Carpathian Journal of Electronic and Computer Engineering 13/1 (2020) 23-29

DOI: 10.2478/cjece-2020-0004

A Review of Processing Methods and Classification Algorithm for EEG Signal

Yu Xie
IT Systems and Networks
Faculty of Informatics
University of Debrecen
Debrecen, Hungary
yu.xie@inf.unideb.hu

Stefan Oniga
IT Systems and Networks
Faculty of Informatics
University of Debrecen
Debrecen, Hungary
oniga.istvan@inf.unideb.hu

***Abstract*—The analysis technique of EEG signals is developing promptly with the evolution of Brain Computer-Interfaces science. The processing and classification algorithm of EEG signals includes three states: pre-processing, feature extraction and classification. The article discusses both conventional and recent processing techniques of EEG signals at the phases of preprocessing, feature extraction and classification. Finally, analyze popular research directions in the future.**

Keywords—EEG; signal processing; feature extraction; classification; deep learning.

I. INTRODUCTION

Human beings' perception of themselves is very superficial compared to other fields. Human beings can now create tools with various functions and continuously improve them, but they are helpless against the common defects of human beings themselves. The most complicated organ of the human body, the brain, is as mysterious and unpredictable as a black box. If the input and output channels of this black box fail, it is almost impossible to solve the problem internally. However, Brain-computer interface (BCI) technology is one of the technologies that can be expected to solve such problems[1]. The BCI device is used to directly convert the patient's intentional conscious behavior into data and control commands, send it to the computer and further drive the related equipment, so that the patient can easily communicate with the outside world and achieve a certain degree of self-care.

BCI technology is a cross-cutting field that combines cutting-edge scientific exploration and application, which has emerged in the 1970s and has only developed rapidly in the past ten years. Although various difficulties and problems have also been encountered, related research work is progressing truly quickly. Among them, EEG, as the principal technology of the BCI, is applied to study the best way of feature extraction. The application of new signal processing methods developed in recent years, such as independent component analysis , wavelet analysis theory, and nonlinear dimensionality reduction algorithms, to EEG feature extraction, is the application and verification of these theories, and it is also expected to deepen the level of theoretical understanding.

As an indispensable technology of the BCI, EEG can be divided into five stages in terms of its application method [2]. The first phase is the acquisition of EEG signals. The second stage is the pre-processing of EEG signals to remove noise interference. The original EEG signals include the noisy signals of eye blinks, heartbeat artifacts, and muscular movements. Removing interference signals can simplify the analysis and processing of subsequent EEG signals. The third stage is the feature extraction of EEG signals. Features are extracted from the pre-processed EEG signals to distinguish different EEG signals and scale down the extent of the vector dimension to prevent the calculational complicatedness simultaneously. The fourth stage is to classify the representative feature using the classification approach. Select a suitable classifier is certainly important to get effective classification results. The fifth stage is to convert the classification results to control external devices or give judgment. EEG signal pre-processing, feature extraction and classification are important in the EEG signal processing and have been extensively investigated[3][4]. This paper address various algorithm at These three stages. This paper also discussed the strengths and weaknesses of each stage. The trend of BCI is discussed at the end of this paper.

II. PRE-PROCESSING OF EEG SIGNALS

The original EEG signal contains noise for instance eye winks, hand movements, and heartbeat artifacts [5]. At the same time, electrical grounding noise is also an important source of EEG artifacts. Since EEG signals are extremely weak physiological signals (the amplitude is at the microvolt level) Therefore, EEG signals are extremely vulnerable and even annihilation to noise. These interference signals will increase the complexity of EEG signal processing and increase the amount of calculation in the rest of the processing. It needs to be stripped before analysis. Here are numerous pre-processing ideas particularly Common Spatial Patterns (CSP), Principle Component Analysis (PCA) [5], Common Average Referencing (CAR) [5][6], Surface Laplacian (SL), adaptive filtering [5][9], Independent Component Analysis (ICA) and digital filter [15].

A. Independent component analysis

ICA is a blind source analysis approach for decomposing a complex dataset into independent sub-parts basic on the statistical technique. According to the ICA algorithm theory, eye movement artifacts, ECG, EMG, and electrical grounding noise are all generated by independent signal sources, which are statistically independent. They can be separated by the

ICA algorithm to extract useful information on EEG signals. The design has been perceived as an impressive mechanism for constrict artifacts and determine mathematical independent cortical measure in the scalp and intracranial EEG reporting[10] meanwhile artifact rejection discards corrupted data tranches, thereby conserving the original bulk of samples, giving in a superior signal-to-noise ratio for consecutive search steps. ICA performs better than PCA when the volume of the data to disintegrate is giant [5].

B. Principal component analysis

PCA is one of the most meaningful dimensionality reduction technology by using the statistical method. It makes usage of scientific procedures that apply an orthogonal transformation to alter a series of measurements of interacted vectors into a series of linearly disrelated vectors called principal components [10]. The dominance of PCA is that it takes care of the reduction of the dimension of the factor vectors and the complexity of signal feature extraction and classification. In the application of EEG signal processing, PCA disintegrates EEG signals into disrelated components. The principal components have the largest variance. And then the EEG signal is reconstructed after separating inference components. Non-principal components with small variance may also contain important information on sample differences, and discarding due to dimensionality reduction may affect subsequent data processing.[7]

C. Common average reference

CAR method is a spatial filtering algorithm that eliminates the common part of the EEG signal of the EEG while retaining the characteristic signals of specific electrodes to improve the Noise Ratio (SNR) [14]. This algorithm can be seen as follows:

$$V_i^{CAR} = V_i^{ER} - \frac{1}{n}\sum_j^n V_j^{ER} \tag{1}$$

Where V_i^{CAR} is the Filtered voltage value for a specific channel, V_i^{ER} is the original voltage value of a specific channel (The original voltage value is the voltage difference between the channel electrode and the reference electrode).$\frac{1}{n}\sum_j^n V_j^{ER}$ is the Average voltage of all channels[8].

D. Common Spatial patterns

CSP filtering is derived from Common Spatial Subspace Decomposition (CSSD). CSSD is a spatial filtering algorithm for multichannel EEG data. The primary idea of the CSSD method is to find a certain direction in a high-dimensional space. When classifying the two cases, the variance of one type is the largest and the variance of the other type is the smallest. Its role is to extract task-related signal components and suppress task-irrelevant components and noise[11]. CSPs are mostly used to process brain imagined EEG data based on the BCI. The basic idea is to design a spatial filter to process the EEG signals to find a series of excellent spatial filters for projection using the diagonalization of the matrix, so as to maximize the difference between the variance values of the two types of signals and obtain a feature vector with higher discrimination. The advantage of this algorithm is that it does not need to select a specific frequency band in advance. However, it is sensitive to noise and demands to be used with various electrodes.

E. Surface Laplacian

SL commonly is the efficient spatial filter based concept in the BCI area. This technique is described as the 2nd order spatial derivative of the surface potential[12]. SL determines the electrode reference problem. It is also robust against artifacts caused at uncovered part by the electrode cap and more is an approach of exploring the EEG signal with high spatial resolution. Nonetheless, the issue is that it is too delicate to the election of spline parameters during spline interpolation [5].

F. Adaptive filter

An adaptive filter is a technique that can automatically adjust parameters without the statistical characteristics of the input signal and noise in advance. During the work process, the required statistical characteristics are gradually estimated to adjust its parameters to achieve the best filtering effect. A complete adaptive filter consists of four components: the signals being handled, the structure that describes the input or output connection, the parameter, which can be iteratively differed to change the filter's input or output relationship.

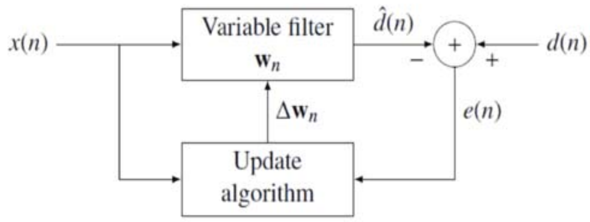

Fig. 1. Example of adaptive filters

Fig.1 is the basis for the realization of these special adaptive filters, such as the least mean square filter (LMS) and recursive least squares filter (RLS).

The input signal is the sum of the desired signal $d(n)$ and the interference noise $v(n)$

$$x(n) = d(n) + v(n) \tag{2}$$

The variable filter has a finite impulse response structure. the impulse response of the structure is equal to the filter coefficient. The coefficients of the p-order filter are defined as

$$w_n = [w_0(0), w_n(1), \dots, w_n(p)]^T \tag{3}$$

The error signal or cost function $e(n)$ is the difference between the desired signal $d(n)$ and Estimated signal $\hat{d}(n)$

$$e(n) = d(n) - \hat{d}(n) \tag{4}$$

In [13] the authors designed a new type of adaptive filter called adaptive Laplace filter. This filter can improve the performance of the motor-computer-based BCI, and has confirmed the accuracy and robustness of the motor-computer-based BCI in a study of 22 subjects.

G. Digital filters

EEG signals are random, non-stationary, and non-linear. Digital filters process EEG signals from the frequency domain and can be divided into low-pass filters, high-pass filters, band-pass filters, and band-stop filters. Digital filters are broadly utilized in artifact processing of EEG signals, specifically in filtering EMG artifacts. However, it is required that EEG signals and artifact signals have different frequency bands, which are difficult to exist in practical situations. So, the use of digital filters is limited. Band-pass filter and notch Filter are frequently utilized in EEG signal processing[16]. We must fully consider which filter to use before performing EEG filtering. Otherwise, we will not get the desired result. For example, Notch filters can be used to filter out 50Hz electrical grounding noise; both FIR and IIR filters will cause a certain delay and this delay will affect the phase relationship in the EEG signal. We know that the phase information of the EEG signal is highly important. Both P300 and SSVEP have a relationship of time-locked and phase-locked in the BCI. EEGLAB is an interactive MATLAB toolbox for processing continuous and event-related EEG, MEG and other electrophysiological data. FIR or IIR filter used in EEGLAB is zero-phase filtering, which is a filter without phase offset. The specific principle can be explained in Google MATLAB 'filtfilt' function.

The EEG signal preprocessing mainly removes the eye movement artifacts, ECG artifacts, electrical grounding noise in the EEG signal and provides "clean" signal data for the subsequent EEG signal feature extraction. The above-mentioned CSP, CAR, PCA, and ICA belong to the spatial domain filter, as well as the time domain filter, such as Fourier analysis and autoregressive analysis. Temporal filters not as useful as spatial filters at the same data volume. Table I. displays the comparison of the pre-processing methods.

TABLE I. COMPARISON OF PRE-PROCESSING APPROACHES

No	Approach	Advantages	Disadvantage
1	ICA	• High calculation efficiency • Works well for large data sized	• Require more computational for decomposition
2	PCA	• Reduction of feature dimension	• Not as good as ICA
3	CAR	• The best among all the reference methods	• Needs sufficient head coverage
4	CSP	• Good effect on EEG data processing based on motor imagination • does not need to select a specific frequency band in advance.	• sensitive to noise • requires multiple electrodes
5	SL	• Robust against the artifacts	• Sensitive to spline pattern • Sensitive to artifacts
6	Adaptive Filtering	• Good for the signals with overlapping spectra	• Needs reference signal
7	Digital Filter	• Easily remove electrical grounding noise	• Needs signal and noise reside distinct frequency bands

III. FEATURE EXTRACTION

We get relatively pure EEG signals from the previous step. However, due to a large amount of EEG signal data, direct processing is too complicated. A compelling array of features are obtained using contain PSD, PCA, ICA, Auto-Regressive (AR), Wavelet Transform (WT), Wavelet Packet Transform (WPT), Fast Fourier Transform (FFT), etc.

A. Power Spectrum Density

PSD is a great mechanism for immobile signal transform and advisable for the narrowband signal. The PSD represents a measure of the power as a function of frequency in a given signal. In [17] the authors propose an effective feature extraction method WDPSD (feature extraction from the Weighted Difference of Power Spectral Density in an optimal channel couple) that can display the time, frequency and spatial characteristics for 2-class motor imagery-based BCI system.

B. Principal Component Analysis

PCA is not only used for EEG data preprocessing but can also be used for feature extraction of data. PCA extracts useful signals from the time series data of EEG signals and removes artifacts to achieve the purpose of dimensionality reduction. BCI P300 speller can be applied as a ground-breaking help for seriously incapacitated individuals in their regular day to day existence. In [18] the authors used PCA and an ensemble of weighted SVM (EWSVM) for classification on the P300 character spelling BCI system. PCA is used to weaken the unnecessary features an ensemble of the weighted classifier for minimizing the classifier adaptability. The effectiveness of classification can be enhanced related to other feature extraction methods by using PCA feature extraction[14]. However, it is unable to process the data if data is linear and continuous and for an intricated set of features[5].

C. Independent Component Analysis

ICA, the same as PCA, can be used for both signal pre-processing and signal feature extraction. ICA decomposes signals into independent components and extracts useful components. As a "blind source" separation technology, ICA helps to identify independent signals and separate noise. The benefits of ICA are fast calculation speed and high efficiency. It is suitable for processing massive sizable signals such as EEG signals. Associating EEG with eye-tracking is an auspicious way to research Neuro-related of native perception. In [19] were appraised and enhanced ICA-based correction for visual search in images and sentence learning with free eye movements. With the developed operation, ICA eliminated nearly all artifacts, covering the spike potentials and its associated spectral broadband artifact from both viewing paradigms, with little deformity of neural movements.

D. Auto-Regressive

AR approach is a time-domain analysis method, which has been used in a number of researches to establish the EEG signal to achieve feature extraction. AR model can be expressed as a linear prediction problem. For time-series data, the predicted value of the current point can be approximated by the linear association of the data at prior marks[20]. It is the most commonly used method for non-stationary data where AR arguments are provided to the model. The expense exists in building the AR parameter model capability [5]. AR models frequently used in EEG signal analysis can be further divided into adaptive models and non-adaptive models.

E. Fast Fourier transform

FFT is a quick method of discrete Fourier transform. In the feature extraction of EEG signals, FFT transforms the EEG signal from the time domain to the frequency domain and performs spectrum analysis or calculates power spectral density[21]. It has a good effect on settled data and in linear irregular systems. Yet, It does not work well for unstable data and cannot calculate both at the time and frequency domain. In [22] Studied the operation control of an electric wheelchair with a BCI. The subject's EEG data was band-pass filtered and FFT to obtain alpha, beta, theta, and delta waves. And then input all feature quantities to the classifier for classification and recognition after calculating standard deviation and entropy value. FFT is also used in the study of driver's EEG signals in fatigue driving and auto driving simulation[23][24].

F. Wavelet transform

WT is a time-frequency transform technique inheriting and developing the idea of localization of short-time Fourier transform. It is applied for extracting data from numerous categories of continued signals such as speech and image data. The WT highlights the characteristics of the signal, and the signal is multi-scaled and refined by the telescopic translation operation, so as to improve the time resolution at high frequencies and the frequency resolution at low frequencies, and automatically adapt to the signal time-frequency analysis requirements. In the WT decomposition process, only the low-frequency part of the signal is re-decomposed. The high-frequency part is no longer decomposed. Wavelet package transform (WPT) can give a more accurate frequency resolution distinguished to WD in the high-frequency part [25]. Compared with the Fourier transform, the disadvantage of the WT is that the wavelet base function is not unique. As a result, one of the difficulties in applying wavelet analysis to practice is the choice of the best wavelet base function.

EEG signal feature extraction is an essential step for the classification and recognition of EEG signals. Therefore, it is not sufficient to utilize elementary ways such as a band stop filter to abstract the useful data since helpful data of brain signals are invisible in a badly noisy situation and information of benefit could be overlapped in time and space due to coinstantaneous movement of the brain. High-dimensional feature vectors will bring incredibly complicated calculations to subsequent classification step. Dimension reduction is required. Generally, PCA or ICA is used to reduce dimensions. The comparison of the feature extraction methods has displayed in Table II.

IV. CLASSIFICATION

TABLE II. THE EXTRACTED VECTORS ARE CLASSIFIED INTO NUMEROUS CLASSIFIERS TO REALIZE THE ANALYSIS AND PREDICTION OF THE EEG SIGNALS AFTER FEATURES HAVE BEEN EXTRACTED. THE MOST FAMOUS CLASSIFIER INCLUDES K-NEAREST NEIGHBOR (K-NN), LINEAR DISCRIMINANT ANALYSIS (LDA), SUPPORT VECTOR MACHINE (SVM), AND NAIVE BAYES (NAIVE BAYES, NB), ARTIFICIAL NEURAL NETWORK (ANN)[26] AND DEEP LEARNING (DL)[27].COMPARISON OF FEATURE EXTRACTION APPROACHES

No	Approach	Advantages	Disadvantages
1	PSD	• Feature stability	• Not suitable for unstable signals • Unable to analyze time-domain signals
2	PCA	• Lossless dimension reduction	• Weak to handle complex set of data
3	ICA	• fast calculation speed • high efficiency	• Needs more calculating for decomposition
4	AR	• Needs small duration of data • Decrease spectra loss issues • Gives well frequency resolution	• Not available for stationary signals • Presents low performance once the estimated model is not suitable and its orders are mistakenly elected
5	FFT	• Applicable for stationary signals • It is appropriate for narrowband signals, such as sine wave	• Not works for nonstationary signals • Sustains greatly noise sensitivity and does not have shorter duration data record
6	WT	• Works well for nonstationary signals • Can analyze signal both in time and frequency domain	• Deficient particular mode to apply to prevalent noise
7	WPT	• Works well to analyze the nonstationary signals • provide precise frequency resolution	• Raise calculating time

A. Artificial Neural Network

Artificial Neural Network(ANN) has been a hot exploration that has emerged in the area of AI from the 1980s. Human neuron networks are abstracted and established corresponding models, and formed various connection according to particular network approaches by ANN. It is a nonlinear classifier composed of a large collection of neurons. It is mainly used to process massive data at the same time to solve classification and regression problems and is a branch of machine learning methods. Multi-Layer Perceptron Neural Network (MLPNN) is the most extensively applied neural network framework consisting of three layers: the input layer, hidden layer(s) and output layer[1][5]. The asset of MLPNN is that its rapid process, ease of completion and thin training sets required. It, however, hard to establish and the capability almost counts on the number of neurons in the hidden layer[5].

ANN is widely used in the area of medical diagnosis, specifically in the detection and analysis of biomedical signals. It can be used to solve problems that are difficult or impossible to solve by conventional methods in biomedical signal processing. It gains "experience" through training on a priori data without paying too much attention to the details and characteristics of the disease[28][29]. In [28] introduced the trained and learned ANN for the diagnosis of tuberculosis with a sensitivity and specificity of 100% and 72% respectively. In [30] the authors used the ANN-based MS-ROM / I-FAST system to extract desired features from EEG to achieve the differential diagnosis of children with autism and achieve good results. The required EEG is only a few minutes of data and does not require any data preprocessing.

B. Deep Learning

Deep learning is a kind of machine learning and an extension of ANN. The meaning of DL refers to the learning, analysis, and processing of a huge quantity of hidden layer neurons contained in ANN. It has many hidden layers, a huge quantity of neurons, and many parameters to be adjusted. In the latest years, not only the application of DL but also many basic disciplines in the field of AI, including machine learning and neural networks have been extensively and deeply studied in the processing and recognition of EEG signals. The research content of EEG classification using DL mainly includes: emotion recognition, motor image recognition, psychological load detection, epilepsy detection, and event correlation Potential detection and sleep scoring. The deep neural networks used mainly contain convolutional neural networks, recurrent neural networks, deep belief networks, and also include stack autoencoders and multilayer perceptron [31].

The current research on basic theories is not optimistic. One is that no more innovative algorithms have appeared in the discipline, and more researches are optimization and change. On the other hand, Neural networks are still a kind of black box for humans. Not knowing how to explain the cause is a serious problem. A convolutional neural network(CNN) is a type of forward neural network. Its structure consists of the following types of layers: the input layer, convolutional layer(s), pooling layer(s), fully connected layer, and output layer. The convolutional layer does feature extraction, the pooling layer reduces network training parameters and reduces the degree of model overfitting. All neurons in the fully connected layer are connected to the last pooling layer. The output layer contains two neurons for binary classification output. The recurrent neural network is used to analyze sequence data. The neuron's input in the structure of the recurrent neural network is affected by the input neuron and the output of previous node neurons.

The most commonly used recurrent neural network is Long Short-Term Memory (LSTM) neural network. The basic structure of the LSTM is a memory module with three gates: input gate, output gate, and forget gates. The three gates determine the input, output, and removal of information. The number of hidden layers in the LSTM and the number of neurons in the hidden layer can be manually adjusted to optimize the prediction performance of the trained LSTM model. In [31] the authors believe that convolutional neural networks, recurrent neural networks, and deep belief networks are superior to stack autoencoders and multilayer perceptron in EEG signal classification accuracy after Summarized 90 studies on deep learning-based EEG signals.

C. *K-Nearest Neighbors*

KNN method is a classification algorithm that is the simplest and most understandable machine learning algorithm, proposed by Cover and Hart in 1968. Over the years, the K-NN algorithm has been improved a lot and already broadly used in the fields of face recognition, text recognition, medical image processing[32][33]. Its core idea is that if most of the k most neighbor samples in a feature space remain with a certain class, then the sample also belongs to that class and has the features of the samples of that class. The K-NN method first determines a training sample set. All sample categories in the sample set are known. For the sample to be classified, find the similarity between the sample and the training sample set. Select the k samples with the highest similarity. The classes of the sample to be classified are determined by the classes of k samples. K-NN is a kind of instance-based learning, or lazy learning, where the data set has classification and eigenvalues in advance and will be processed directly after receiving new samples. The expense of KNN is excessive calculating complicatedness. The computational complexity is proportional to the volume of data in the data set. Therefore, KNN is generally suitable for data sets with a small volume of samples.

D. *Linear Discriminant Analysis*

LDA is a linear learning method proposed by Fisher in 1936. The main idea of LDA is: find the proper projection direction to project the samples onto a straight line so that the projection points of the same type are as concentrated as possible, and the projection points of different types are as far as possible. Then use the same method to classify the new sample, and determine the type of the new sample based on the position of the projection point of the new sample on the straight line. LDA has low computational complexity and is easy to use. It provides a good outcome. LDA may not sustain the complicated construction in the data of non-Gaussian distributions. LDA will fail if the discriminatory function is not in the mean but the variance of the data [3]. LDA is usually divided the sample set into two categories when used for pattern recognition[34].

E. *Support Vector Machine*

SVM is a type of widespread linear classifier that achieves binary classification of data according to supervised learning[35]. The basic principle is to find the optimal decision surface in space so that different types of data can be distributed on both sides of the decision surface to achieve classification. SVM can be divided into linear separable SVM, linear SVM, and nonlinear SVM according to its construction model. In [36] the author uses the SVM classification method to classify the degree of driving fatigue. It is concluded that the key node to distinguish between mild fatigue and severe fatigue is that the continuous driving time is at least 2 hours, and the algorithm has a better recognition effect for severe fatigue. In [37] the authors optimized the epilepsy EEG data and used PCA for feature extraction. On this basis, Naive Bayes, K nearest neighbor algorithm, linear discriminant analysis, and least squares SVM was used to classify the feature vectors. The classification accuracy rate of LS-SVM is up to 100%, which is 7.10% higher than the accuracy of existing epilepsy EEG data classification algorithms.

F. *Naive Bayes*

NB classifier is an uncomplicated and practical classifier based on Bayes' theorem. In some fields, its efficiency is comparable to the efficiency of other classifiers [38][39][40]. The main idea of NB is: for a given item to be classified, solve the probability of each category appearing under the condition that this item appears. The item to be classified belongs to the category with the supreme possibility. The NB algorithm considers that the samples are independent and uncorrelated [41]. NB classifiers have the outstanding characteristics of fast speed, high efficiency and simple algorithm structure with processing high-dimensional data [42]. Based on the NB algorithm, researchers have proposed various improved algorithms, such as tree-enhanced NB algorithm and network-enhanced NB algorithm, to improve algorithm performance and classification accuracy.

At present, none of the algorithms has absolute advantages. Only by selecting the appropriate classification algorithm for specific sample data can the best classification effect be achieved. There are two factors to consider when choosing a classifier: data dimension and bias and variance balance. The amount of training data required increases geometrically with the dimension of the eigenvector. The classifier tends to have high deviation and low variance. The comprehensive use of various classifier algorithms is the development trend of classifier algorithms[3]. The comparison of the classification approaches is shown in Table III.

TABLE III. COMPARISON OF CLASSIFICATION APPROACHES

No	Approach	Advantages	Disadvantages
1	ANN and DL	● High accuracy ● flexible structure	● Performance depends on the number of neurons in the hidden layer
2	K-NN	● Easy to understand, ● simple to implement	● Sensitive to irrelevant and redundant features ● requires large storage space
3	LDA	● Easy to use ● Low computational complexity	● Requires a linear model
4	SVM	● Performance is better compared to other linear classifiers	● Low computational complexity
5	NB	● Easy to understand	● Independent variables

V. DISCUSS

The BCI establishes a direct connection path between external devices and the human brain. EEGs are broadly researched and implemented as a non-invasive method for detecting physiological signals. The BCI system relies on EEG to judge the instructions issued by the human brain. In this process, the acquisition, data processing and pattern classification of the EEG signals determine the performance of the BCI system. EEG data processing includes data preprocessing and feature extraction. Data processing affects the speed and accuracy of subsequent pattern classification. The choice of classifier determines the result of pattern classification.

EEG signals are complex non-stationary irregularity signals. The analysis and processing of EEG data is the key technology of the BCI system. It is the focus and difficulty of research and determines the development prospect of the BCI system. In recent years, methods for pre-processing, feature extraction, and classification and recognition of EEG data are extensively and deeply studied. A variety of EEG data processing methods have been applied to BCI systems.

The choice of data processing method in specific practice is mainly determined by the research subject. The data preprocessing, feature extraction methods and classification approaches used are determined according to the characteristics and amount of EEG data required for the research.

From the perspective of pattern recognition, the validity of the original input information and accurate interpretation of the original data determine the overall performance of the system. The signal source of the EEG signal system is the EEG signal monitored by non-invasive scalp electrodes. Due to the constraints of acquisition methods, the low spatial resolution, instability, large individual differences and susceptibility to the interference of EEG signals make it extremely difficult to accurately read EEG information.

BCI researchers have carried out a lot of research work on EEG signal processing and pattern recognition methods, scalp electrode distribution and optimization of the number of leads. However, the lag in the precise acquisition and interpretation of neural activity information based on EEG is still the biggest obstacle restricting the development of EEG.

VI. CONCLUSION

This paper has Presented the varieties of techniques implemented for EEG signals at preprocessing, feature extraction and classification stages. Deep learning is a branch of artificial neural networks, a new field of machine learning, and a current research hotspot in the processing and classification of EEG data. With the collection of a huge quantity of clinical electrophysiological data, deep learning will be more widely studied and applied.

REFERENCES

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain-computer interfaces, a review," Sensors, vol. 12, pp. 1211-1279, 2012.

[2] Michael X Cohen. Where Does EEG Come From and What Does It Mean? [J]. Trends in Neurosciences, 2017(40):208-218.

[3] Mohd Zaizu Ilyas, Puteh Saad & Muhammad Imran Ahmad. A Survey of Analysis and Classification of EEG Signals for Brain-Computer Interfaces[J]. 2015 2nd International Conference on Biomedical Engineering(ICoBE), 30-31 March 2015, Penang.

[4] Namita.R.Tambe, Ajitkumar Khachane. Mood Based E-learning using EEG[J]. 2nd International Conference on Computing, Communication, Control and Automation,12-13 AUG 2016.

[5] M. R. Lakshmi, D. T. V. Prasad, and D. V. C. Prakash, "Survey on EEG Signal Processing Methods," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, pp. 84-91, 2014.

[6] Mohammed J. Alhaddad, "Common Average Reference (CAR) Improves P300 Speller," International Journal of Engineering and Technology, vol. 2, pp. 451-489, 2012.

[7] Tiejun Liu, Dezhong Yao.Removal of the ocular artifacts from EEG data using a cascaded Spatio-temporal processing[J]. Computer methods and programs in biomedicine, 83 (2006):95–103.

[8] Luis Alfredo Moctezuma, Alejandro A. Torres-García, et al.Subjects identification using EEG-recorded imagined speech[J]. Expert Systems With Applications,118 (2019):201–208.

[9] M. K. Ahirwal, A. Kumar, and G. K. Singh, "Adaptive filtering of EEG/ERP through Bounded Range Artificial Bee Colony (BR-ABC) algorithm," Digital Signal Processing, vol. 25, pp. 164-172, 2// 2014.

[10] G. Korats, S. Le Cam, R. Ranta, and M. Hamid, "Applying ICA in EEG: Choice of the Window Length and of the Decorrelation Method," in Biomedical Engineering Systems and Technologies. vol. 357, J. Gabriel, J. Schier, S. Van Huffel, E. Conchon, C. Correia, A. Fred, et al., Eds., ed: Springer Berlin Heidelberg, 2013, pp. 269-286.

[11] L. Velásquez-Martínez, A. M. Álvarez-Meza, and C. G. Castellanos-Domínguez, "Motor Imagery Classification for BCI Using Common Spatial Patterns and Feature Relevance Analysis," in Natural and Artificial Computation in Engineering and Medical Applications. vol. 7931, J. Ferrández Vicente, J. Álvarez Sánchez, F. de la Paz López, and F. J. Toledo Moreo, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 365-374.

[12] D. J. McFarland, "The advantages of the surface Laplacian in brain-computer interface research," International Journal of Psychophysiology, 2014.

[13] Jun Lu, Dennis J. McFarland, Jonathan R. Wolpaw. Adaptive Laplacian filtering for sensorimotor rhythm-based brain-computer interfaces[J]. Journal of neural engineering,2013(10),p016002.

[14] X. Yu, P. Chum, and K.-B. Sim, "Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system," Optik - International Journal for Light and Electron Optics, vol. 125, pp. 1498-1502, 2// 2014.

[15] J. Suto, S. Oniga, Music Stimuli Recognition in Electroencephalogram Signal. Elektronika ir Elektrotechnika, 24(4), 2018, pp.68-71. http://dx.doi.org/10.5755/j01.eie.24.4.21482.

[16] Shayan Motamedi-Fakhr, Mohamed Moshrefi-Torbati, Martyn Hill, Catherine M. Hill, Paul R.White.Signal processing techniques applied to human sleep EEG signals-A review[J]. Biomedical Signal Processing and Control, 10 (2014):21–33.

[17] Chungsong Kim, Jinwei Sun and Dan Liu, "An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI ", Medical & Biological Engineering & Computing volume 56, pages1645–1658(2018)

[18] Sourav Kundu, Samit Ari.P300 Detection with Brain-Computer Interface Application using PCA and Ensemble of Weighted SVMs[J]. IETE Journal of Research, 2017.

[19] Dimigen, O., Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments, NeuroImage (2019), DOI: https://doi.org/10.1016/j.neuroimage.2019.116117

[20] V. Lawhern, W. D. Hairston, K. McDowell, M. Westerfield, and K. Robbins, "Detection and classification of subject-generated artifacts in EEG signals using autoregressive models," Journal of Neuroscience Methods, vol. 208, pp. 181-189, 7/15/ 2012.

[21] 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 Neuroscience, vol. 2014, p. 7, 2014.

[22] Md Fahim Ansari, Damodar Reddy Edla.Brain-Computer Interface for wheelchair control operations: An approach based on Fast Fourier Transform and On-Line Sequential Extreme Learning Machine[J]. Clinical Epidemiology and Global Health,2018.

[23] Mejdi Ben Dkhil, Ali Wali, and Adel M. Alimi.Drowsy Driver Detection by EEG Analysis Using Fast Fourier Transform[J].Electrical Engineering and Systems Science,2018.

[24] Liu Yang, Rui Ma.Driving behavior recognition using EEG data from a simulated car-following Experiment[J].Accident Analysis and Prevention,116 (2018):30–40.

[25] Damodar Reddy Edla, Md Fahim Ansari, et al. Classification of Facial Expressions from EEG signals using Wavelet Packet Transform and SVM for Wheelchair Control Operations[J]. Procedia Computer Science, 132 (2018):1467–1476.

[26] J. Suto, S. Oniga, Efficiency investigation of artificial neural networks in human activity recognition, Journal of Ambient Intelligence and Humanized Computing 9 (4), 1049-1060, August 2018. DOI:https://doi.org/10.1007/s12652-017-0513-5. WOS:000440310900012.

[27] J. Suto, S. Oniga, C. Lung, I. Orha, Comparison of offline and real-time human activity recognition results using machine learning techniques, Neural Computing and Applications, March 2018. https://doi.org/10.1007/s00521-018-3437-x.

[28] Payal Dande, Purva Samant.acquaintance to Artificial Neural Networks and use of artificial intelligence as a diagnostic tool for tuberculosis: A review[J]. Tuberculosis, 108 (2018):1-9.

[29] Errikos.Ventouras, Efstratia A. Monoyiou, Periklis Y. Ktonas, Thomas Paparrigopoulos, Dimitris G.Dikeos, Nikos K.Uzunoglu, Constantin R.Soldatos. Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: A feasibility study[J]. Computer Methods and Programs in Biomedicine,78 (2005):191-207.

[30] Enzo Grossi, Chiara Olivieri, Massimo Buscema. Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study[J].Computer Methods and Programs in Biomedicine,142 (2017):73–79.

[31] Alexander Craik, Yongtian He, Jose L. Contreras-Vidal.Deep Learning for Electroencephalogram (EEG) Classification Tasks: A Review[J]. Journal of Neural Engineering,2019.

[32] Pascal Soucy, Guy W.Mineau. A Simple KNN Algorithm for Text Categorization[J].Proceedings of the IEEE International Conference on Data Mining,2001(2):647-648.

[33] K.Murphy, B.van Ginneken, A.M.R.Schilham, B.J.de Hoop, H.A. Gietema, M.Prokop.A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbor classification[J].Medical Image Analysis, 13 (2009):757–770.

[34] L.F.Nicolas-Alonso and J.Gomez-Gil. Brain computer interfaces, a review.Sensor,2012(12):1211-1279.

[35] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.

[36] Yun Li.Recognition Algorithm of Driving Fatigue Related Problems Based on EEGSignals[J].NeuroQuantology,2018,16(6):517-523.

[37] Siuly Siuly, Yan Li. Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification[J]. Computer methods and programs in biomedicine,119 (2015):29–42.

[38] Farajollah Tahernezhad-Javazm, Vahid Azimirad and Maryam Shoaran.A review and experimental study on the application of classifiers and evolutionary algorithms in EEG-based brain-machine interface systems[J]. Journal of Neural Engineering,15(2018):021007 (39pp).

[39] Juliano Machado, Alexandre Balbinot. Executed Movement Using EEG Signals through a Naive Bayes Classifier[J]. Micromachines, 2014(5):1082-1105.

[40] Raja Majid Mehmood, Ruoyu Du, Hyo Jong LEE.Optimal Feature Selection and Deep Learning Ensembles Method for Emotion Recognition From Human Brain EEG Sensors[J]. Digital Object Identifier 10.1109/ACCESS,2017,2724555.

[41] DR.Mohammad A Obeidat, DR.Ayman M.Mansour.EEG Based Epilepsy Diagnosis System using Reconstruction Phase Space and Naive Bayes Classifier[J]. Wseas Transactions on Circuits and Systems,2018(17):159-168.

[42] Mr.Vijay D.Katkar, Mr.Siddhant Vijay Kulkarni.A novel parallel implementation of Naive Bayesian classifier for big data[J]. 2013 International Conference on Green Computing, Communication and Conservation of Energy (IEEE),p847–852.