Robotica (2022), **40**, pp. 77–101
doi:10.1017/S0263574721000382

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Recent Approaches on Classification and Feature Extraction of EEG Signal: A Review

Pooja, SK Pahuja and Karan Veer*

Department of Instrumentation & Control Engineering, DR B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India
*Corresponding author. Email: veerk@nitj.ac.in

Received: 7 April 2020; **Revised:** 2 February 2021; **Accepted:** 20 March 2021; **First published online:** 4 May 2021

Keywords: Electroencephalography, Brain waves, Feature extraction, TDFs, FDFs, TFDFs, Classification

Abstract

Objective: Electroencephalography (EEG) has an influential role in neuroscience and commercial applications. Most of the tools available for EEG signal analysis use machine learning to extract the required information. So, the study of robust techniques for feature extraction and classification is an important thing to understand the practical use of EEG. The paper aims that if there is any special tool for a particular task. Which feature domain or classifier has a significant role in EEG signal analysis?
Approach: It presents a detailed report of the current trend for bio-electrical signals classification focusing on various classifiers' advantages and disadvantages. This study includes literature from 2000 to 2021 with a brief description of EEG signal origin and advancement in classification techniques.
Results: Randomly used classifiers for EEG signal can be categorized into five classes, namely Linear Classifiers, Nearest Neighbor Classifiers, Nonlinear Bayesian Classifiers, Neural Networks, and Combinations of Classifiers. Approximately 40% of studies use Support Vector Machine, Nearest Neighbor, and their combination with others. For specific tasks, particular classifiers are recommended in the survey. Features can be defined into four categories, namely TDFs, FDFs, TFDFs, and statistical features, where 39% of studies used TFDFs. Multi-domains features are preferred when the required information cannot be obtained from one domain.
Significance: The paper summarizes the recent approaches for feature extraction and classification of EEG signals. It describes the brain waves with their classification, related behavior, and task with the physiological correlation. The comparative analysis of different classifiers, toolbox, the channel used, accuracy, and the number of subjects from various studies can help the practitioners choose a suitable classifier. Furthermore, future directions can cope up with the relevant problems and can lead to accurate classification.

1. Introduction

The biological signals are measured and monitored from the different parts of human body. A few of them are Electromyography (EMG) (for muscular contraction), Electrocardiography (ECG) (for heart waves), Electrooculography (EOG) (to record eye dipole field), Electroencephalography (EEG) (for brain waves), and Electrogastrography (EGG) (to record muscular activity of stomach). With it, various imaging techniques are also available as Single-Photon Emission Computed Tomography (SPECT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Computed Tomography (CT), and Functional Magnetic Resonance Imaging (fMRI). But the EEG signals play a significant role in the biomedical field because it directly measures the electrical activity of the brain. [1] Mostly cited and surveyed articles regarding feature extraction and EEG signal classification are briefly described in Table I, which focuses on the significance of EEG signal.

EEG is an imaging technique that scans electrical activity of brain. If EEG signal is measured from the cortical surface (exposed surface) of the brain called as electrocorticogram, when it is recorded

Table I. Some important papers on EEG feature extraction and classification with citation.

Author/Year	Key focus	No. of citation
A S Al-Fahoum et al./2014. [2]	Conventional techniques for EEG signal feature extraction	96
Y L Chong et al./2015. [3]	An improved version of quantum support vector machine with an unknown kernel, which predicts the label of EEG signal	39
A Al-Nafjan et al./2017. [4]	To review the studies on emotion detection and recognition and classification	31
F Lotte et al./2018. [5]	Various EEG signal classification approaches used in BCIs are described.	289
M Hamada et al./2018. [6]	Systematic review for emotion classification, feature extraction, comparison of groups using EEG signal of the human brain	9
A Craik et al./2019 . [7]	Which classification task and input formulations are performed with deep learning? Which specific deep learning network structures are suitable for a particular task?	47
Q Gao et al./2020. [8]	Three emotions (neutral, happiness, and sadness) classification using a fused feature extraction method	32

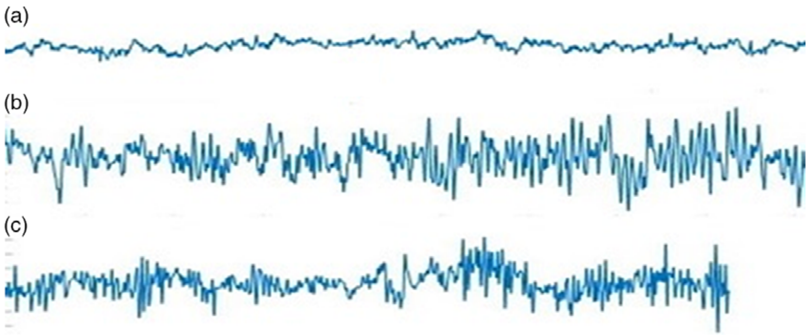

Figure 1. *Distinctive waves of EEG signal. (a) Movement/motor imagery (eye open/close)dataset, (b) mental arithmetic task, (c) rapid serial visual presentation [9, 10, 11].*

from scalp only called as electroencephalogram. Nowadays, it is a noninvasive procedure that can be recorded repeatedly from patients, children, and adults without any risk or limitations. Figure 1 shows some raw waveforms of EEG signals of three activities (motor imagery task, mental arithmetic task, and rapid serial visual presentation) of the human brain. Figure 2 shows the acquisition and analysis of EEG signals.

This survey is systematically performed for the existing literature to collect information from prestigious journals and conferences. Applying the related search strings gives about 1200 relevant articles, out of which 96 are identified as the relevant articles after excluding the duplicate and irrelevant ones, which represents the classification and feature extraction of brain waves. For the brief introduction of EEG signal and its progression with time, few papers are included from 1924 to 2000. For the statistical

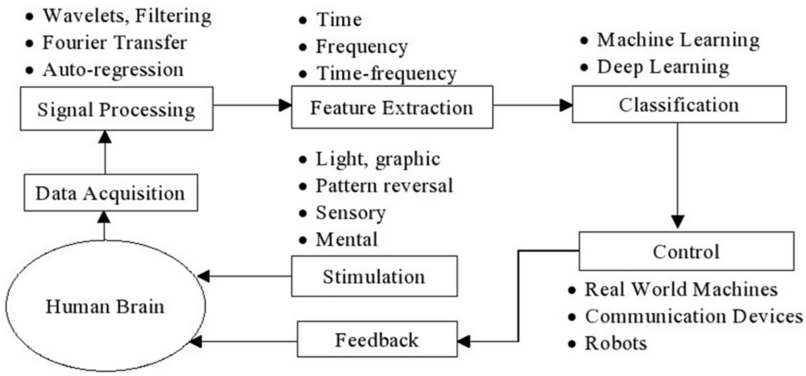

Figure 2. *EEG data acquisition with interfacing setup.*

analysis, relevant articles are taken from the years 2000 to 2021. This survey provides detailed information about all the brain waves. It presents how the classification of EEG signals started and became more accurate with time. The paper shows the statistical analysis of feature domains and classifiers. It also explains the research challenges in terms of future perspectives that should be focused on accurate classification.

1.1. Brain wave classification

Usually, EEG signal is recorded with the eye open and eye closure in relax conditions. These signals are measured from peak to peak and have amplitude of 0.5–100 μ Volt. EEG signal can be differentiated into alpha (α), beta (β), gamma (γ), theta (θ), and delta (δ). If the power spectrum of raw EEG signal is derived, sine wave contribution with frequency difference can be seen. The frequency spectrum is obtained by applying Fourier transform that is described in Table II [12].

1.2. A little bit of history

EEG signal was identified in 1924 by Hans Berger. A few years later, in 1956, a neurological model and time series representation of EEG was described. [13] Cross-correlation gives a relation of the electrical activities of two locations. The autocorrelation of EEG signal was described on behalf of dominant frequencies. [14] Near 1960 with the evolution of machine learning, Hirsch and colleagues recorded evoked potential for the first time. J. J. Denier, V. D. Gon and Strackee, in 1966, elaborated the mathematical frequency analysis of EEG signal. [15] With the resonant frequency, three characteristics (efficiency, time, and frequency selectivity) of filters were observed more significant for this type of analysis. In 1969, a new technique spatiotemporal phenomenon was evolved by Thelema Estrin and Robert Uzgalis to determined the characteristic pattern of EEG signal. [16] Simultaneously, in another study, the EEG signal was classified into two parts; the first part is described as the feature reduction technique to reduce data handling for linear classification. The second part carries the algorithm to classify the unknown samples into any linearly inseparable classes. [17] The signal was classified in the form of abnormal transients like sharp and spike waves in 1975 to treat epileptic patients. [18] With the hierarchy of previous models' help, quantitative analysis of EEG was demonstrated by Arthur C et al.; an operator-independent model could compress the data for complex signals and comfortable for classification techniques. [19] In 1982, the invention of Recurrent Neural Network was like a breakthrough in the classification field. Joseph D Bronzin O published the feature extraction and data reduction technique such as compressed spectral arrays and spectral analysis. [20] A modal in,[21] Ocular Artifacts Removal (OAR), was capable of eliminating noises from almost all physiological situations, which helped in error-free analysis of EEG signal.

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

Table II. EEG waveform analysis based on frequency domain.

EEG signal range (in Hz)	Delta wave (0.1–3)	Theta wave (4–8)	Alpha wave (8–12)	Beta wave (12–30)	Gamma wave (>30)
Classification	Generally broad or spread out	Normally local, may contain two or more lobes	Local, mostly contain full lobe	Confined	Very Confined
Feeling of Subject	An unconscious, deep, and dreamless sleep	Very natural, imagery, and laziness	Relaxing mode without any agitation	Alertness, Disturbance	Focused but arousal
Related task and behaviors	Sleepy without any movement or attention	Inventive, emotional, without focus	Concentration, deep thought but not too much active	Mind activity, for example, mathematics	Deep and focused thought processing
Physical correlation	Less movement, no encouragement	Remedial, assimilation of mind or body	Calm, curing	Attentive and effective	Knowledge-rich tasks performed

In 1986, a tree-structured approach, Classification and Regression Trees (CART), was investigated where the authors implemented a partition method using iterative research for best binary "splits" of data. [22] A. S. Gevins and N. H. Morganva described a model with five detectors for maximal classification accuracy in a pattern recognition analysis. [23] In 1989, the evoked potential waveforms were classified into different patterns, and 90% accuracy was reported. [24] A Time-Domain Analysis Tool (TDAT) was evolved for displaying and analyzing multiple-channel EEG data. It was found that data modification and quantification is relatively easy with the computer. The author concluded that more features and productive results could be calculated using Digital Signal Processing (DSP). [25] The revolution in machine learning classifiers was brought about by Support Vector Machine (SVM) in 1995. In ref. [26], nonlinear effects were observed in EEG signal classification using Artificial Neural Network (ANN), and results were compared with the linear Autoregressive (AR) model. The study concluded that nonlinearity is an essential factor for EEG signals. The ocular effect was reduced using two adaptive filters and a neural network. [27] This procedure was more accurate for real-time processing.

The breakthrough of EEG signal classification for different applications came in trend near 2000. In 2000, Bigan C. and Woolfson M.S. increased the accuracy by combining the polynomial modeling and phase compensation methods. [28] A comparison of SVM classifier was made with LDA and conventional neural networks in a study of linear and nonlinear methods where SVM executed better than others. [29] A new Adaptive Neuro-Fuzzy Inference System (ANFIS) application with five classifiers was described in 2005 for a more accurate classification. [70]

A Common Spatial Pattern (CSP) algorithm reduced artifacts and improved the data transfer rate (number of bits per trial) in an experiment of imagined limb movement activity in Brain–Computer Interface (BCI). [31] A multi-class support vector machine and Probabilistic Neural Network (PNN) classification results were better than a multilayer perceptron neural network. [32] In 2008, a comparison of different spectral signal representations was made by Pawel Herman et al. [71] The band power features were extracted and used for motor imagery classification.In ref. [34], data of epileptic, motor imagery, and mental imagery tasks were classified using the Clustering-Based Least Square SVM (CT-LS-SVM) method with an accuracy of 94.18%, 84.17%, and 61.69%, respectively.

In ref. [35], the author described a kernel-based ensemble learning algorithm that can automatically compute the most discriminative part of the EEG channel for internal emotion recognition. Three techniques having ensemble type architecture are compared in ref. [36] to identify the best classifier where KNN contributed more accurate results. Three feature extraction methods used in a study are Wavelet-based Energy and Entropy (EngEnt), Band Power (BP), and Adaptive Autoregressive (AAR). [37] Several techniques extracted from literature for classification are shown in Fig. 3.

2. Features and extraction methods

For EEG signal, classification and feature extraction are key technology. Classifiers use feature values as input and predict the class for the classifier. A classifier has different parameters that can be learned from training data. [38] Features are extracted in the form of. [39] TDFs, FDFs, TFDFs, and statistical features are shown in Fig. 4. The feature extraction techniques are shown in Fig. 5.

Waveform morphological features contain First Half-Wave Duration (FHWD), First Half-Wave Slope (FHWS), First Half-Wave Amplitude (FHWA), Second Half-Wave Duration (SHWD), Second Half-Wave Amplitude (SHWA), and so on. These characteristics also include Waveform Length (WL), Simple Square Integral (SSI), Root Mean Square (RMS), Mean Absolute Value (MAV), Variance (Var), Average Amplitude Change (AAC), etc. Depending on the applications, various feature extraction techniques are Fourier spectral analysis, Short-Time Fourier Transform (STFT), Discrete Wavelet Transform (DWT). Fourier analysis represents EEG waveform in the frequency domain, and STFT extracts the density of spectrum of the signal using one of the available sliding windows. DWT describes the waveform features in both time or frequency domain. The nonlinear analysis of EEG wave gives the Approximate Entropy (ApEn), Lyapunov Exponents (LyEx), Correlation Dimension (CorDim), etc. [37] Statistical features include standard deviation, correlation coefficient, average value, and normal distribution.

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

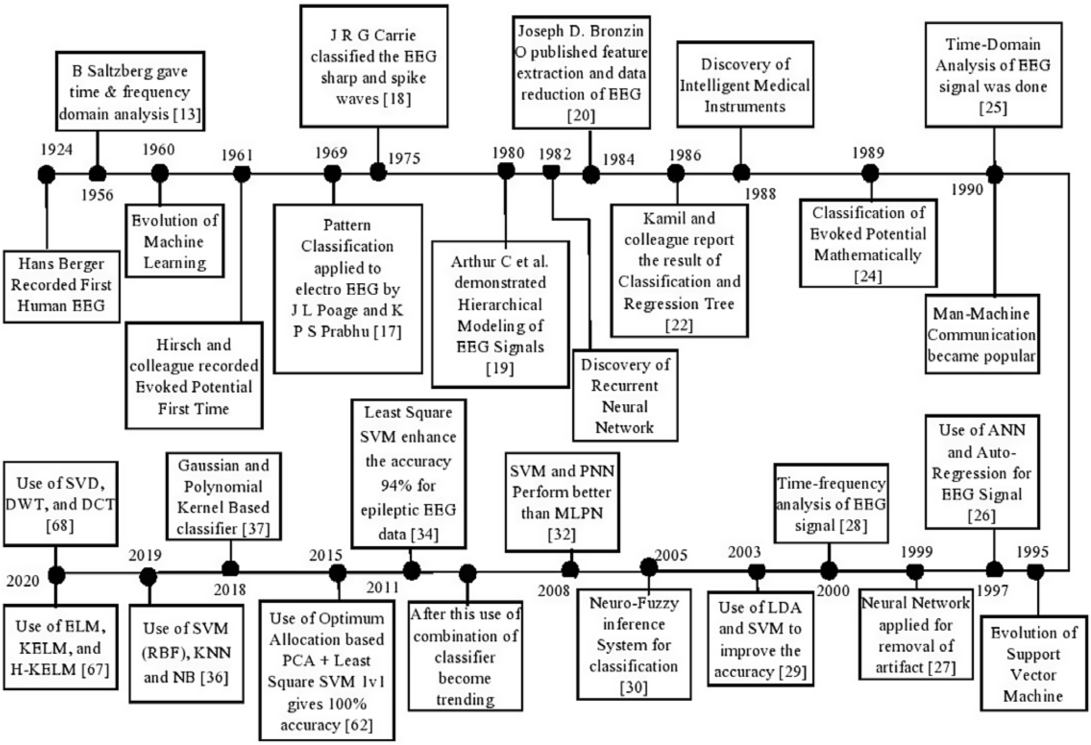

Figure 3. *Breakthrough in classifiers history.*

2.1. Time-domain features (TDFs)

Among time-domain parameters, one variable features are Mean, Standard Deviation, RMS,[40] Peak to Peak, Autocorrelation, Integral of Absolute Value, Zero Crossing (ZC), and Mean of Amplitude (MA). Some more advanced features in the time domain are described below:

1. Histogram – It shows the distribution of EEG data samples.
2. Kurtosis – Kurtosis describes the sharpness of the peak of the frequency distribution curve compared to a Gaussian distribution curve.
3. Skewness – It denotes asymmetry of the distribution curve in comparison to a Gaussian one.
4. Fractal Dimension – It is also called as Hurst Exponent, which is relevant to the long-term memory of a time series.
5. Entropy – Entropy defines the regularity of waves and the unpredictability of fluctuations over our time series.

2.2. Frequency domain features (FDFs)

2.2.1. Fast Fourier transform (FFT)

It is a mathematical tool that represents the signal from time to frequency domain. The examined EEG signal features are calculated by estimating the Power Spectral Density (PSD). [2] The PSD is computed by the autocorrelation function of EEG signal using Welch's method. The information about the sequence is as follows:

$$X_i(n) = x(n + iD),\ here\ n = 0, 1, 2\ldots, M\text{-}1,\ and\ i = 0, 1, 2\ldots.\ L\text{-}1.$$

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

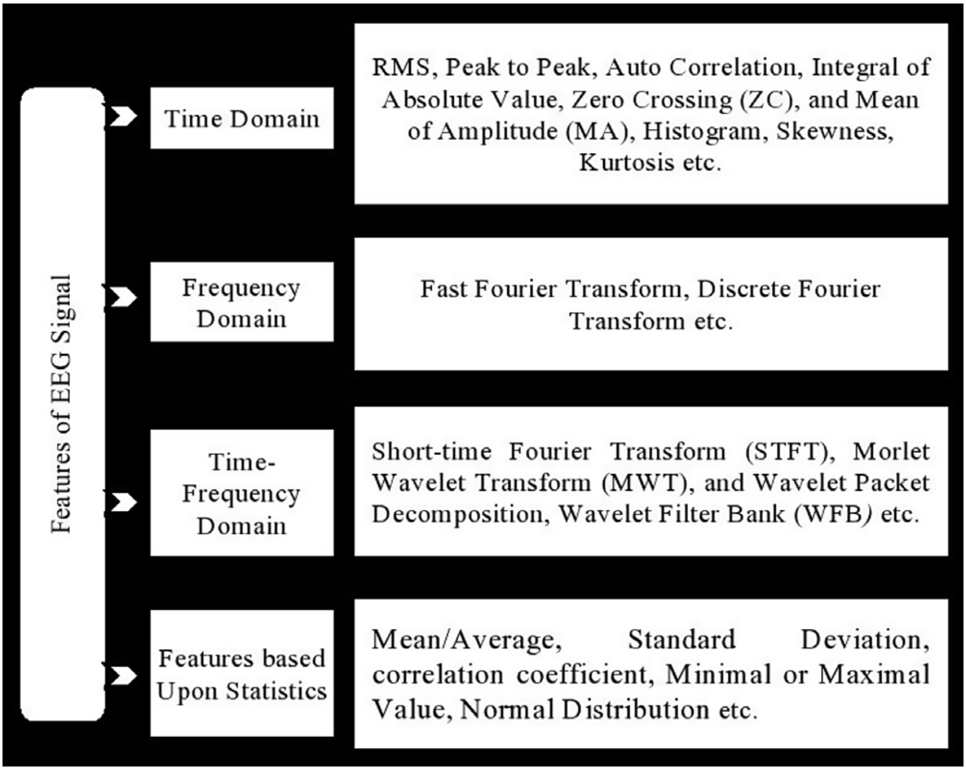

Figure 4. *Feature extraction using different approaches .*

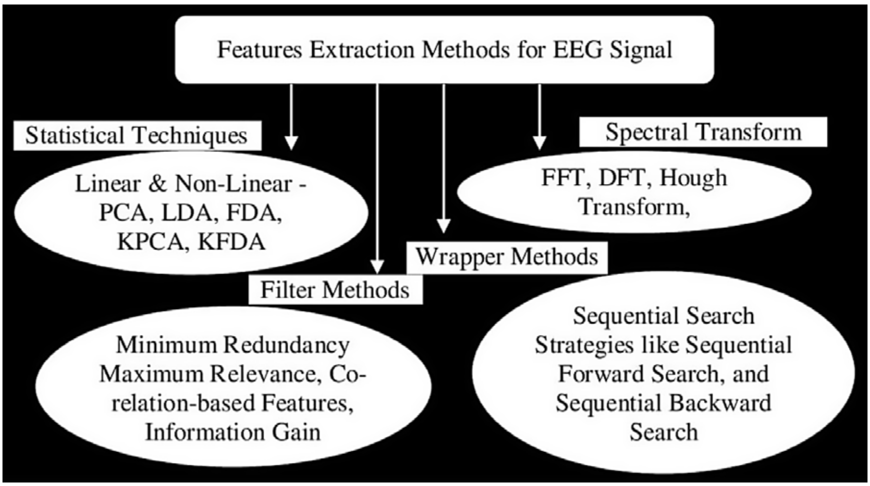

Figure 5. *Frequently adapted methods for feature extraction of EEG signal.*

If $X_i(n)$ is the sequence, iD will be the first point, and L shows the length of 2M, which is a segment of information. The output is presented as

$$P_{xx}^{\approx(t)}(f) = \frac{1}{MU}\left|\sum_{n=0}^{M-1} x_i(n)w(n)e^{-j2\pi fn}\right|^2. \tag{1}$$

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

In the above window function, U is the regularization feature of the power and it is denoted by

$$U = \frac{1}{M}\sum_{n=0}^{M-1} w^2(n) \tag{2}$$

Here, w (n) is window function that describes Welch's power spectral as

$$P_{xx}^{w} = \frac{1}{L}\sum_{t=0}^{L-1} P_{xx}^{\approx(t)}(f). \tag{3}$$

2.3. Time and frequency domain features (TFDFs)

2.3.1. Short-time fourier transform (STFT)

It is used to determine the phase and frequency of the localized part of the signal when it shifts with time. STFT divides the real frame into several frames, which are examined with a time-shifting window frame. It derives a time-varying spectrum of the framed part using Discrete Wavelet Transform (DWT). This framed part moves throughout the recorded signal. Using this method, spectral characteristics behave as constant for each constant frame. STFT is exerted to EEG signal using the following equation:

$$X(u, w) = \sum_{v=-\infty}^{\infty} x(v)\psi(v-\mu)e^{-jwv}. \tag{4}$$

Here, ψ (v) denotes the windowing function. [42]

2.3.2. Morlet wavelet transform (MWT)

The Morlet wavelet transform (MWT) is an efficient means of detecting and analyzing transient signals. MWT is used to find the wavelet power spectrum for a linear frequency axis. The peak response displays on a lower frequency value than the actual value. [94] On computational behalf, its efficiency is low. When it is applied to a signal, it gives a complex and real part.

$$\psi(t) = e^{\frac{-w_\circ^2 t^2}{2}} e^{jw_\circ t}. \tag{5}$$

2.3.3. Wavelet packet decomposition (WPD)

In electroencephalogram signal analysis, high-frequency and low-frequency components are related to time and frequency domain, respectively. This type of signal can be solved using wavelets in both domains (time or frequency).

$$W(j, k) = \sqrt{\sum_{n=1,\ldots L} x_{j,k}^2(n)/L}, \tag{6}$$

where L denotes the number of coefficients, and it can be calculated from depth level and the length of the examined data as L = $M2^{-j}$. [44]

2.3.4. Wavelet filter bank (WFB)

Discrete Wavelet Transform is used to reconstruct the signal. DWT decomposition gives two pairs of functions on both sides are; scaling coefficients (xk+1(n)), and wavelet coefficients (yk+1(n)). The below equation shows the use of two coefficients

$$X^{k+1}(n) = \sum_{i=1}^{2n} h(2n-i)X^k(n), y^{k+1}(n) = \sum_{i=1}^{2n} h(2n-i)X^k(n).$$

Here k denotes the scaling coefficient.

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

2.4. Autoregressive method

This method (AR) determines the Power Spectral Density (PSD) of the given brain wave. PSD is calculated from coefficients of linear system parameters. In autoregressive method, parametric approach is used so it does not have a spectral leakage problem and provides better frequency resolution. AR needs less time to record the data. For the automatic regression technique, creating model properties is a bit complex process. Inappropriate model order selection gives poor estimated results.

2.5. Independent component analysis (ICA)

It can also be defined as a Blind Source Separation (BSS) method. In BSS, without any prior information, a source signal is recovered from mixed signals. If there are n linear mixtures $x_{1,}x_{2,}\ldots x_n$ of n independent components. Then vector x will be x = AS.

Where A denotes a mixing matrix of size n*n, and s is an independent component vector. The goal of ICA is to finds a matrix W, which will be an inverse matrix of A to inverse the mixing effect. [45] It will give an independent component as shown in the equation: y = Wx. ICA depends upon the linearity of the signal and requires more area for computation when it decomposes the signal into a fixed and independent component.

2.6. Principal component analysis (PCA)

It is a generally used technique to extract features and reduce the dimensions. The purpose of reducing the dimensions is to decrease the degree of freedom and decrease space and time complications. This classifier aims to present the data in space so that error variation can be observed. For a defined dataset with z dimensions, mean vector μ and covariance matrix (z*z) summation are calculated. Then, eigen values and eigen vectors are found and sorted toward decreasing eigenvalue. Name these eigenvectors as e_1 for value $k_{1,}e_2$ for k_2, and so on. Usually, there will be a single dimension. [46] PCA accepts data as linear and continuous, so it fails for more complex data. [44]

2.7. Empirical mode decomposition (EMD)

In the 90s, a new concept came for the classification of EEG signal was EMD. It is the application of Hilbert transform and is called the Hilbert–Huang Transform. It extracts information about time-frequency for a nonlinear and nonstationary signal. This classifier works on the principle of decomposition of EEG signal into Intrinsic Mode Function (IMF). Due to mode mixing, this method is more complicated than others. [47] EMD is a data-driven, adaptable technique, and multi-resolution presentation of the signal. It is very sensitive to noise; it may enhance the complications because of mode mixing.

3. Survey of classifier used for EEG signal

This section encapsulated the various classification methods used for EEG signal classification. These classifiers are divided into five classes: Linear and Nonlinear Bayesian Classifiers, Nearest Neighbor Classifiers, Neural Networks (NN), and fusion of classifiers. [48] A brief introduction of mostly used classifiers is described below.

3.1. Linear classifier

Those classifier uses the linear discriminant algorithm to differentiate the classes and are called as linear classifiers. The mostly used algorithms for BCIs are linear discriminant analysis and support vector machine.

3.1.1. Linear discriminant analysis (LDA)

It uses a hyperplane to distinguish the data in various classes. For two classes, feature vector classification depends on where the vector lies either below or above the hyperplane. LDA assumes an equal covariance matrix for both classes. For N-class problems, different hyperplanes are required. This strategy distinguishes each class from all others using "one vs rest method" (OVR). This classifier aims to generate a new variable that contains the original predictors [49] It is done by maximizing the difference between the predetermined group and a new variable. The aim is to collect the predictor score to form a new combined variable and discriminant score. LDA is more suitable for online BCIs because of its low computation requirement. Usually, it gives good results with the drawback of its linearity, so it is not useful for nonlinear data classification.

3.1.2. Support vector machines (SVMs)

SVM also contains a hyperplane to distinguish the classes; this hyperplane maximizes the margin. As the margin is maximized, the accuracy of classification will increase. SVM uses linear boundaries for classification, so-called linear SVM. For synchronous BCIs problems, it gives relatively successful classification results. By expanding the bit complexity of the classifier and varying the kernel value, it can also be used for nonlinear data classification. It makes the input patterns fit into a higher dimensional space by applying some nonlinear mapping. Then, linear decision surface is created in the high-dimensional feature space. This classifier is linear and used for classification and regression problems, but it becomes a nonlinear classifier when data are mapped into nonlinear mapping. [50] Support Vector Classifier is too much similar to the perceptron classifier.

3.2. Neural network

The most used category of the classifier for classification is Neural Network. It is a combined form of various artificial neurons that help in creating nonlinear boundaries for decision. The below section briefly represents the widely used classifier for EEG, that is, multilayer perceptron and other neural networks.

3.2.1. Multilayer perceptron

MLP is a group of interconnected nodes and looks like a network of neurons in the brain. So it can be defined as interconnections between different layers of the system. It is a three-layered structure; the inner layer, the middle, the hidden one, and the output layer. The first layer transmits data through synapses to the second one. This second layer of neurons sends data to the third layer (output layer) through synapses. MLP and NN can approximate any continuous function by assembling the appropriate neurons and layers. This classifier is compatible with all types of BCIs, synchronous or asynchronous, linear or nonlinear, binary, and multi-class. It is more sensitive to noise and nonlinearity. Without hidden layers, MLP behaves as a perceptron, which is the same as LDA.

3.3. Nonlinear bayesian classifiers

The mainly used Bayesian classifiers for EEG data are Hidden Markov Model (HMM) and Bayesian Graphical Network (BGN). These classifiers also give nonlinear boundaries for decision-making. They can reject uncertain samples more efficiently than the above classifiers.

3.3.1. Bayes quadratic

Bayesian classifiers assign the feature vector to that class, which has the maximum probability. When this vector is associated with an existing class, the computed value is called posterior probability. As its name suggests, it contains quadratic decision boundaries. This classifier is more suitable for motor imagery tasks and mental tasks classification.

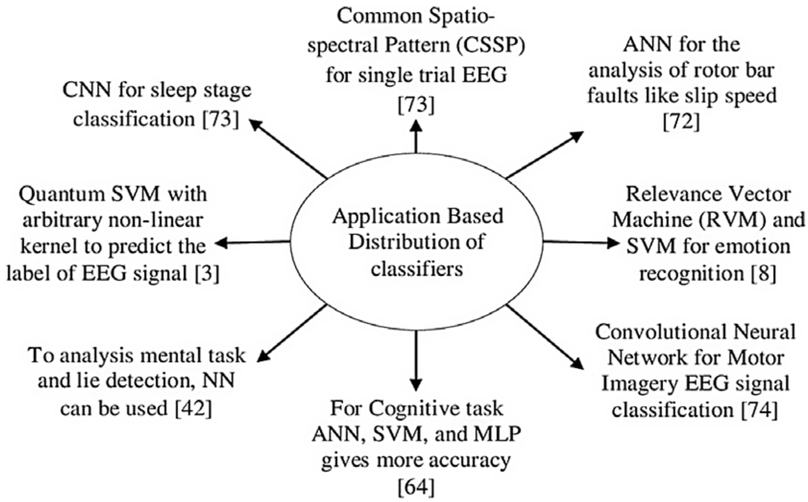

Figure 6. *Representation of individual classifier for particular application.*

3.3.2. Hidden markov model

HMM is famous for EEG analyzing data in speech recognition. [48] It is the probabilistic automation where the best sequence is selected from the given sequence of feature vectors. These probabilities are usually Gaussian Mixture Models (GMM) for EEG signals. HMM algorithms are suitable for time series classification. This classifier is helpful for temporal sequence and raw EEG signal classification.

Another type of HMM is Input–Output HMM (IOHMM). It is a discriminative classifier, not generative. It can discriminate various classes, but in the case of HMM, individual HMM is needed for each class.

3.4. Nearest neighbor classifiers

These are relatively simple classifiers than others. Here the feature vectors are assigned to Nearest Neighbor(s) class. If this feature vector belongs to a class from the training set, it is called K-Nearest Neighbor (KNN).

3.4.1. K-nearest neighbors

It is a nonlinear technique in which a feature vector is assigned to the nearest neighbor within the training set. For a large value of K and the sufficient training samples, KNN gives better results for any function. It had been used for statistical estimation and pattern recognition.

4. Results and analysis

From the present survey of different feature extraction and classification techniques, some classifiers are used repeatedly for a particular area, as shown in Fig. 4. SVM is recognized as a multi-purpose classifier in applications like emotion recognition (sadness, happiness, and neutral). Here, it predicts the label of EEG as signal recording during complex cognitive tasks and rest conditions with eyes open and close. [3, 8] From literature, it can be said that neural network helps for mental task classification. [42, 72] Cognitive problems can be solved using MLP, SVM, and ANN depending upon the individual requirement. [62, 72] Convolution Neural Networks perform more accurately than other classifiers to classify the sleep stages and motor imagery tasks,[64, 3] as shown in Fig. 6.

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

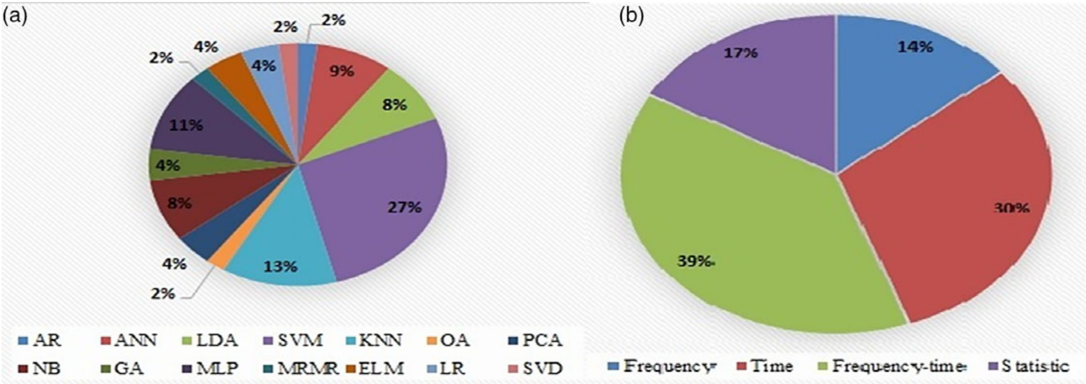

Figure 7. *(a) Statistical analysis of classifiers used in previous studies and (b) statistical analysis of features extracted in various studies.*

Statistical analysis of features extracted and classifiers used in previous studies are shown in Fig. 7(a) and (b), respectively. Here, AR, ANN, LDA, SVM, KNN, NB, GA, MLP, MRMR, and ELM represents Autoregression, Artificial Neural Network, Linear Discriminant Analysis, Support Vector Machine, K-Nearest Neighbor, Naïve Bayes, Genetic Algorithm, Multilayer Perceptron, Minimum Redundancy Maximum Relevance, and Extreme Learning Machine. The most used classifier is SVM, which is used in 27% of studies. Table III mentioned the sensitivity, specificity, and accuracy of previously used classifiers. Table IV shows the number of channels used for data recording, toolbox, number of subjects, and features (TDFs, FDFs, and TFDFs) extracted in various studies. It can be said that repeatedly used classifiers in Table III. give more accuracy which helps the practitioner in a random analysis to determine the suitable classifier. Features that can be discriminated against on time, frequency, and time–frequency domains are shown in Fig. 7(b). The mostly used features are TFDFs. After that, TDFs features also play a significant role in extracting the information. Different classifiers can be categorized as supervised, unsupervised, and probabilistic learning. The pros and cons of each category classifiers are described in Table V.

5. Conclusion

This paper presents a literature survey of widely used feature domains, extraction techniques, and classifiers used in biomedical applications for EEG signals. The mostly used feature domain is TFDFs, which is analyzed using STFT or Wavelet Transform (WT) and it is used in 39% of studies out of total extracted articles. Generally, TFDFs are applied in joint with statistical and amplitude-related parameters. TDFs are also used in various studies, but different feature domains are combined in a single feature vector to achieve complementary information. The classification approaches are divided into four main categories: Linear, Neural Network, Nearest Neighbor, and Nonlinear Bayesian classification. Some miscellaneous techniques are also used to enhance the classification accuracy. Nonlinear classifiers are very accurate to improve reliability. In a number of studies, Support Vector Machine (SVM) classifier with different kernels gives better results in terms of accuracy, sensitivity, and specificity. Comparison tables of extracted studies with different parameters are explained to understand the feature domains and classification in detail. Statistical analysis of surveyed feature domains and classifiers are illustrated graphically to understand the significance of the individual. No classifier guarantees for every classification, but individuals give accuracy for a particular field.

Table III. *Detailed description of accuracies, sensitivity, and specificity of classifiers.*

Researcher	Year	Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
B O Peters, G Pfurtscheller, and H Flyvbjerg. [51]	2001	Autoregressive Model (AR) Artificial Neural Network (ANN)	–	–	For Male (A4) – 94 For Female (B6) – 95 For Male (B8) – 91
E Yom-Tov, G F Inbar. [52]	2002	Genetic Algorithms Movements of Two Limbs Multiple Limb	–	–	87 63
D Garrett et al. [53]	2003	Linear Discriminant Analysis (LDA) Neural Networks SVM	–	–	66 69.4 72.0
I Guler and E D Ubeyli. [55]	2007	Support Vector Machine (SVM)	Data A 99.25	99.84	75.6
			B 99.38	99.84	
		Probabilistic Neural Network (PNN)	C 99.25	99.75	
			D 99.38	99.65	72.0
			E 99.13	100.00	
		MLPNN	Data A 98.25	99.62	
			B 98.13	99.56	
			C 98.00	99.40	
			D 98.13	99.12	68.8
			E 97.75	99.84	
			Data A 93.25	98.42	
			B 93.63	98.36	
			C 94.00	98.16	
			D 94.13	97.17	
			E 93.13	99.54	

Table III. *Continued.*

Researcher	Year	Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
P Herman et al. [55]	2008	LDA	-	-	SVM perform
		Regularized Fischer Discriminant (RFD)			
		SVM			Better than others
J Li et al. [56]	2009	PSD	-	-	61.0
		Common Spatial Patterns			75.6
		Nonnegative Multiway Factorization			68.8
		Tensor-Based Scheme (TbS)			76.3
Y U Khan, F Sepulveda. [57]	2010	Radial Basis Function (RBF)	89 %	11 %	About 90
L Guo et al. [58]	2011	1.Normal and seizure EEG classification			88.6
		KNN-alone classifier			99.2
		GP-KNN classifier			67.2
		2. Normal, seizure-free, and seizure EEG classification			93.5
		KNN-alone classifier			
		GP-KNN classifier (Normal)			
10pt] Siuly, Y Li, and P Wen. [34]	2011	Least Square SVM (EEG Dataset)			
		Epilepsy	94.92	93.44	94.18
		Motor Imagery	83.98	84.37	84.17
		Mental Imagery Tasks	64.61	58.77	61.69
W Yi et al. . [59]	2013	Support Vector Machine (SVM)	-	-	84
X W Wang, D Nie, and B L Lu. [60]	2014	SVM with different kernels	-	-	
		Linear			87.53
		Polynomial			82.09
		RBF			72.43
M H Alomari et al. [61]	2014	Neural networks (NN)			

Table III. *Continued.*

Researcher	Year	Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
S Siuly, Y Li. [62]	2015	Optimum Allocated PCA with	-	-	89.11
		Least Square SVM ECOC			99.97
		Least Square SVM MOC			99.97
		Least Square SVM 1v1			100.0
		Least Square SVM 1vA			99.96
		NB classifier			99.24
		KNN classifier			98.82
		LDA classifier			87.79
H U Amin. [63]	2015	For approximate coefficients (A4)			
		SVM	100%	97.50%	98.75
		MLP	100%	96.40%	98.21
		K-NN	98.60%	96.80%	98.21
		Naive Bayes	75.00%	92.10%	83.57
		For detailed coefficients (D4)			
		SVM	99.60	96.80	98.21
		MLP	99.60	97.60	98.57
		K-NN	98.90	95.40	97.14
		Naive Bayes	84.60	81.40	83.03
J Atkinson, D Campos. [64]	2016	Minimum Redundancy Maximum Relevance (mRMR)	-	-	
		1 Arousal			60.72
		2 Valence			62.39
		Genetic Algorithm SVM			
		1 Arousal			
		2 Valence			56.69
					53.46

Table III. Continued.

Researcher	Year	Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
H U Amin. [65]	2017	Dataset 1			
		(For approximate coefficients)			
		SVM	99.28	98.93	99.11
		MLP	96.48	97.83	97.14
		NB	88.24	90.77	89.63
		k-NN	96.88	99.63	89.63
		(For detailed coefficients)			
		SVM			
		MLP	97.89	99.28	98.57
		NB	88.70	94.98	91.60
		KNN	82.71	79.59	81.07
		Dataset 2	98.22	98.56	98.39
		(For approximate coefficients)			
		SVM	87.93	85.48	86.67
		MLP	89.93	88.52	89.17
		NB	77.42	79.31	78.33
		KNN	80.95	84.21	82.50
		(For detailed coefficients)			
		SVM	87.50	92.86	90
		MLP	90.48	94.74	92.50
		NB	82.54	85.96	84.17
		KNN	91.94	94.83	93.33
Y Zhang et al. [37]	2018	1. MLP with one hidden layer	-	-	75.9
		2. Conventional SVM			76.3
		3. SVM (Gaussian Kernel)			76.7
		4. SVM (Polynomial Kernel)			76.5
		5. Multi-kernel SVM (Gaussian and Polynomial Kernels)			77.9

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

Table III. *Continued.*

Researcher	Year	Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)
		6. Extreme Learning Machine			77.0
		7. ELM (Gaussian Kernel)			77.5
		8. ELM (Polynomial Kernel)			77.9
		9. Multi-Kernel ELM (Gaussian and Polynomial Kernels)			78.9
A Datta and R Chatterjee. [36]	2019	KNN (K=5)	-	-	79.76
		KNN (K=7)			79.52
		KNN (K=9)			78.81
		SVM (RBF)			69.29
		NB			76.43
C Ieracitano et al. [66]	2019	Autoencoder, Multilayer Perceptron, Logistic Regression, SVM with three feature sets	Continuous Wavelet Transform, and bi-spectrum representation (CWT + BiS) fusion of two gives more accuracy than individual feature set		
K Venkatachalam et al. [67]	2020	PCA	H-KELM gives the highest – 95.64		
		Fisher's linear discriminant			
		Extreme Learning Machine (ELM)			
		Kernel ELM			
		H-KELM (Hybrid-Kernel ELM)			
T H M Delsy et al. [68]	2020	Singular Value Decomposition (SVD)	75	90.41	84
		Framelet	97.12	97.95	97.6
		DWT	100	98.6	99.17
		DCT	100	99.3	99.58
R G Andrzejak et al. [69]	2021	KNN	-	-	99.20

Table IV. *Summary of feature extraction, system channel, and toolbox used in different studies for analysis purpose.*

Study	**Year**	**System channel/ electrode used**	**Toolbox**	**No. of subjects**	**No. of features**
E Yom-Tov, G F Inbar. [52]	2002	8	MATLAB	Four males and one female	Autoregressive Coefficients, Power Spectral Density, Mean, Amplitude, Standard Deviation
D Garrett et al. [29]	2003	6 Channel	Lab Master 12-bit A/D converter mounted in an IBM-AT computer	2	Autoregressive Coefficients
I Guler and E D Ubeyli. [70]	2007	128- Channel Amplifier System	MATLAB 7.0 with neural-networks	5	Wavelet Transform (WT)
P Herman et al. [71]	2008	2 Channel	Pentium IV 3 GHz, 512MB RAM)	11	Different Band Power Features
J Li et al. [56]	2009	62 Channel	SynAmps2, Neuro scan, Charlotte, NC	9	Wavelet Transform (WT)
Y U Khan and F Sepulveda. [57]	2010	64 Channel	Biosemi Active Two hardware and MATLAB	5	Discrete Wavelet Transform (DWT)
S Siuly, Y Li, and P P Wen. [34]	2011	118 Electrode used from 128	MATLAB (version7.7, R2008b)	13 healthy volunteers and 5 epileptic patients	Mean, Median, and Mode, Minimum, Maximum, First and Third quartile
W Yi et al. [59]	2013	64	Neuroscan SynAmps2	3 males and 7 females	Event-Related Spectral Perturbation (ERSP), Power Spectral Entropy (PSE), and Spatial Distribution Coefficient

Table IV. Continued..

Study	Year	System channel/ electrode used	Toolbox	No. of subjects	No. of features
X W Wang, D Nie, B L Lu. [60]	2014	128	ESI-128, Neuro Scan Labs), SCAN4.2software, and 64-channel Quick Cap with electrodes Ag/AgCl	3 males and 3 female	Power Spectrum, Wavelet and Nonlinear Dynamical Analysis
M H Alomari et al. [61]	2014	64	BCI2000,MATLAB Toolbox EEGLAB	109 healthy subjects	RMS, Variance of EEG, MAV, Integrated EEG, Simple Square Integral (SSI), Average Amplitude Change
S Siuly, Y Li. [62]	2015	128	MATLAB 7.14, R2012a LS-SVM lab toolbox (version 1.8)	5 healthy volunteers 5 epileptic patients	Hyper parameters (γ, σ^2)
H U Amin. [63]	2015	128	HydroCel Geodesic Net Polygraph Input Box (PIB)	8 males	Wavelet Relative Energy
J Atkinson, D Campos. [64]	2016	14 Channel used from 32	BCI device	32 human subjects	Hjorth parameters, Band power for different frequencies, and fractal dimension for a Channel
H U Amin. [65]	2017	128	IBM-AT	Dataset 1 8 subjects Dataset 2 7 ubjects	Relative energy features
Y Zhang et al. [37]	2018	118	-	BCI Competition III dataset IVa- 5 Subject BCI Competition IV dataset IIb-9 Subject	Using Common spatial pattern (CSP)

Table IV. *Continued..*

Study	Year	System channel/ electrode used	Toolbox	No. of subjects	No. of features
A Datta, R Chatterjee. [36]	2019	3	MATLAB	6 healthy females	Wavelet energy and entropy Band power (BP) Adaptive Autoregressive Parameters (AAR)
C Ieracitano et al. [66]	2019	19	MATLAB toolbox EEGLab	189 subjects (63 patients of each – Alzheimer's disease and Mild Cognitive Impairment), 63 healthy controls	Continuous Wavelet Transform (CWT), Bi-spectrum, and (CWT + BiS)
K Venkatachalam et al. [67]	2020	-	Matrix Laboratory version R2016a	5 physically fit persons	Band Power in the μ and β rhythms
R G Andrzejak et al. [69]	2021	1	MATLAB 2018a	100 segments	Hamsi-Pat and iterative neighborhood component analysis (INCA) based feature generation

Table V. *Pros/cons of frequently used classifiers.*

Sr. No	Classifier	Supervised Learning
1.	ANN	Pros – It can handle the incomplete and noisy data. Capable to tolerate faults in the system. Cons – It is hardware dependent. Unable to work with small data. [75]–[78]
2.	SVM	Pros – It reduces the overfitting problems. Due to kernel function it can solve difficult issues. Cons – But to choose an appropriate kernel is a trouble. Perform less accurately on small data. [79]–[82]
3.	RF	Pros – RF is an ensemble machine learning tool. It decreases the overfitting. It is flexible and gives more accurate results. Data scaling is not mandatory. Cons – But consumes more time and few times interpretation becomes difficult. [96]–[84]
4.	KNN	Pros – Gives more approximate results if training samples are large. Muscular for noisy data. Its implementation is easy. Cons – "K"' values must be known. It has high computation cost. [85]–[87]
5.	DT	Pros – It is an upgraded version of the C4.5 classifier. It is simple to understand and can handle missing values. It makes regression and classification like a tree structure. Cons – It gives expectation based on results that may lead to inaccuracy. [86],[88],[89]
6.	LR	Pros – It is an easily handable tool that can minimize the overfitting problem. Cons – Outlier sensitivity is high. [91]
		Unsupervised Learning
7.	K-Means	Pros – It can solve missing data problems. Less cost for computation. Cons – Unable to choose cluster numbers. Gives less efficiency to work with global cluster. [83],[92]
		Deep Learning
8.	RNN, CNN	Pros – Accuracy is better than others. Weight sharing characteristic is also good. Cons – Needs large dataset to analyze the results and computational cost is a bit high. [93],[94]
		Probabilistic
9.	NB	Pros – It is easy to implement that can perform prediction also. It needs less training data. It can execute probabilistic prediction also. Cons – Provide less accurate result. [96],[82],[85],[95]
10.	HMM	Pros – This classifier is more comfortable to fit any dataset. Gives better compression. Cons – Interpretation is difficult. Requires comparatively more time and memory for computation. [96]

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

Future Directions. Research cannot be assumed sufficient until and unless the problem remains unresolved in the literature. It makes the basic to accomplish further research by proposing a novel approach to fulfill the gaps pinpointed.

(1) During training, the runtime is calculated by any technique like Hamilton's simulation and phase estimation module. During testing, runtime depends upon the complexity of the system. But various classifiers follow the logarithmic or polylogarithmic complexity in the classification part. So, this runtime period of the classifiers should be reduced to decrease the complexity of the algorithm.
(2) Hybrid designs are giving reliable results for classification accuracy and precision still, it needs more in-depth research for the use of a combination of classifiers.
(3) Excepting the classifier design, it also needs how the existing network interacts with noisy data. The demands for algorithms that can work online and with nonstationary data are increasing day by day.
(4) By analyzing the quantum information, various other features can be extracted. Kernel-based nonlinear models can be explored more for dimension reduction and nonlinear regression.

Acknowledgments. We are grateful to anonymous referees for their valuable suggestions and also to the authors of all the research papers we have been through.

References

[1] S. Mantri, V. Dukare, S. Yeogle, D. Patil and V. M. Wadhai, "A survey: fundamentle of EEG," C. Science and M. Studies, *Int. J. Adv. Res. Comput. Sci. Manage. Stud.* **1**, 83–89 (2013).
[2] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains," *ISRN Neurosci.* **2014**, 1–7 (2014).
[3] Y. C. Li, R. Zhou, R. Q. Xu, J. Luo and S. X. Jiang, "A quantum mechanics-based framework for EEG signal feature extraction and classification," *IEEE Trans. Emerging Top. Comput.* **14**(8), 1–11 (2020).
[4] A. Al-Nafjan, M. Hosny, Y. Al-Ohali and A. Al-Wabil, "Review and classification of emotion recognition based on EEG brain-computer interface system research: A systematic review," *MDPI Appl. Sci.* **7**(12), 1–34 (2017).
[5] 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," *J. Neural Eng.* **15**(3), 1–28 (2018).
[6] M. Hamada, B. B. Zaidan and A. A. Zaidan, "A systematic review for human EEG brain signals based emotion classification, feature extraction, brain condition, group comparison," *J. Med. Syst.* **42**(162), 1–25 (2018).
[7] A. Craik, Y. He and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.* **16**(3), 1–28 (2019).
[8] Q. Gao, C. han Wang, Z. Wang, X. lin Song, E. zeng Dong, and Y. Song, "EEG based emotion recognition using fusion feature extraction method," *Multimedia Tool Appl.* **79**, 27057–27074 (2020).
[9] G. Schalk, D. J. McFarland, T. Hinterberger, B. Birbaumer and J. R. Wolpaw, "BCI2000: A general-purpose Brain-Computer Interface (BCI) system," *IEEE Trans. Biomed. Eng.* **51**(6), 1034–1043 (2004).
[10] I. Zyma, S. Tukaev, I. Seleznov, K. Kiyono, A. Popov, M. Chernykh and O. Shpenkov, "Electroencephalograms during mental arithmetic task performance," *Data MPDI* **4**(1), 2–7 (2019).
[11] A. Matran-Fernandez and R. Poli, "Towards the automated localisation of targets in rapid image-sifting by collaborative braincomputer interfaces," *PLoS One* **12**(5), 1–28 (2017).
[12] T. Jung, "Introduction to electroencephalogram basic physics of EEG," BIOE 280A, Center for Advanced Neurological Engineering and Swartz Center for Computational Neuroscience and University of California San Diego, USA and Department of Computer Science National Chiao-Tung University, Hsinchu, Taiwan. https://cfmriweb.ucsd.edu/ttliu/be280a_12/BE280A12_IntrotoEEG.pdf. [Accessed on 7 September, 2020]
[13] B. Saltzberg, N. R. Burch, A. Miles and E. G. Correll, "A new approach, to signal analysis in electroencephalography," *National Electronics Conference, Chicago, Illinois*, (1956) pp. 24–30.
[14] J. S. Barlow, "Autocorrelation and cross-correlation analysis," *IRE Trans. Med. Electron.* **ME-6**(3), 179–183 (1959).
[15] M. C. Houston, "Some aspects of A College Health Service," *Am. J. Nurs.* **42**(10), 1183–1189 (1942).
[16] P. L. Nunez, "Representation of evoked potentials by Fourier-Bessel expansions,"*IEEE Trans. Biomed. Eng.* **74**, 372–374 (1973).
[17] J. L. Poage and K. P. S. Prabhu, "Pattern classification applied to electro-encephalographs," *Technical Report No. **1** harvard University Cambridge, Massachusetts*, (1956) pp. 1–56.
[18] J. R. G. Carrie, "Computer methods for detecting and classifying EEG spikes and sharp waves," *Am. J. EEG Technol.* **15**(2), 68–74 (1975).
[19] A. C. Sanderson, J. Segen and E. Richey, "Hierarchical modeling of EEG signals," *IEEE Trans Pattern Anal. Mach. Intell.* **PAMI-2**(5), 405–415 (1980).
[20] J. D. Bronzino, "Quantitative analysis of the EEG—general concepts and animal studies," *IEEE Trans. Biomed. Eng.* **31**(12), 850–856 (1984).

[21] E. C. Ifeachor, B. W. Jervis, E. L. Morris, E. M. Allen and N. R. Hudson, "A new microcomputer-based online ocular artefact removal (OAR) system," *IEEE Proc. Phys. Sci. Meas. Instrum. Manage. Educ. Rev* **133**(5), 291 (1986).

[22] K. A. Grajski, L. Breiman, G. V. Di Prisco and W. J. Freeman, "Classification of EEG spatial patterns with a tree-structured methodology: CART," *IEEE Trans. Biomed. Eng.* vol. **BME-33**(12), 1076–1086 (1986).

[23] N. H. Morgan, "Classifier-directed signal processing," *IEEE Trans. Biomed. Eng.* **33**(12), 1054–1068 (1986).

[24] V. H. Clarson and J. J. Liang, "Mathematical classification of evoked potential waveforms," *IEEE Trans. Syst. Man Cybern.* **19**(1), 68–73 (1989).

[25] S. Park, J. C. Principe, J. R. Smith and S. A. Reid, "TDAT—time domain analysis tool for EEG analysis," *IEEE Trans. Biomed. Eng.* **37**(8), 803–811 (1990).

[26] N. Hazarika, "Non-linear considerations in EEG signal classification," *IEEE Trans. Signal Process.* **45**(4), 829–836 (1997).

[27] S Selven and R Srinivasan, "Removal of ocular artifacts from EEG using an efficient neural network based adaptive filtering technique," *IEEE Signal Process. Lett.* **6**(12), 330–332 (1999).

[28] C. Bigan and M. S. Woolfson, "Time-frequency analysis of short segments of biomedical data," *IEEE Proc. Sci. Meas. Technol.* **147**(6), 368–373 (2002).

[29] D. Garrett, D. A. Peterson, C. W. Anderson and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**(2), 141–144 (2003).

[30] I. Guler and E. D. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *J. Neurosci. Methods* **148**(2), 113–121 (2005).

[31] S. Lemm, B. Blankertz and G. Curio, "Spatio-spectral filters for robust classification of single trial EEG," *IEEE Trans. Biomed. Eng.* **52**(9), 1–7 (2002).

[32] E. D. Guler and I. Ubeyli, "Multi-class support vector machines for EEG- multi-class support vector machines for EEG-signals classification," *IEEE Trans. Inf. Technol. Biomed.* **11**, 117–126 (2016).

[33] P. Herman, G. Prasad, T. M. McGinnity and D. Coyle, "Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**(4), 317–326 (2008).

[34] S. Siuly, Y. Li and P. P. Wen, "Clustering technique-based least square support vector machine for EEG signal classification," *Comput. Methods Programs Biomed.* **104**(3), 358–372 (2011).

[35] H. Ullah, M. Uzair, A. Mahmood, M. Ullah, S. D. Khan and F. A. Cheikh, "Internal emotion classification using EEG signal with sparse discriminative ensemble," *IEEE Access* **7**(c), 40144–40153 (2019).

[36] A. Datta and R. Chatterjee, "Emerging technologies in data mining and information security," *Adv. Intell. Syst. Comput.* **755**, 145–154 (2019).

[37] Y. Zhang, Y Wang, G Zhou, J Jin, B Wang, X Wang and A Cichocki, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Syst. Appl.* **96**, 302–310 (2018).

[38] S. M. R. Islam, A. Sajol, X. Huang and K. L. Ou, "Feature extraction and classification of EEG signal for different brain control machine," *3rd International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2016* (2017).

[39] W. Ren, M. Han, J. Wang, D. Wang and T. Li, "Efficient feature extraction framework for EEG signals classification," *7th International Conference on Intelligent Control and Information Processing, ICICIP 2016 - Proceedings*, pp. 167–172 (2017).

[40] K. Veer, "A technique for classification and decomposition of muscle signal for control of myoelectric prostheses based on wavelet statistical classifier," *Measurement* **60**, 283–291 (2015).

[41] A. S. Al-Fahoum and A. A. Al-Fraihat, "Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains," *ISRN Neurosci.* **2014**, 1–7 (2014).

[42] C. Uyulan and T. T. Erguzel, "Analysis of time - frequency EEG feature extraction methods for mental task classification," *Int. J. Comput. Intell. Syst.* **10**(1), 1280 (2017).

[43] L. Sun, B. Jin, H. Yang, J. Tong, C. Liu and H. Xiong, "Unsupervised EEG feature extraction based on echo state network," *Inf. Sci.* 475, 1–17 (2019).

[44] A. Sadeghian, Z. Ye and B. Wu, "Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks," *IEEE Trans. Instrum. Meas.* **58**(7), 2253–2263 (2009).

[45] I. Rejer, P. Górski, I. Rejer, P. Górski, I. Component and E. E. G. Data, "Independent component analysis for EEG data preprocessing - algorithms comparison," *Int. Fed. Inf. Process.* (2017).

[46] M. K. Lakshmanan, H. Nikookar and H. Nikookar, "Construction of optimum wavelet packets for multi-carrier based spectrum pooling systems," *Wireless Press Commun.* **54**, 95–121 (2010).

[47] A. R. Mane, P. S. D. Biradar and P. R. K. Shastri, "Review paper on feature extraction methods for EEG signal analysis," *Int. J. Emerging Trend Eng. Basic Sci. (IJEEBS)* **2**(1), 545–552 (2015).

[48] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *J. Neural Eng.* **4**(2), 1–13 (2007).

[49] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Syst. Appl.* **37**(12), 8659–8666 (2010).

[50] P. N. Kumar and H. Kareemullah, "EEG signal with feature extraction using SVM and ICA classifiers," *Int. Conf. Inf. Commun. Embedded Sys. ICICES 2014* **85**(3), 1–7 (2015).

https://doi.org/10.1017/S0263574721000382 Published online by Cambridge University Press

[51] B. O. Peters, G. Pfurtscheller and H. Flyvbjerg, “Automatic differentiation of multichannel EEG signals,” *IEEE Trans. Biomed. Eng.* **48**(1), 111–116 (2001).
[52] E. Yom-Tov and G. F. Inbar, “Feature selection for the classification of movements from single movement-related potentials,” *IEEE Trans. Neural Sys. Rehabil. Eng.* **10**(3), 170–177 (2002).
[53] A. S. Rodionov and A. A. L’vov, “Comparison of linear, non-linear and feature selection methods for EEG signal classification,” *Conf. Proc. – Int. Conf. Actual Probl. Electron. Devices Eng. APEDE’* **2004**(1), 436–439 (2004). doi:10.1109/apede.2004.1393604.
[54] W. Min, H. Cui, H. Rao, Z. Li and L. Yao, “Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics,” *IEEE Access* **6**(Cccv), 9324–9335 (2018).
[55] R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieva and E. Moya-Albor, “A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-Fall detection dataset,” *Comput. Biol. Med.* **115** (2019).
[56] J. Li, L. Zhang, D. Tao, H. Sun and Q. Zhao, “A Prior neurophysiologic knowledge free tensor-based scheme for single trial EEG classification,” *IEEE Trans. Neural Sys. Rehabil. Eng.* **17**(2), 107–115 (2009).
[57] Y. U. Khan and F. Sepulveda, “Brain-computer interface for single-trial EEG classification for wrist movement imagery using spatial filtering in the gamma band,” *IET Signal Process.* **4**(5), 510–517 (2010).
[58] L. Guo, D. Rivero, J. Dorado, C. R. Munteanu and A. Pazos, “Automatic feature extraction using genetic programming: An application to epileptic EEG classification,” *Expert Syst. Appl.* **38**(8), 10425–10436 (2011).
[59] W. Yi, S. Qiu, H. Qi, L. Zhang, B. Wan and D. Ming, “EEG feature comparison and classification of simple and compound limb motor imagery,” *J. Neuro Eng. Rehabil.* **10**(1), 1–12 (2013).
[60] X. W. Wang, D. Nie and B. L. Lu, “Emotional state classification from EEG data using machine learning approach,” *Neurocomputing* **129**, 94–106 (2014).
[61] M. H. Alomari, E. A. Awada, A. Samaha and K. Alkamha, “Wavelet-based feature extraction for the analysis of EEG signals associated with imagined fists and feet movements,” *Comput. Inf. Sci.* **7**(2), 8–12 (2014).
[62] S. Siuly and Y. Li, “Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification,” *Comput. Methods Programs Biomed.* **119**(1), 29–42 (2015).
[63] H. U. Amin, AS Malik, RF Ahmad, N Badruddin, N Kamel, M Hussain, WT Chooi, “Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques,” *Australas. Phys. Eng. Sci. Med.* **38**(1), 139–149 (2015).
[64] J. Atkinson and D. Campos, “Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers,” *Expert Syst. Appl.* **47**, 35–41 (2016).
[65] H. U. Amin, W. Mumtaz, A. R. Subhani, M. N. M. Saad and A. S. Malik, “Classification of EEG signals based on pattern recognition approach,” *Front. Comput. Neurosci.* **11**, 1–12 (2017).
[66] C. Ieracitano, N. Mammone, A. Hussain and F. C. Morabito, “A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia,” *Neural Netw.* **123**, 176–190 (2020).
[67] K. Venkatachalam, A. Devipriya, J. Manirajc, M. Sivaramd, A. Ambikapathye and I. S. Amiri, “A novel method of motor imagery classification using eeg signal,” *Artif. Intell. Med.* **103**, 101787 (2020).
[68] T. T. M. Delsy, N. M. Nandhitha and B. S. Rani, “Feasibility of spectral domain techniques for the classification of non-stationary signals,” *J. Ambient Intell. Hum. Comput.* (2020). doi:10.1007/s12652-020-02220-7.
[69] T. Tuncer, “A new stable non-linear textural feature extraction method based EEG signal classification method using substitution Box of the Hamsi hash function: Hamsi pattern,” *Appl. Acoust.* **172**, 107607 (2021).
[70] I. Guler and E. D. Ubeyli, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients,” *J. Neurosci. Methods* **148**(2), 113–121 (2005).
[71] P. Herman, G. Prasad, T. M. McGinnity and D. Coyle, “Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification,” *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**(4), 317–326 (2008).
[72] S. Sarkar, U. Taraphder, S. Datta, S. P. Swain and D. Saikhom, “Multivariate statistical data analysis-principal component analysis (PCA),” *Int. J. Livest. Res.* **7**(5), 60–78 (2017).
[73] S. Lemm, B. Blankertz, G. Curio and K. R. Müller, “Spatio-spectral filters for improving the classification of single trial EEG,” *IEEE Trans. Biomed. Eng.* **52**(9), 1541–1548 (2005).
[74] F. Li, F. He, F. Wang, D. Zhang, Y. Xia and X. Li, “A novel simplified convolutional neural network classification algorithm of motor imagery EEG signals based on deep learning,” *Appl. Sci.* **10**(5), 2–14 (2020).
[75] O. Ņupa, A Procházka, O Vyšata, M Schätz, J Mareš, M Vališ and V Mařík, “Motion tracking and gait feature estimation for recognising Parkinson’s disease using MS Kinect,” *Biomed. Eng. Online* **14**(1), 1–20 (2015).
[76] L. Dranca, LD de Mendarozketa, A Goñi, A Illarramendi, IN Gomez, MD Alvarado and MC Rodríguez-Oroz, “Using Kinect to classify Parkinson’s disease stages related to severity of gait impairment,” *BMC Bioinf.* **19**(1) (2018).
[77] H. Lee, L. Guan and I. Lee, “Video analysis of human gait and posture to determine neurological disorders,” *Eurasip J. Image Video Process.* **380867**, 1–12 (2008).
[78] N. M. Tahir and H. H. Manap, “Parkinsons disease gait classification based on machine learning approach,” *J. Appl. Sci.* **12**(2), 180–185 (2012).

[79] T. T. Verlekar, L. D. Soares and P. L. Correia, "Automatic classification of gait impairments using a markerless 2D video-based system," *mdpi.com*.

[80] T. Khan, J. Westin and M. Dougherty, "Motion cue analysis for parkinsonian gait recognition," *Open Biomed. Eng. J.* **7**(1), 1–8 (2013).

[81] S. Aich, P. M. Pradhan, J. Park and H. C. Kim, "A machine learning approach to distinguish Parkinson's disease (PD) patient's with shuffling gait from older adults based on gait signals using 3D motion analysis," *Int. J. Eng. Technol. (UAE)* **7**(3), 153–156 (2018).

[82] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J. Biomed. Health Inf.* **19**(6), 1794–1802 (2015).

[83] A. Kuhner, T. Schubert, C. Maurer and W. Burgard, "An online system for tracking the performance of Parkinson's patients," *IEEE Int. Conf. Intell. Robots Syst.* **2017**, 1664–1669 (2017).

[84] S. Soltaninejad, A. Rosales-Castellanos, F. Ba, M. A. Ibarra-Manzano and I. Cheng, "Body movement monitoring for Parkinson's disease patients using a smart sensor based non-invasive technique," *IEEE 20th Int. Conf. e-Health Netw. Appl. Serv. Healthcom* **2018**, 1–6 (2018).

[85] A. Procházka, O. Vyšata, M. Vališ, O. Upa, M. Schätz and V. Mařík, "Bayesian classification and analysis of gait disorders using image and depth sensors of Microsoft Kinect," *Digital Signal Process.: Rev. J.* **47**, 169–177 (2015).

[86] T. T. Verlekar, L. D. Soares and P. L. Correia, "Automatic classification of gait impairments using a markerless 2D video-based system," *Sensors* **18**(9), 1–16 (2018).

[87] M. Pistacchi, M Gioulis, F Sanson, E De Giovannini, G Filippi, F Rossetto and SZ Marsala, "Gait analysis and clinical correlations in early Parkinson's disease," *Funct. Neurol.* **32**(1), 28–34 (2017).

[88] S. Hwang, Y. Woo, S. Y. Lee, S. S. Shin and S. Jung, "Augmented feedback using visual cues for movement smoothness during gait performance of individuals with parkinson's disease," *J. Phys. Ther. Sci.* **24**(6), 553–556 (2012).

[89] H. H. Manap, N. M. Tahir and R. Abdullah, *"Parkinsonian gait motor impairment detection using decision tree," Proceedings - UKSim-AMSS 7th European Modelling Symposium on Computer Modelling and Simulation*, *EMS 2013* (2013), pp. 209–214.

[90] B. Shao, X. Li and G. Bian, "A survey of research hotspots and frontier trends of recommendation systems from the perspective of knowledge graph," *Expert Syst. Appl.* 113764 (2020).

[91] M. Ambrus, J. A. Sanchez and M. Fernandez-del-Olmo, "Walking on a treadmill improves the stride length-cadence relationship in individuals with Parkinson's disease," *Gait Posture* **68**, 136–140 (2019).

[92] S. Nõmm, A. Toomela, M. Vaske, D. Uvarov and P. Taba, "An alternative approach to distinguish movements of parkinson disease patients," *IFAC-PapersOnLine* **49**(19), 272–276 (2016).

[93] M. H. Li, T. A. Mestre, S. H. Fox and B. Taati, "Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation," *J. Neuro Eng. Rehabil.* **15**(1), 1–13 (2018).

[94] R. Sun, Z. Wang, K. E. Martens and S. Lewis, "Convolutional 3D attention network for video based freezing of gait recognition," *Int. Conf. Digital Image Comput.: Tech. Appl., DICTA 2018* **15**(97), 1–13 (2019).

[95] L. Dranca, LD de Mendarozketa, A Goñi, A Illarramendi, IN Gomez, MD Alvarado and MC Rodríguez-Oroz, "Using Kinect to classify Parkinson's disease stages related to severity of gait impairment," *BMC Bioinf.* **19**(1), 1–15 (2018).

[96] S. Aich, PM Pradhan, J Park, N Sethi, VS Vathsa and HC Kim, "A validation study of Freezing of Gait (FoG) detection and machine-learning-based fog prediction using estimated gait characteristics with a wearable accelerometer," *Sensors* **18**(3287), 3–16 (2018).

Cite this article: Pooja, SK. Pahuja and K. Veer (2022). "Recent Approaches on Classification and Feature Extraction of EEG Signal: A Review", Robotica **40**, 77–101. https://doi.org/10.1017/S0263574721000382