among many others as discussed in [34]–[37].

C. Time-Frequency Domain Features

In general Time-Frequency representations of the signal are the most complete and widely used methods for non-stationary EEG based BCI studies. Table III elaborates the Time-Frequency domain features along with their properties.

1) Matched Filtering: The Matched Filtering (MF) is a feature extraction technique which can detect the specified pattern from the unknown EEG signals based on its match with the templates, where templates are the known signals. The Correlation between templates and unknown EEG signals represents user's intention. The better correlation intends higher matching between intention of the user and template. In general, as shown in equation 12 every matched filter easily modelled as addition of sinusoidal components which are the harmonically related.

$$MF(n) = \sum_{k=1}^{N} a_k \cos\left(\frac{2\pi k f_F}{f_s} n + \phi_k\right) \tag{12}$$

where n denotes the template sample number, N is the total number of harmonics to model, the sampling frequency f_s and fundamental frequency of template f_F. The FFT spectrum is used to derive a_k and ϕ_k which represents the amplitude and phase of individual harmonics respectively.

2) Autoregressive Model: Autoregressive models (AR) based on the fact that its natural tendency of the signals to be likely correlated over time or even other dimensions such as space. Thus it is possible to predict the future measurements based on the past few values. The AR model prediction of the current signal measurement x_t based on past values as in equation 13.

$$x_t = \sum_{i=1}^{p} a_i x_{t-1} + \xi \tag{13}$$

where a_i is the set of coefficients, ξ is presumed as zero mean white noise process that accounts for the differences between the signal and its linear weighted sum approximation. The parameter p denotes the order of the AR model and determines the window of past inputs used for predicting the current input. The traditional AR model assumes the statistical properties of the signal are stationary so that a single set of coefficients a_i can be used. However, brain signals tend to be nonstationary, and one constantly requires a time-varying set of coefficients $a_{i,t}$. This leads to a multivariate adaptive Autoregressive model (MVAAR) as shown in equation 14.

$$x_t = \sum_{i=1}^{p} a_{i,t} x_{t-1} + \xi_t \tag{14}$$

The time-varying coefficients $a_{i,t}$ can be updated using recursive least square optimization algorithm, kalman filtering etc. The coefficients $a_{i,t}$ capture the local statistic structure of the signal as it evolves over time.

3) Short Time Fourier Transform: The Fourier transform represents an original signal with basis functions namely, sines and cosines of different frequencies. However, because sines and cosines occupy an infinite temporal extent, the Fourier transform does a poor job of representing signals that are finite and non-periodic or having sharp peaks and discontinuities. However, brain signals such as EEG are typically non-stationary (i.e statistical properties vary over time), breaking the assumption of a stationary signal in Fourier analysis. One solution is to perform Fourier analysis over short-time windows, a procedure known as short-term Fourier transform (STFT). The STFT deals with the problem of window size, where small windows providing good temporal resolution however poor frequency resolution, while large windows providing better frequency resolution yet poor temporal resolution. This realization leads to wavelet transform which achieves the excellent trade-off among temporal and frequency resolution.

4) Wavelet: The wavelet transform (WT) utilizes finite basis functions called wavelets, such are translated and scaled copies of a single finite length waveform known as mother wavelet. By using basis functions at different scales, the wavelet transform allows a signal to be analysed at multiple resolutions which allow representing signals that are non-periodic or have sharp discontinuities. The wavelets are a most suitable and powerful tool for the transient EEG signal analysis. The Continuous Wavelet Transform (CWT) of signal $x(t)$ with wavelet function as in equation 15.

TABLE III: Time-Frequency Domain Features

Method	Properties
Matched Filtering (MF)	•Detect specific waveform with temporal characteristics •Suitable for specific pattern detection based on its matches with known templates
Autoregressive Model (AR)	•Short time segments with High frequency resolution •Adequate for stationary signals •Spectrum model •MVAAR: Adaptive version of AR
Short Time Fourier Transform (STFT)	•Adequate for nonstationary signal •Provides frequency as well as temporal information •deals with window size problem
Continuous Wavelet Transform (CWT)	•Adequate for non-stationary signals •Provides frequency as well as temporal information
Discrete Wavelet Transform (DWT)	•Reduces the complexity and redundancy of CWT •Provides frequency as well as temporal information •Adequate for non-stationary signals

$$W(S,T) = \int_{-\infty}^{\infty} x\left(t\right) \psi^{*}_{S,T}\left(t\right) dt \tag{15}$$

Where $W(S,T)$ is the wavelet coefficient of the signal having frequency with scale S and time T of the wavelet function $\psi^{*}_{S,T}\left(t\right)$. The symbol * denotes the complex conjugate. The wavelet function $\psi^{*}_{S,T}\left(t\right)$ is a dilated and shifted version of mother wavelet $\psi\left(t\right)$.

$$\psi_{S,T}\left(t\right) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-T}{s}\right) \tag{16}$$

A mother wavelet has ability to make multiple shapes and satisfies the condition in equation 17.

$$W(S,T) = \int_{-\infty}^{\infty} \psi\left(t\right) dt \tag{17}$$

Moreover, The CWT is like template matching or matched filter where the cross variance between the signals is calculated. Wavelet template is always preferable than classical template matching due to its special properties. However, CWT is more complex and redundant since it involves the signal analysis with multiple dilations and shifting to mother wavelet at a high number of frequencies. The Discrete Wavelet Transform (DWT) translates and dilates mother wavelet at discrete values and hence overcome the shortcoming of CWT. Though the DWT is more popular in BCI studies, CWT is still applied in EEG based BCI research. One of the reason may be CWT provides subtle information where dwt inadequate to extract such information.

Wavelet analysis is done by corresponding coefficients. Recent signal-processing packages include the wavelet transform as one of the available options and provide a variety of choices for mother wavelet. Mother wavelet selection depends on the BCI application and the sort of features required to extract from the signal.

5) State-of-the-art: Liang et al. [38] proposed an automatic sleep scoring method combining autoregressive (AR) models and multiscale entropy (MSE) for single-channel EEG. Zhang et al. [39] proposed EEG signals Classification based on wavelet packet decomposition and autoregressive model. Hatamikia et al. [40] proposed EEG based emotion recognition system using the autoregressive model and sequential forward feature selection. Moreover, the AR has also applied in event-related paradigm [41], [42]. Gomez-Herrero et al. [43] proposed the study on measuring directional coupling between EEG sources using MVAR. STFT is applied in Epilipsy patients EEG studies [44]–[46]. Behnam et al. [47] investigated in their study regarding the EEG activity in Autism spectrum disorder using FFT and STFT. Hadjidimitriou and Hadjileontiadis [48] employed three methods, namely STFT based spectrogram (SPG) the Zhao-Atlas-Marks (ZAM) distribution, and the Hilbert-HuangSpectrum (HHT) in the study regarding the recognition of EEG based music like preference. Faust et al. [49] explored the review of CWT and DWT regarding epilepsy diagnosis and seizure detection. The DWT is widely used in EEG studies rather than CWT. Though it is not possible to list all the dwt based research, we attempt to list dwt based EEG studied for different applications [12]–[14], [106]–[109]. Matched filter is employed for new-born sizure detection [110] and event related brainwave extraction from EEG signals [111].

D. Common Spatial Pattern

Common Spatial Pattern (CSP) has been emerged as a popular feature extraction method for EEG based BCI, where the similarities between classes are minimized and differences are emphasized. CSP finds spatial filters which can transform the input data into resulting feature vectors that enhance the discriminability between classes. Though fundamentally CSP has been intended for the multichannel data allied to two class problems, few extensions have been also suggested for multiclass BCI data [50]. CSP is more suitable for synchronous BCI which is restricted to time frames. However, it is unable to provide the similar performance in asynchronous BCIs. Also, the spatial resolution affects the performance since the few electrode positions provide more discriminating information for particular brain activities compared to others. Considering these issues, the methods have been suggested to enhance the performance of CSP: Common Sparse Spectral-Spatial Pattern (CSSSP) [53], Common Spatio-Spectral Pattern (CSSP) [52] and Wavelet Common Spatial Pattern (WCSP) [51]. Table IV describes the CSP properties.

Consider the data $\{X_c^i\}_{i=1}^k$, where i is the trial for the class $c\epsilon\{1,2\}$. Each X_c^i is a $N \times T$ matrix, where N denotes number of channels and the number of samples per channel are denoted by T. The CSP aims at finding M spatial filters

TABLE IV: Common Spatial Patterns

Method	Properties
Common Spatial Patterns (CSP)	•Spatial filter designed for 2-class problems Multiclass extensions are also exist •WCSP, CSSP, CSSSP: Enhanced versions of CSP •Spatial resolution affects the CSP performance

where a $N \times M$ matrix W, which is linearly transforming the input signal as in equation 18.

$$X_{csp}(t) = W^T X(t) \tag{18}$$

where $X(t)$ is the all channels input signals at time t. With the intention of finding all filters the conditional covariance matrices of two classes are calculated as in equation 19.

$$R_i = \frac{1}{K}\sum_{i=1}^{k} X_c^t (X_c^i)^T \tag{19}$$

The matrix W can be determined for $c\epsilon\{1, 2\}$ as in equation 20 and 21.

$$W^T R_1 W = \Lambda 1 \tag{20}$$

$$W^T R_2 W = \Lambda 2 \tag{21}$$

where the Λi are diagonal metrices and $\Lambda 1 + \Lambda 2 = I$, where I is the identity matrix. It can be achieved by resolving generalized eigenvalue problem as shown in equation 22.

$$R_1 W = \lambda R_2 W \tag{22}$$

where the generalized eigenvectors $w = w_j$ satisfy the equation 22 and computes the coloums of W which represents the CSP spatial filters. $\lambda_1^j = W_j^T R_1 W_j$ and $\lambda_2^j = W_j^T R_2 W_j$are the generalized eigenvalues which form the diagonal elements of Λ_1 and Λ_2 respectively. Since $\lambda_1^j + \lambda_2^j = 1$, a high value for λ_1^j represents that the filter output based on filter w_j generates a higher variance for input signals in class 1 and low variance for signals in class 2 or vice versa.

1) State-of-the-art: Till date lot of motor imagery based BCI research has been done using CSP [56], [57]. In general, the spatial filtering used in conjunction with time point features or band power for EEG based BCIs. Several variants have been proposed which are robust to nonstationarity or noise, using robust data averaging, regularization approaches and/or new divergence measures [55]. Nevertheless, various extensions of this algorithm are proposed to optimize spatial and spectral filters simultaneously as proposed in [54], [58].

E. Feature Selection

Appropriate selection of features from the set of extracted features having potential benefits in EEG based BCI as:
1. Feature selection removes redundant and irrelevant features of the targeted mental states.
2. Feature selection improving the performance of the learning process since fewer parameters need to be optimized by the classifier.
3. The number of training samples is small in BCI studies; hence feature selection overcome the overfitting problem by reducing the number of features.
4. It reduces the data dimensionality which helps the faster model building of classifier, reducing computational cost and often producing more compact and easier to model interpretability.
5. From the set of selected features, it can be easy to observe more correlated features w.r.t. intended mental state.

In general, the feature selection methods divided into the filter, wrapper and embedded models. The Filter model selects the features independently of the classifiers to be used where the bias of the classifiers and feature selection algorithm are independent. It depends on the training data characteristics such as correlation, dependency, information, distance and consistency. Koprinska [59] has evaluated five filter methods namely Correlation-Based Feature Selection (CFS), 1R Ranking (1RR), information Gain Ranking (IG), ReliefF and Consistency-Based Feature Selection on BCI competition dataset. Ang et al. [4] proposed filtering method based frequency band selection for motor imagery.

Wrapper method selects a subset of features and fed it to the classifier during the learning phase. The resulting performance has been observed and proposes a new subset if the criteria are not satisfied or stop the search as per the stopping criteria. The various wrapper methods are used in EEG based BCI studies are linear regression for knowledge extraction [60], support vector machine for channel selection [61], P300 based feature selection [62], multiresolution analysis based evolutionary algorithms for feature selection [63] and genetic algorithms for spectral feature selection [64]. The shortcoming of wrapper methods is they are computationally expensive for a large number of features.

Since the Filter approaches are computationally inexpensive but having the drawback of redundant feature selection and wrapper methods avoid redundant feature selection but they are computationally expensive, Embedded models are proposed to bridge the gap between them. Initially, as per filter models, it selects several features with a particular cardinality and later on, it selects the subset of features having the highest classification accuracy. Thus embedded model archives comparable computational efficiency to filter model and comparable accuracy compared to the wrapper model. In short Embedded methods achieves the model fitting and feature selection simultaneously. Krusienski et al. [65] used embedded method based on the linear discriminant analysis for P300-BCI research. Moreover, the metaheuristic techniques such as tabu search, ant colony optimization, simulated annealing are widely used for feature selection to avoid curse-of-dimensionality [66], [67].

F. Dimension Reduction

This section contains a brief description of dimension reduction techniques such as PCA and LDA. Since dimension reduction techniques remove the insignificant and redundant information, computational costs are reduced. Table V summarizes dimension reduction techniques along with their properties.

TABLE V: Dimension Reduction

Method	Properties
Principal Component Analysis (PCA)	•Linear transformation •Transforms the data from existing feature space into low dimension feature space •Optimal data representation with reference to minimal mean-square-error •Not always suitable for classification
Independent Component Analysis (ICA)	•Robust and powerful tool for artifact removal •May corrupt the power spectrum •Mixed signals are splitted into its sources

1) Principal Component Analysis: PCA is simple yet popular and useful linear transformation technique which converts the possibly correlated input data into a set of new data points which are uncorrelated. PCA attempts to find the dominant direction of statistical variability of the data. Once these dominant directions corresponding to a low dimensional subspace of the original subspace have been found, new data points can be projected along with these principal directions, where each projection is called the principal component. In the EEG measurements from N electrodes placed on the head, measurements from nearby electrodes may be correlated or there may underlying rhythms that appear across multiple electrodes. Such redundancies can be exploited using PCA. PCA does not always guarantee the discriminative features and thus not optimal for classification purpose. However, with respect to EEG based BCI applications, PCA has been reasonably successful for reconstructing signals without artefacts [68]. Moreover, PCA has been applied for the feature space dimensionality reduction [69].

Consider the training data $t = [t_1, t_2, ..., t_n]$ where t_i is i^{th} d-dimension training sample, and the number of samples n. PCA projects the input training data into m-dimension space of m-dimension vector computed from the covariance matrix C_v as in equation 23.

$$C_v = \sum_{i=1}^{n} (p_i - m_v)(p_i - m)^t \tag{23}$$

where, m_v is the mean vector of the training sample t_i computed as in Equation 24.

$$m_v = \frac{1}{n}\sum_{i=1}^{n} t_i \tag{24}$$

The covariance matrix C_v is symmetric and real matrix $l \times l$, where l are distinct eigenvectors and eigenvalues. The eigenvectors having the highest eigenvalue represent the principal components of the training dataset t. PCA decides eigenvectors q, where $q < l$. These selected eigenvectors provides a projection matrix P which can be useful to extract features from the test set u. The Matrix P having the coloums of sorted l eigenvectors with the maximal eigenvalue is stored at the first coloum of P. Finally, PCA transforms the test data u into new subspace by computing feature vector v from the data in matrix P as in Equation 25.

$$v = P^t (u - m) \tag{25}$$

where, m is the mean computed as in Equation 24.

2) Independent Comonent Analysis: Independent Component Analysis (ICA) is a method to compute independent sources from the set of mixed signals, where it assumes that underlying unknown sources are mutually statistically independent. The EEG signal is the mixture of independent sources signals including artefacts. Consider the EEG signal $x(t)$, where ICA attempts to find independent sources as per equation 26.

$$x(t) = W(s(t)) + n(t) \tag{26}$$

Where W is the mixing matrix, $s(t)$ and $n(t)$ are source signals and noise vectors respectively. The dimension of $x(t)$ depends on number of channels. The number of sources decide the dimension of $s(t)$. Furthermore, by considering the input data is noiseless, the $s(t)$ and W are obtained by considering Infomax [70] or modifying Infomax [71]. It is noteworthy that Unlike PCA, where the dimensionality of the output vector is always smaller than the dimensionality of the input vector, the feature vector dimension of ICA can be larger than, equal to or greater than the number of input dimensions. Moreover, the resultant vectors which form the matrix W are no longer are any longer constrained to be orthogonal, ICA has proved useful in a variety of settings of BCI applications such as ocular artefact removal and classification [72]–[75].

G. Advanced Techniques

Recently BCI research has gain attraction towards using advanced techniques such as Riemannian geometry, tensors and Deep learning instead of using feature vectors of EEG data.

1) Riemannian geometry: The Riemannian geometry in BCI is emerging and increasing attention due to it's accuracy, simplicity, robustness and generalization capabilities. Instead of extracting features using time domain or frequency domain from EEG signal, Riemannian geometry maps the data directly to space, where it is possible the manipulation of data for various intends such as smoothing, averaging, extrapolating and interpolating and classifying. In BCI, this mapping deals with the covariance matrix. Riemannian geometry deals with smooth curved surfaces which can be linearly and locally approximated. The curved space is called as manifold and a linear approximation at every point is called tangent space. Riemannian Classifiers either operate directly on the manifold i.e Riemannian distance to mean or the data projection to the tangent space. The Riemannian geometry has been used in EEG based BCI for event related potentials (ERPs) [81], [82], mental imagery [76]–[78] and Steady state visually evoked potential (SSVEP) [79], [80].

2) Tensors: EEG is a primary tool for brain imaging modality. In general EEG signals are represented with a vector or matrix to enable certain processing and analysis of data with methods such as spectral analysis, time-series analysis and matrix decompositions. Indeed, naturally EEG signals are with two modes i.e time and space, they can be represented by tensors (multi-way array). Though tensor tools are emerging tools for EEG signal analysis such as feature extraction, classification and clustering [83]–[87], till yet it is not very well matured and explored. The BCI applications are having problems such as overfitting due to high dimensional data and information loss for the structured data. One of the major reason behind these problems is small sample size. Tensors for BCI data can help to mitigate these problems as the tensors inherit the information regarding the structure of data which can help for the reduction of the futile features in the learning model. These tensor representations have been used in EEG based BCI for P300 [88], [89], motor imagery [90] and SSVEP [91]–[94].

3) Deep Learning: Nowadays deep learning is very popular and has attracted attention in pattern recognition systems due to its ability to detect features or latent structures from data. The fundamental principle of deep learning is the automatic extraction of features without human intervention. In deep learning feature extraction and modelling is done simultaneously where each layer train on a set of features depending on the previous layer output. As we proceed with the layers the complex features are trained which are the aggregated output of the previous layers. Convolutional neural nets (CNN) are very popular deep learning technique. It is a multilayered feedforward neural network where the weights are updated through the error backpropagation. CNN's are extensively used in several applications due to its ability to learn the most significant features. However, the performance of CNN's depend on their learning hyperparameters and architecture. Deep neural nets (DNNs) explored for EEG based BCI systems such as SSVEP [95], P300 [96], [98] and motor imagery [97], [99]–[101].

IV. Discussion

Different feature extraction techniques for EEG based BCI are surveyed in this paper. Table VI summarizes feature extraction and feature selection techniques for several dominant works in the BCI field. Time domain features consider the EEG samples over all the channels. Time domain features are easy and fast to compute and require selection of less number of parameters. In general Time domains features are used in conjunction with other feature extraction technique like spatial features or AR to have a complete performance analysis. Frequency domain features employ the mathematical technique to EEG data analysis. Frequency domain features are widely used in mental and motor imagery BCI. Nevertheless, it is also helpful in decoding emotions or for SSVEP based BCI systems. Indeed the other feature extraction techniques have been explored and used. Time-frequency feature extraction methods are most suitable for the EEG signals analysis. However, Discrete Wavelet analysis with statistical features is most adopted time-frequency feature extraction technique where wavelets are derived from the mother wavelet by dilation and translation process. CSPs are suitable for motor imagery BCI where it detects the EEG pattern by constructing spatial patterns and tuning the variance among tasks. The CSP used the multiple electrodes and performance can be affected by changing electrode positions. Finally, It is recommended to use a combination of features instead of single features extraction technique, such as time domain with frequency domain or frequency domain with connectivity features, which improves the classification accuracy of the system.

It is noteworthy that instead of using feature vectors, recent research has also explored the Riemannian geometry and Tensor approaches. The idea of Riemannian geometry evolves with the fundamental idea of directly mapping data onto a geometrical space furnished with a suitable metric. Riemannian geometry is a very popular and promising approach for various BCI problems including P300, motor imagery and SSVEP. Current research has also explored the use of tensors or covariance matrices for EEG signal classification. These tensors or covariance matrices are the linear combinations of various time points or data from various sensors. Recently the tensor approaches are emerging and promising. However, it requires more research to increase its effectiveness. DNNs are having the potential to extract features automatically and build the classification model. DNN requires large training sets, but due to the unavailability of large training sets in BCI, DNNs are the suboptimal solution for it.

A. Open Challenges

This section presents the open challenges which are important for further progress in BCI.

1) Small Training sets: The EEG based BCI applications require the training where a user has to concentrate on specific mental activity and the signals are acquired using the EEG equipment. The main challenge is the user acceptance of this training process. Due to this usability issue, training sets are small in EEG based BCI applications. Moreover, the training process is a time-consuming activity. Hence the single-trial is preferable over multi-trial which causes the small training sets.

2) High Dimension Curse: In BCI systems the EEG signals are acquired through multiple channels to preserve high spatial resolution, and thus deals with high dimensionality curse. Due to small training sets its challenging to develop BCI system without an overfitting problem. In general to describe the data properly the number of training samples needs to increase exponentially with the dimension of EEG signals [102], [103]. However, EEG based BCI systems lagging this solution due to the small training set and high dimension curse.

3) Information Transfer Learning: The primary hypothesis in machine learning is that same feature space and probability distribution are considered for modelling the train data and evaluation of test data. Unfortunately, this hypothesis is not sustained in BCI due to data is acquired from different sessions and across different subjects. To make this hypothesis true solution is Information transfer learning which is essential to develop a generalized system. Information transfer learning is the improvement in learning where the system can apply the previous tasks learned knowledge to novel tasks. It is always relevant to transfer

TABLE VI: Summary of EEG based BCI Studies

Author	EEG Pattern	Electodes	Feature Extraction	Feature Selection
Oh et al.(2014) [9]	Motor Imagery	C3, Cz, and C4	Hjoth parameters+Band pass filtering	Principal Frequency band
Vidaurre et al.(2009) [10]	Motor Imagery	C3, Cz and C4	Hjorth parameters	-
Hu et al.(2016) [11]	Attention	C3, C4, Cz, P3, P4, Pz	Hjorth parameters+AR+nonlinear features	Correlation based feature selection
Sharma et al.(2017) [104]	Epilipsy patients	32 Electrodes(International Standard (IS))	Fractal dimension	Student's t-test
Hosseinifard et al.(2013) [105]	Depression patients	Fz, Cz, Pz, Fp1,Fp2, F3, F4, F7, F8, C3, C4, T3, T4, P3, P4, T5, T6, O1,O2	detrended fluctuation analysis (DFA)+ higuchi fractal+ correlation dimension + lyapunov exponent	Genetic algorithm
Hsu(2010) [19]	Motor Imagery	64 Electrodes(IS)	Fractal dimension+DWT	-
Oikonomou et al.(2007) [21]	Epilepsy Patients	16 Electrodes(IS)	Kalman filter	-
Luke and Wouters(2016) [20]	ASSRs pattern	64 Electrodes(IS)	Kalman filter	-
Mierlo et al.(2013) [22]	Epilepsy Patients	64 Electrodes(IS)	AR+Kalman filtering	-
Mohseni et al.(2009) [23]	Epilepsy Patients	25 Electrodes(IS)	AR+Particle filter+DWT	-
Monajemi et al.(2017) [24]	P300	F3,F4,P3,P4)	Particle filtering	-
Murugappan et al.(2010) [13]	Emotions	32 Electrodes(IS)	Statistical features+Fractal dimension	-
Monajemi et al.(2017) [24]	Emotions	64 channels(IS)	DWT+Statistical features	-
da Silveira et al.(2017) [12]	Emotions	Fpz-Cz and Pz-Oz	Statistical features+Fractal dimension	Random feature selection
Kevric and Subasi(2017) [108]	Motor Imagery	118 channel(IS)	Statistical features	-
Polat and Gunes(2013) [28]	Epilipsy Patients	128 channel(IS)	PSD(Welch)	-
Wang et al.(2017) [30]	Tinnitus patients	128 channel(IS)	FFT	-
Li and Lu(2013) [25]	Emotions	62 Electrodes(IS)	FFT	CSP, SVM Wrapper
Wen and Zhang(2017) [29]	Epilipsy Patients	128 channel(IS)	FFT+nonlinear features	GA
Reuderink et al.(2013) [33]	Emotion	32 Electrodes(IS)	PSD(Welch)	Correlation analysis
Djamal et al.(2008) [31]	Motor Imagery	Fp1	FFT	-
Samiee et al.(2015) [44]	Epilipsy Patients	128 channel(IS)	DSTFT	-
Kiymik et al.(2005) [45]	Epilipsy Patients	C3-A2	STFT+CWT	-
Behnam et al.(2007) [47]	ASD patients	Fp1, Fp2, F7, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2, A2, A1	STFT+FFT	-
Fu et al.(2014) [46]	Epilipsy Patients	128 channel(IS)	HHT	-
Liang et al.(2012) [38]	Sleep EEG	F3-A2, F4-A1, C3-A2, C4-A1, P3-A2, P4-A1	multiscale entropy (MSE) and autoregressive (AR)	LDA
Zhang et al.(2015) [58]	Mental tasks	C3, C4, P3, P4, O1,O2	AR+approximate entropy AR+ WPD	-
Hatamikia et al.(2014) [40]	Emotion	32 electrodes(IS)	AR	-
Li et al.(2015) [41]	P300	16 channels(IS)	AR	Sequential forward feature selection
Zhang et al.(2017) [58]	ERP(P300)	128-channel(IS)	AR	-
Gomez-Herrero et al.(2008) [43]	relaxed wakefulness state	59 electrodes (IS)	MVAR + ICA	-
Li et al.(2017) [106]	Epilipsy patients	128-channel(IS)	DWT+Statistical features	-
Mohammadi et al.(2017) [107]	Emotion	F3-F4, F7-F8, FC1-FC2, FC5-FC6, FP1-FP2	DWT + Statistical features	-
Sereshkeh et al.(2017) [14]	Imagined Speech	64 Electrodes(IS)	DWT+Statistical features	-
Khlif et al.(2013) [110]	Newborn Seizure	128-channel(IS)	MF	-
Lafuente et al.(2017) [111]	Imagined Speech	64 Electrodes(IS)	MF	-
Samek et al.(2012) [54]	Motor Imagery	118-channel(IS)	CSP	-
Kevric and Subasi(2017) [108]	Motor Imagery	118 Electrodes(IS)	DWT+Statistical features	-
Falzon et al.(2012) [55]	SSVEP	27-Electrodes(IS)	CSP	-
Sannelli et al.(2010) [56]	Motor Imagery	119 channels(IS)	CSP	-
Yue et al.(2012) [57]	Motor Imagery	FC1, FC2, FC5, FC6, Cz, C1, C2,C3, C4, CP1, CP2, CP5, CP6, P3, P4	CSP	-

the knowledge between the same tasks (P300, P300) instead of different tasks (P300, motor imagery).

4) Nonstationary: The non-stationary nature of EEG signals is the major obstacle in developing BCI systems. The continuous variations are occurred over time either within or between recording sessions. Moreover, concentration and fatigue are also considered as intrinsic nonstationary aspects. The undesirables signals are also included due to environmental noise, alterations in electrode placement and environmental noise as well as various artefacts like Electromyography (EMG), Electrooculography (EOG) and ocular artefacts.

5) Universality: The expertise and knowledge from various disciplines are required for successful BCI development. The electrical engineering, physics, medicine and biology are required for the accurate EEG measurement, electrode placement and the selection of proper input signals. The computer science engineering and mathematics are required to build models which include preprocessing and designing of neural nets. As per the extended output requirement, the other disciplines like telecommunications or mechanical engineering might be required. Indeed for the efficient and effective BCI development required the team of experts from various disciplines.

B. BCI Applications

This section begins by clarifying the distinction between BCIs as a vital preamble to address the potential practical applications using BCI technology. Present BCIs classified as exogenous or endogenous according to the nature of the recorded signal. Exogenous BCI systems depend on the neuronal activity evoked by an external stimulus such as auditory or visual evoked potentials. Exogenous BCIs may not require intensive training since it is easy to set up their control signals (P300 and SSVEPs). In contrast, the endogenous BCI system depends on brain rhythms and other potentials instead of external stimuli. Endogenous BCIs require extensive training in which a user learns the skill of producing a specific pattern which is decoded by the system.

BCI tools have potential applications in the following different areas:

1) Communication and control: Human beings communicate with each other by means of verbal and visual expressions. However, paralysed patients having various neurological diseases like brainstem infarcts, brain injury, stroke and advanced amyotrophic lateral sclerosis (ALS), their conditions prohibit normal communication, contrarily affecting their life. Due to the security issue, in a few situations, it would be fascinating to interact using brain signals. In this context, the brain-computer interface (BCI) is promising to use as a communication technology via a variety of methods such as in spelling applications [112], semantic categorization [113], or silent speech communication [114].

The hands-free applications like BCI aided mind controlling machines can bring comfort and ease to human beings. These applications can accomplish a set of commands using brain signals and thus don't require any muscle movement. Disabled users can get support from BCI assistive robots in professional and daily life and can build their life better [115].

2) Neuroergonomics and smart environment: BCIs are also exploited in building smart environments such as workplaces, houses or transportation with further offering such as luxury, safety and physiological control to humans daily life. However, they have expected the cooperation between BCI technologies and the Internet Of Things (IoT) [116], [117]. Initially, BCI applications were focused on the aim to yield a communication channel for disabled users who have speaking or mobility issues. But later on, BCI applications have shown remarkable footprints in the world of healthy people. It acts as a physiological measuring tool about a users cognitive, affective or emotional state [118]–[120].

3) Neuromarketing and advertisement: The BCI researchers have also shown interest in the marketing field. The EEG evaluation benefits for the TV advertisements related to political and commercial fields are explained in [121]. The research has been also carried to measure the attention of accompanying watching activity [122] and the estimation of the memorization of TV advertisements [123].

4) Educational and self-regulation: Neurofeedback is an encouraging approach for improvement in brain performance by analyzing brain waves. The educational systems can determine the degree of clearness of studied information via utilizing brain electrical signals. The personalized interaction took pace with each learner according to the response [124]. Self-regulation through learning using noninvasive BCI has been also studied [125], [126].

5) Games and entertainment: BCIs are not only popular in the medical field but also in the entertainment and gaming applications. Researchers have shown interest in providing multi-brain entertainment experience by combining the existing games features with brain controlling capabilities [127]. The players can join a competitive or collaborative football game with two BCIs, whereby right or left-hand movements players can score goals. On the other hand, few games have been developed for neuroprosthetic rehabilitation and emotional control. The brainball game is created with the intent to reduce stress lever; where users can move the ball with the relaxation [128].

6) Security and authentication: Electrophysiology or cognitive biometrics are used as the detection of suspicious objects and abnormal behavior [129], [130]. Electrophysiology signals improve the resistance of biometric systems to spoofing attacks since they are difficult to synthesize and can not be casually acquired by external observers. Researchers have also worked on the EEG signal based authentication system as part of the smart driving system using driving behavior or mental-task conditions [131], [132].

V. Conclusions

This paper has reviewed various feature extraction techniques employed in EEG based BCI applications. As EEG changes over time, the adoption of fast and reliable feature extractors is very important in BCI research. We summarize the time-domain, frequency-domain and time-frequency domain features with their properties. The findings suggest that each technique has its specific benefits and flaws; hence the optimum feature extraction method for any application might be different. Currently rather than using feature extraction techniques, recent research has been also explored around tensors, Riemannian Geometry and deep learning. Tensors and Riemannian geometry are very helpful to improve BCI reliability. Deep learning networks don't appear to be effective for BCI applications; one possible reason is limited training data. Open challenges have been also discussed; which we think are important for further development in BCI. In the future, the work related to feature extraction methods for EEG based BCI should focus on developing more robust and efficient feature extraction techniques which are theoretically motivated, biologically realistic and useful for real-time BCI applications.

REFERENCES

[1] G. Pfurtscheller, G. R. Muller-Putz, R. Scherer and C. Neuper, "Rehabilitation with brain-computer interface systems," *J. Computer*, vol. 41, no. 10, pp. 58-65, Oct 2008.

[2] J. R. Wolpaw, D. J. McFarland, G. W. Neat and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," *J. Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 3, pp. 252259, Mar 1991.

[3] D. Coyle, J. Principe, F. Lotte and A. Nijholt, "A. Guest Editorial: Brain/neuronal-Computer game interfaces and interaction," *IEEE Trans. Comput. Intell. AI Games*, vol. 5, no. 2, pp. 77-81, Jun 2013.

[4] K. K. Ang and C. Guan, "Guest Editorial: Brain/neuronal-Computer game interfaces and interaction," *J. Comput. Sci. Eng.*, vol. 7, pp. 139-146, 2013.

[5] O. G. Selfridge, "Pattern recognition and modern computers," *Western Joint Computer Conference*, pp. 91-93, Mar 1955.

[6] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, p.R1, Jan 2007.

[7] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, p.031005, Apr 2018.

[8] M. Bueno-Lopez, P. A. Munoz-Gutierrez, E. Giraldo and M. Molinas, "Electroencephalographic Source Localization based on Enhanced Empirical Mode Decomposition," *IAENG International Journal of Computer Science*, vol. 46, no. 2, pp. 228-236, 2019.

[9] S. H. Oh, Y. R. Lee and H. N. Kim, "A novel EEG feature extraction method using Hjorth parameter," *International Journal of Electronics and Electrical Engineering*, vol. 2, no. 2, pp. 106-110, Jun 2014.

[10] C. Vidaurre, N. Kramer, B. Blankertz and A. Schlogl "Time domain parameters as a feature for EEG-based brain-computer interfaces," *Neural Networks*, vol. 22, no. 9, pp. 1313-1319, Nov 2009.

[11] B. Hu,X. Li, S. Sun and M. Ratcliffe, "Attention recognition in EEG based affective learning research using CFS+ KNN algorithm," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 1, pp. 38-45, Oct 2016.

[12] T. L. da Silveira, A. J. Kozakevicius and C. R. Rodrigues, "Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain," *Medical & Biological Engineering & Computing*, vol. 55, no. 2, pp. 343-352, Feb 2017.

[13] M. Murugappan, N. Ramachandran and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," *Journal of Biomedical Science and Engineering*, vol. 3, no. 4, p.390, Apr 2010.

[14] A. R. Sereshkeh, R. Trott, A. Bricout and T. Chau "EEG classification of covert speech using regularized neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 12, pp. 2292-2300, Nov 2017.

[15] R. Khosrowabadi and A. Rahman, A "Classification of EEG correlates on emotion using features from Gaussian mixtures of EEG spectrogram," *International Conference on Information and Communication Technology for the Moslem World*, pp. E102-E107, Dec 2010.

[16] K. Ansari-Asl, G. Chanel and T. Pun, "A channel selection method for EEG classification in emotion assessment based on synchronization likelihood," *in Proc.* 15^{th} *Eur. Signal Process. Conf*, pp. 1241-1245, Sep 2007.

[17] Y. Liu and O. Sourina, "Real-time fractal-based valence level recognition from EEG," *Trans. Comput. Sci. XVIII*, pp. 101-120, 2013.

[18] O. Sourina and Y. A. Liu, "A fractal-based algorithm of emotion recognition from EEG using arousal-valence model," *in Proc. Biosignals*, pp. 209-214, Jan 2011 .

[19] W. Y. Hsu, "EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features.," *Journal of Neuroscience Methods*, vol. 189, no. 2, pp. 295-302, Jun 2010.

[20] R. Luke, J. Wouters, "Kalman filter based estimation of auditory steady state response parameters," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 196-204, Apr 2016.

[21] V. P. Oikonomou, A. T. Tzallas and D. I. Fotiadis, "A Kalman filter based methodology for EEG spike enhancement. Computer methods and programs in biomedicine," *Computer Methods and Programs in Biomedicine*, vol. 85, no. 2, pp. 101-108, Feb 2007.

[22] P. Van Mierlo, E. Carrette, H. Hallez, R. Raedt, A. Meurs, S. Vandenberghe, D. Van Roost, P. Boon, S. Staelens and K. Vonck, "Guest Editorial: Ictal-onset localization through connectivity analysis of intracranial EEG signals in patients with refractory epilepsy," *Epilepsia*, vol. 54, no. 8, pp. 1409-1418, Aug 2013.

[23] H. R. Mohseni, K. Nazarpour, E. L. Wilding and S. Sanei, "The application of particle filters in single trial event-related potential estimation," *Physiological Measurement*, vol. 30, no. 10, p.1101, Sep 2009.

[24] S. Monajemi, D. Jarchi, S. H. Ong and S. Sanei, "Cooperative Particle Filtering for Tracking ERP Subcomponents from Multichannel EEG," *Entropy*, vol. 19, no. 5, p.199, May 2017.

[25] M. Li and B. L. Lu, "Emotion classification based on gamma band EEG," *IEEE Int. Conf. Eng. Med. Biol. Soc.*, pp. 1323-1326, Sep 2009.

[26] Z. Khalili and M. H. Moradi, "Emotion detection using brain and peripheral signals," *IEEE Int. Biomedical Eng. Conf.*, pp. 1-4, Dec 2008.

[27] S. Makeig, G. Leslie, T. Mullen and D. Sarma, Bigdely-Shamlo, N. and Kothe, C., "First demonstration of a musical emotion BCI.," *in Proc. Int. Conf. Affect. Comput. Intell. Interaction*, pp. 487-496, Oct 2011.

[28] K. Polat and S. Gune, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1017-1026, Apr 2007.

[29] T. Wen and Z. Zhang, "Effective and extensible feature extraction method using genetic algorithm based frequency domain feature search for epileptic EEG multiclassification," *Medicine*, vol. 96, no. 19, May 2017.

[30] S. J. Wang, Y. X. Cai, Z. R. Sun, C. D.Wang and Y. Q. Zheng, "Tinnitus EEG classification based on multi frequency bands," *In International Conference on Neural Information Processing*, pp. 788-797, Nov 2017.

[31] E. C. Djamal, M. Y. Abdullah and F. Renaldi, "Brain Computer Interface Game Controlling Using Fast Fourier Transform and Learning Vector Quantization. ," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, pp. 71-74, Jun 2017.

[32] E. Kroupi, A. Yazdani and T. Ebrahimi, "EEG correlates of different emotional states elicited during watching music videos," *in Proc. Int. Conf. Affect. Comput. Intell. Interaction.*, pp. 457-466, Oct 2011.

[33] B. Reuderink, C. Muhl and M. Poel, "Valence, arousal and dominance in the EEG during game play, " *Int. J. Autonom. Adaptive Commun. Syst.*, vol. 6, no. 1, pp. 45-62, Jan 2013.

[34] M. Grosse-Wentrup, "Understanding brain connectivity patterns during motor imagery for brain computer interfacing," *Advances in Neural Information Processing Systems*, pp. 561-568, 2009 .

[35] T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, G. Cauwenberghs and T. P. Jung, "Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG," *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 2184-2187, Jul 2013.

[36] Q. Wei, Y. Wang, X. Gao and S. Gao, "Amplitude and phase coupling measures for feature extraction in an EEG-based braincomputer interface," *Journal of Neural Engineering*, vol. 4, no. 2, p.120, Mar 2007.

[37] H. Zhang, R. Chavarriaga and J. D. Millan, "Discriminant brain connectivity patterns of performance monitoring at average and single-trial levels NeuroImage," *NeuroImage*, vol. 120, pp. 64-74, Oct 2015.

[38] S. F. Liang, C. E. Kuo, Y. H. Hu, Y. H. Pan and Y. H. Wang, "Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models," *IEEE Trans. Instrumentation and Measurement*, vol. 61, no. 6, pp. 1649-1657, Mar 2012.

[39] Y. Zhang, B. Liu, X. Ji and D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Processing Letters*, vol. 45, no. 2, pp. 365-378, Apr 2017.

[40] S. Hatamikia, K. Maghooli and A. M. Nasrabadi, "The emotion recognition system based on autoregressive model and sequential

forward feature selection of electroencephalogram signals.," *Journal of Medical Signals and Sensors*, vol. 4, no. 2, p.194, Jul 2014.
[41] P. Li, X. Wang, F. Li, R. Zhang, T. Ma,Y. Peng, X. Lei, Y. Tian, D. Guo, T. Liu and D. Yao,"Autoregressive model in the Lp norm space for EEG analysis," *Journal of Neuroscience Methods*, vol. 240, pp. 170-178, Jan 2015.
[42] Z. G. Zhang, Y. S. Hung and S. C. Chan, "Local polynomial modeling of time-varying autoregressive models with application to time-frequency analysis of event-related EEG," *IEEE Trans. on Biomedical Engineering*, vol. 58, no. 3, pp. 557-566, Oct 2010.
[43] G. Gomez-Herrero, M. Atienza, K. Egiazarian and J. L. Cantero, "Measuring directional coupling between EEG sources," *Neuroimage*, vol. 43, no. 3, pp. 497-508, Nov 2008.
[44] K. Samiee, P. Kovacs and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform," *IEEE Trans. on Biomedical Engineering*, vol. 62, no. 2, pp. 541-552, Sep 2014.
[45] M. K. Kiymik, I. Guler, A. Dizibuyuk and M. Akin, "Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application," *Computers in Biology and Medicine*, vol. 35, no. 7, pp. 603-616, Oct 2005.
[46] K. Fu, J. Qu, Y. Chai and Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM," *Biomedical Signal Processing and Control*, vol. 13, pp. 15-22, Sep 2014.
[47] H. Behnam, A. Sheikhani, M. R. Mohammadi, M. Noroozian and P. Golabi, "Analyses of EEG background activity in Autism disorders with fast Fourier transform and short time Fourier measure," *Int. Conf. on Intelligent and Advanced Systems*, pp. 1240-1244, Nov 2007.
[48] S. K. Hadjidimitriou and L. J. Hadjileontiadis, "Toward an EEG-based recognition of music liking using time-frequency analysis," *IEEE Trans. on Biomedical Engineering*, vol. 59, no. 12, pp. 3498-3510, Sep 2012.
[49] O. Faust, U. R. Acharya, H. Adeli and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56-64, Mar 2015.
[50] M. Grosse-Wentrup and M. Buss, "Multiclass common spatial patterns and information theoretic feature extraction.," *IEEE Trans. Biomed. Eng*, vol. 55, no. 8, pp. 1991-2000, Mar 2008.
[51] E. A. Mousavi, J. J. Maller, P. B. Fitzgerald and B. J. Lithgow, "Wavelet common spatial pattern in asynchronous offline brain computer interfaces," *Biomedical Signal Processing and Control*, vol. 6, no. 2, pp. 121-128, Apr 2011.
[52] S. Lemm, B. Blankertz, G. Curio and K. R. Muller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541-1548, Aug 2005.
[53] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch,G. Curio and K. R. Muller,"Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2274-2281, Oct 2006.
[54] W. Samek, C. Vidaurre, K. R. Muller and M. Kawanabe, "Stationary common spatial patterns for braincomputer interfacing," *Journal of Neural Engineering*, vol. 9, no. 2, p.026013, Feb 2012.
[55] O. Falzon, K. P. Camilleri and J. Muscat, "The analytic common spatial patterns method for EEG-based BCI data," *Journal of Neural Engineering*, vol. 9, no. 4, p.045009, July 2012.
[56] C. Sannelli, T. Dickhaus, S. Halder, E. M. Hammer, K. R. Muller and B. Blankert, "On optimal channel configurations for SMR-based braincomputer interfaces," *Brain Topography*, vol. 23, no. 2, pp. 186-193, Jun 2010.
[57] J. Yue, Z. Zhou, J. Jiang, Y. Liu and D. Hu "Balancing a simulated inverted pendulum through motor imagery: an EEG-based real-time control paradigm," *Neuroscience letters*, vol. 524, no. 2, pp. 95-100, Aug 2012.
[58] Y. Zhang, G. Zhou, J. Jin, X. Wang and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface," *Journal of Neuroscience methods*, vol. 255, pp. 85-91, Nov 2015.
[59] I. Koprinska, "Feature selection for brain-computer interfaces," *Pacific-Asia conference on knowledge discovery and data mining*, pp. 106-117, 2009.
[60] N. Liang and L. Bougrain, "Decoding finger flexion from band-specific ECoG signals," *Humans Front. Neurosci*, vol. 6, p.19, Jun 2012.
[61] T. N. Lal, M. Schroder, M., T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer and B. Scholkopf, "Support vector channel selection in BCI," *IEEE Transactions on Biomedical engineering*, vol. 51, no. 6, pp. 1003-1010, May 2004.
[62] B. Dal Seno, M. Matteucci and L. Mainardi, "A genetic algorithm for automatic feature extraction in P300 detection," *IEEE Int. Joint Conf. on Neural Networks*, pp. 3145-3152, Jun 2008.
[63] J. Ortega, J. Asensio-Cubero, J. Q. Gan and A. Ortiz, "Classification of motor imagery tasks for BCI with multiresolution analysis and multiobjective feature selection," *Int. J. Biomedical engineering*, vol. 15, no. 1, p.73, July 2016.
[64] R. Corralejo, R. Hornero and D. Alvarez, "Feature selection using a genetic algorithm in a motor imagery-based brain computer interface.," *Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 7703-7706, Aug 2011.
[65] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan and J. R. Wolpaw, "A comparison of classification techniques for the P300 speller ," *J. Neural Eng*, vol. 3, no. 4, pp. 299-305, Oct 2006.
[66] C. J. Lin and M. H. Hsieh, "Classification of mental task from EEG data using neural networks based on particle swarm optimization," *Neurocomputing*, vol. 72, pp. 1121-1130, Jan 2009.
[67] C. J. Tu, L. Y. Chuang, J. Y. Chang and C. H. Yang, "Feature Selection using PSO-SVM," *IAENG International Journal of Computer Science*, vol. 33, no. 1, pp. 111-116, 2007.
[68] O. G. Lins, T. W. Picton, P. Berg and M. Scherg, "Guest Editorial: Ocular artifacts in recording EEGs and event-related potentials II: Source dipoles and source components," *Brain Topography*, vol. 6, no. 1, pp. 65-78, Sep 1993.
[69] A. T. Boye, U. Q. Kristiansen, M. Billinger, O. F. do Nascimento and D. Farina, "Identification of movement-related cortical potentials with optimized spatial filtering and principal component analysis," *Biomed. Signal Process. Control*, vol. 3, no. 4, pp. 300-304, Oct 2008.
[70] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129-1159, Nov 1995.
[71] T. W. Lee, M. Girolami and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural Computation*, vol. 11, no. 2, pp. 417-441, Feb 1999.
[72] J. F. Gao, Y. Yang, P. Lin, P. Wang and C. X. Zheng, "Automatic removal of eye-movement and blink artifacts from EEG signals," *Brain Topography*, vol. 23, no. 1, pp. 105-114, Mar 2010.
[73] A. Flexer, H. Bauer, J. Pripfl and G. Dorffner, "Using ICA for removal of ocular artifacts in EEG recorded from blind subjects," *Neural Networks*, vol. 18, no. 7, pp. 998-1005, Sep 2005.
[74] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne and T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745-1758, Oct 2000.
[75] S. Chiappa and D. Barber, "EEG classification using generative independent component analysis," *Neurocomputing*, vol. 69, pp. 769-777, Mar 2006.
[76] A. Barachant, S. Bonnet, M. Congedo and C. Jutten, "Multi-class brain computer interface classification by riemannian geometry.," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920-928, Oct 2011.
[77] F. Yger, "A review of kernels on covariance matrices for BCI applications," *IEEE Int. Workshop on Machine Learning for Signal Processing*, pp. 1-6, Sep 2013.
[78] A. Barachant, S. Bonnet, M. Congedo and C. Jutten, "Classification of covariance matrices using a Riemannian-based kernel for BCI applications," *Neurocomputing*, vol. 112, pp. 172-178, July 2013.
[79] S. Chevallier, E. Kalunga, Q. B. Elemy and F. Yger, "Riemannian classification for SSVEP-based BCI Brain Computer Interfaces," *Handbook: Technological and Theoretical Advances*, 2018.
[80] E. K. Kalunga, S. Chevallier, Q. Barthelemy, K. Djouani, E. Monacelli and Y. Hamam, "Online SSVEP-based BCI using Riemannian geometry," *Neurocomputing*, vol. 191, no. 2, pp. 55-68, May 2016.
[81] A. Barachant and M. Congedo, "A plug and play P300 BCI using information geometry," *arXiv preprint arXiv:1409.0107*, Aug 2014.
[82] L. Mayaud, S. Cabanilles, A. Van Langhenhove, M. Congedo, A. Barachant, S. Pouplin, S. Filipe, L. Petegnief, O. Rochecouste, E. Azabou and C. Hugero, "Brain-computer interface for the communication of acute patients: a feasibility study and a randomized controlled trial comparing performance with healthy participants and a traditional assistive device," *Brain-Computer Interfaces*, vol. 3, no. 4, pp. 197-215, Oct 2016.
[83] A. Cichocki, "Tensor decompositions: a new conceptin brain data analysis? ," *J. Soc. Instrum. Control Eng.*, vol. 58, pp. 507-516, 2011.
[84] A. Cichocki, N. Lee, I. Oseledets, A. H. Phan, Q. Zhao and D. P. Mandic, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions,"

Foundations and Trends in Machine Learning, vol. 9, pp. 249-429, 2016.
[85] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145-163, 2015.
[86] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, A. H. Phan, S. Choi, H. Lee, Q. Zhao, L. Zhang and Y. Li, "Noninvasive BCIs: Multiway signal-processing array decompositions," *Computer*, vol. 41, no. 10, pp. 34-42, Oct 2008 .
[87] A. Cichocki, R. Zdunek, A. H. Phan and S. I. Amari, "Nonnegative matrix and tensor factorizations: applications to exploratory multiway data analysis and blind source separation," *John Wiley & Sons*, July 2009.
[88] A. Onishi, A. H. Phan, K. Matsuoka and A. Cichocki, "Tensor classification for P300-based brain computer interface," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 581-584, Mar 2012.
[89] Y. Washizawa, H. Higashi, T. Rutkowski, T. Tanaka and A. Cichocki, "Tensor based simultaneous feature extraction and sample weighting for EEG classification," *Int. Conf. on Neural Information Processing*, pp. 26-33, Nov 2010.
[90] A. H. Phan and A. Cichocki, "Guest Editorial: Tensor decompositions for feature extraction and classification of high dimensional datasets," *Nonlinear Theory and its Applications*, vol. 1, pp. 37-68, 2010.
[91] Y. U. Zhang, G. Zhou, J. Jin, X. Wang and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *International Journal of Neural Systems*, vol. 4, no. 4, p.1450013, Jun 2014.
[92] Y. Zhang, G. Zhou, J. Jin, X. Wang and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface," *Journal of Neuroscience Methods*, vol. 255, pp. 85-91, Nov 2015.
[93] Y. Zhang, G. Zhou, J. Jin, X. Wang and A. Cichocki, "Sparse Bayesian multiway canonical correlation analysis for EEG pattern recognition," *Neurocomputing*, vol. 225, pp. 103-110, Feb 2017.
[94] Y. Zhang, G. Zhou, J. Jin, X. Wang and A. Cichocki, "Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs," *Int. Conf. on Neural Information Processing*, pp. 287-295, Nov 2011.
[95] N. S. Kwak, K. R. Muller and S. W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PloS one*, vol. 12, no. 2, p.e0172578, Feb 2017.
[96] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain-computer interfaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 433-445, jun 2010.
[97] N. Lu,T. Li, X. Ren, H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil.Eng.*, vol. 25, no. 6, pp. 566-576, Aug 2016.
[98] R. Manor and A. B. Geva, "Convolutional neural network for multi-category rapid serial visual presentation BCI," *Frontiers in Computational Neuroscience*, vol. 9, p.146, Dec 2015.
[99] R. T. Schirrmeister, L. Gemein, K. Eggensperger, F. Hutter and T. Ball, "Deep learning with convolutional neural networks for decoding and visualization of EEG pathology," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391-5420, Nov 2017.
[100] I. Sturm, S. Lapuschkin, W. Samek and K. R. Muller, "Interpretable deep neural networks for single-trial EEG classification," *Journal of Neuroscience Methods*, vol. 274, pp. 141-145, Dec 2016.
[101] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, vol. 14, no. 1, p.016003, Nov 2016.
[102] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners.," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 1, no. 3, pp. 252-264, Mar 1991.
[103] A. K. Jain and B. Chandrasekaran, "Dimensionality and sample size considerations in pattern recognition practice," *Handbook of Statistics*, vol. 2, pp. 835-855, Jan 1982.
[104] M. Sharma, R. B. Pachori and U. R. Acharya, "A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension," *Pattern Recognition Letters*, vol. 94, pp. 172-179, July 2017.
[105] B. Hosseinifard, M. H. Moradi and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal," *Computer Methods and Programs in Biomedicine*, vol. 109, no. 3, pp. 339-345, Mar 2013.
[106] M. Li, W. Chen and T. Zhang, "Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble," *Biomedical Signal Processing and Control*, vol. 31, pp. 357-365, Jan 2017.
[107] Z. Mohammadi, J. Frounchi and M. Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Computing and Applications*, vol. 28, no. 8, pp. 1985-1990, Aug 2017 .
[108] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomedical Signal Processing and Control*, vol. 31, pp. 398-406, Jan 2017.
[109] D. Pawar and S. N. Dhage, "Recognition of unvoiced human utterances using brain-computer interface," *Int. Conf. on Image Information Processing (ICIIP)*, pp. 1-4, Dec 2017.
[110] M. S. Khlif, P. B. Colditz and B. Boashash, "Effective implementation of time-frequency matched filter with adapted pre and postprocessing for data-dependent detection of newborn seizures," *Medical Engineering and Physics*, vol. 35, no. 12, pp. 1762-1769, Dec 2013.
[111] V. Lafuente, J. M. Gorriz, J. Ramirez and E. Gonzalez, "P300 brainwave extraction from EEG signals: An unsupervised approach," *Expert Systems with Applications*, vol. 74, pp. 1-10, May 2017.
[112] Y. Lelievre, Y. Washizawa and T. M. Rutkowski, "Single trial BCI classification accuracy improvement for the novel virtual sound movement-based spatial auditory paradigm.," *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 1-6, 2013.
[113] W. Wang, A. D. Degenhart, G. P. Sudre, D. A. Pomerleau and E. C. Tyler-Kabara, "Decoding semantic information from human electrocorticographic (ECoG) signals," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6294-6298, 2011.
[114] D. Pawar and S. N. Dhage, " Multiclass covert speech classification using extreme learning machine," *Biomedical Engineering Letters*, pp. 1-10, Mar 2020.
[115] N. Prataksita, Y. T. Lin, H. C. Chou and C. H. Kuo, "Brain-robot control interface: Development and application," *IEEE International Symposium on Bioelectronics and Bioinformatics*, pp. 1-4, apr 2014.
[116] M. C. Domingo, "An overview of the Internet of Things for people with disabilities," *Journal of Network and Computer Applications*, vol. 35, no. 2, pp. 584-596, 2012.
[117] C. T. Lin, B. S. Lin, F. C. Lin and C. J. Chang, "Brain computer interface-based smart living environmental auto-adjustment control system in UPnP home networking," *IEEE Systems Journal*, vol. 8, no. 2, pp. 363-370, Apr 2012.
[118] R. N. Roy, S. Bonnet, S. Charbonnier and A. Campagne, "Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI," *International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6607-6610, July 2013.
[119] D. G. Duru, A. D. Duru, D. E. Barkana,O. Sanli and M. Ozkan, "Assessment of surgeon's stress level and alertness using EEG during laparoscopic simple nephrectomy," *International IEEE/EMBS Conference on Neural Engineering*, pp. 452-455, Nov 2013.
[120] Y. Dong, Z. Hu, K. Uchimura and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 596-614, Dec 2010.
[121] G. Vecchiato, L. Astolfi, F. D. V. Fallani, S. Salinari, F. Cincotti, F. Aloise, D. Mattia, M. G. Marciani, L. Bianchi,R. Soranzo and F. Babiloni, "The study of brain activity during the observation of commercial advertsing by using high resolution EEG techniques," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 57-60, Sep 2009.
[122] M. Yoshioka, T. Inoue and J. Ozawa, "Brain signal pattern of engrossed subjects using near infrared spectroscopy (NIRS) and its application to TV commercial evaluation," *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 1-6, Jun 2012.
[123] G. Vecchiato, F. Babiloni,L. Astolfi, J. Toppi, P. Cherubino, J. Dai, W. Kong and D. Wei, "Enhance of theta EEG spectral activity related to the memorization of commercial advertisings in Chinese and Italian subjects," *IEEE International Conference on Biomedical Engineering and Informatics*, vol. 3, pp. 1491-1494, Oct 2011.
[124] K. A. Sorudeykin, "An educative brain-computer interface," arXiv preprint arXiv:1003.2660, Mar 2010.
[125] S. J. Johnston, S. G. Boehm, D. Healy, R. Goebel and D. E. Linden, "Neurofeedback: a promising tool for the self-regulation of emotion networks," *Neuroimage*, vol. 49, no. 1, pp. 1066-1072, Jan 2010.
[126] V. Zotev, R. Phillips, H. Yuan, M. Misaki and J. Bodurka, "Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback," *NeuroImage*, vol. 85, pp. 985-995, Jan 2014.

[127] L. Bonnet, F. Lotte and A. Lecuyer, "Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 185-198, Jan 2013.
[128] D. S. Tan and A. Nijholt, "Brain-computer interfaces," Springer-Verlag London Limited, pp. 3-19, 2010.
[129] D. T. Karthikeyan and B. Sabarigiri, "Enhancement of multi-modal biometric authentication based on iris and brain neuro image coding," *International Journal of Biometrics and Bioinformatics*, vol. 5, no. 5, pp. 249-256, 2011.
[130] I. Svogor and T. Kisasondi, "Two factor authentication using EEG augmented passwords," *IEEE International Conference on Information Technology Interfaces*, pp. 373-378, Jun 2012.
[131] I. Nakanishi, S. Baba, K. Ozaki and S. Li, "Using brain waves as transparent biometrics for on-demand driver authentication," *IJBM*, vol. 5, pp. 288305, Jan 2013.
[132] I. Nakanishi, K. Ozaki and S. Li, "Evaluation of the brain wave as biometrics in a simulated driving environment.," *BIOSIG-Proceedings of the International Conference of Biometrics Special Interest Group*, pp. 1-5, Sep 2012.

Dipti Pawar received her M.E degree from Vishwakarma Institute of Technology, Pune, India. She is currently a PhD research scholar at the Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, India. Her research interests include pattern recognition, machine learning and brain-computer interface. She has published various research papers at national and international journals, and conference proceedings.

Sudhir Dhage received his PhD in Computer Engineering from the Veermata Jijabai Technological Institute (VJTI), Mumbai, India. He is currently working as a Professor in the Computer Engineering Department at Sardar Patel Institute of Technology, Mumbai, India. His area of interest is computational science, machine intelligence, parallel and distributed systems, multimedia systems. He has 25 years of experience in academic and research. In his credit 02 patents, 41 publications in refereed journals as well as international and national conferences proceedings.