Review

A Review of the Role of Machine Learning Techniques towards Brain–Computer Interface Applications

Saim Rasheed

Department of Information Technology, Faculty of Computing and Information Technology,
King Abdulaziz University, Jeddah 23443, Saudi Arabia; srahmed@kau.edu.sa

Abstract: This review article provides a deep insight into the Brain–Computer Interface (BCI) and
the application of Machine Learning (ML) technology in BCIs. It investigates the various types of
research undertaken in this realm and discusses the role played by ML in performing different BCI
tasks. It also reviews the ML methods used for mental state detection, mental task categorization,
emotion classification, electroencephalogram (EEG) signal classification, event-related potential
(ERP) signal classification, motor imagery categorization, and limb movement classification. This
work explores the various methods employed in BCI mechanisms for feature extraction, selection,
and classification and provides a comparative study of reviewed methods. This paper assists the
readers to gain information regarding the developments made in BCI and ML domains and future
improvements needed for improving and designing better BCI applications.

Keywords: brain–computer interface; BCI; EEG signals; emotion state; ERP signals; ML; mental state;
motor imagery

1. Introduction

In the last few years, vast developments have occurred in automated control and monitoring applications. These developments have promoted the discovery of advanced technologies such as Brain–Computer Interfaces (BCIs). BCIs provide a methodology for manipulating computers and devices to operate based on signals and thoughts generated. It is a direct link between the human brain and computers. BCI applications generally record the human brain signals and transmit them to a machine that can execute the anticipated task. BCI is a breakthrough innovation in the domain of brain mapping. It generally interprets the message of neurons and employs it for executing the tasks [1]. This direct bridge between the machine and the brain has numerous applications in the medication domain for physically disabled or locked-in people. It is not less than a boon for locked-in people whose brains are paralyzed due to physical disorders. It also possesses the ability to assist people beyond medical fields, such as gaming, entertainment, and experimentation with learning. BCI connects the human brain with peripheral devices by creating a direct interacting link between the outer world and the brain and creating a bi-directional communication interface between the outer environment and the brain. They provide a muscle-free medium for conveying persons' purposes of actions to outer/external devices such as computers, neural prostheses, and other assistive appliances. Unlike the classical input devices such as keyboard, pen, and mouse, the BCI reads the signals generated from the human brain at distinct locations, translates them into actions and commands through which computer(s) can be controlled for executing desired control/monitoring tasks [2].

A comprehensive BCI system normally includes four components: collecting and intensifying brain signals, algorithm identifying and categorizing the brain signals, algorithm transmitting the detected commands to the controlling equipment and the tool/device transmitting back the feedback. The brain produces signals for controlling the user's intentions and the BCI system translates such signals into manageable output commands

for controlling external devices. In brain–machine fusion mechanisms, the BCI is a vital element of information exchange. In BCI, there is no need to transmit signals via peripheral muscles and nerves, instead, EEG signals can be used directly for establishing the link between peripheral devices and the brain. In general terms, BCI allows people to interact with the outer world or surroundings through utilizing control signals produced from EEG activity, without involving peripheral muscles and nerves. The user's intentions are reflected employing EEG waves, which are later transformed into the desired output forms. A standard BCI system comprising a signal acquisition unit, signal processing unit, controlling unit, and application & feedback unit is depicted in Figure 1 [1,3–6]. The signal processing unit further includes three parts, namely, preprocessing, feature extraction, and feature classification. In the signal acquisition phase, brain signals are captured. In the preprocessing phase, noise and other artifacts are removed and signals are prepared in a form suitable for processing. In the feature extraction phase, discriminative data in the recorded brain signals are identified. After measuring, the signal is then mapped to a vector comprising discriminant and effective features from the signals observed. In the classification phase, signals with feature vectors are classified. The selection of better discriminative features is essential for achieving good pattern recognition and deciphering the user's intentions. The signals thus classified are translated into useful instructions for a connected device such as a computer or any other assistive appliance. BCIs can be independent BCI and dependent BCI. In an independent BCI, the ordinary output paths of peripheral muscles and nerves do not possess any role. These BCIs are useful for people suffering an acute neuromuscular disability. In the case of dependent BCI, the brain's ordinary paths are not utilized for carrying messages instead the brain activity is employed for generating EEG.
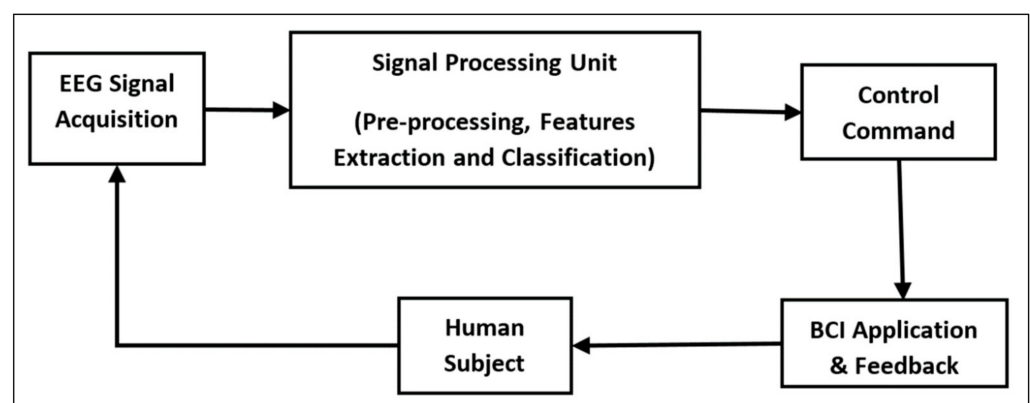

Figure 1. Standard BCI system.

Research in the BCI domain started late back in the 1970's. The foremost researches were conducted on monkeys and rat cortices at the University of Washington, School of Medicine, making an effort for imitating their neural actions. Until this phase, BCI research was constrained to only the medical diagnosis and examining the brain function. However, with technological advances, numerous developments have occurred in the BCI area over time. Now the scope of BCI has been extended to multiple domains including artificial intelligence (AI) which can provide a new way of recording human brain wave activity for understanding the improved functional operation of human brain activity, emotion recognition, human limb/hand movement detection, etc. The evolution and availability of cost-effective microelectronics will facilitate BCI users to execute complex tasks via embedded circuits. Also, the latest surge in self-decision making and improved algorithms such as Machine Learning (ML) may aid to explore more and surpass the bounds of brain-controlled applications. This has motivated the researchers to employ ML in BCI systems for executing its tasks in a more precise and useful manner.

ML is one of the crucial branches of statistics and computer science, which exploits computers to perform tasks without being programmed explicitly. ML studies computer architectures and algorithms, which are learned from generally observed facts. Typically, ML techniques are considered to be one of the important methods for implementing AI, which develops hardware, software, machines, and devices with human-type intelligence. It is a powerful tool exploited in diverse BCI applications. With the advances in modern science, the ML domain has made rapid progress and is widely exploited in monitoring, detection, classification, and other tasks. Moreover, these algorithms have become imperative for accomplishing huge developments towards human-level AI. ML techniques play a predominant part in data analysis in BCI systems [7]. Through these techniques, useful knowledge can be automatically obtained and this knowledge can be applied to the target task. Since existing BCI systems are not sufficient and capable enough for learning, understanding, and interpreting complex brain activities, the use of ML technology with BCIs can create wonders in understanding the complex brain signals and recognizing the actions more effectively.

1.1. Motivation

Diverse challenges are encountered by BCIs when exploited in real-world applications. The EEG, ERP, and other brain signals in BCI systems play a vital part in controlling, monitoring, and detecting human actions. However, many BCI systems fail to recognize and categorize these signals effectively. Inaccurate classification of these signals often degrades the BCI system performance. Moreover, the lack of methods for recognizing human mental states and emotions further adds to complexities in BCI systems employed for emotion detection and mental condition recognition. Additionally, the inadequacy of precise methods for recognizing hand/limb motions of individuals in BCI systems makes human motion identification more complex and weaker. All these BCI tasks depend on algorithms employed for feature extraction, selection, and categorization or detection of recognition parameters. Thus, the improvement of these algorithms can aid in the enhancement of BCI systems and their detection abilities. Therefore, utilization of better algorithms for feature extraction, selection, classification, and advanced AI algorithms such as ML can aid in improving BCI system performance and in achieving better outcomes, thereby supporting to deal with BCI real-life challenges more effectively and smoothly. This has given the motivation to explore better methods for analyzing and performing BCI tasks with greater precision and accurate results and to investigate the role of these methods in enhancing the performance of existing BCI applications.

1.2. Paper Organization

The rest of the paper is structured as follows: Section 2 investigates the existing works conducted on BCI and ML. Section 3 discusses the significance and role played by ML in diverse BCI tasks such as EEG signal classification, motor imagery classification, mental state recognition, etc. Section 4 discusses the different techniques employed for feature extraction, selection, and classification. Section 5 presents the findings of this work.

1.3. Contributions of the Study

The important contributions of this work include:

- To explore the role of ML in BCI applications.
- To determine the advantages of employing ML in various BCI tasks such as motor imagery classification, ERP signal classification, emotional and mental state detection, and several others.
- To study the diverse feature extraction, selection, and classification methods exploited for BCI applications.

2. Literature Survey

Currently, BCIs have drawn ample attention among researchers owing to their capability of understanding neural communications and functions within the brain. Motivated by this attractive feature of BCI systems numerous research studies have been conducted over the past several years. In this section, some existing works on BCIs have been explored.

In [4], the role of BCIs in control and communication was discussed. The features of BCI and its crucial parts were discussed. Furthermore, the different sorts of BCI based on utilization of electrophysical signals were described and the critical problems in BCI-based control and communication systems were highlighted. In [5], the role of BCIs in communication and motor rehabilitation was explained. This study discussed the use of BCIs for communication in individuals suffering from locked-in disorder or paralysis. BCI use in motor rehabilitation after spinal cord impairment and severe stroke was also been described. This study reported the promising advantages of BCIs in clinical applications. In [8], a study on P300-dependent BCI systems was presented. This study revealed that despite its acceptance, several P300-BCI systems still require improvement prior to considering them as substitute communication instruments for individuals suffering from locked-in problems. This work also recommended improvements in diverse regions of P300-BCI operations for optimizing the convenience, consistency, accuracy, and speed of existing systems, particularly for real-life environments. In [9], four application realms, namely, motor recovery, control and communication, entertainment, and motor substitution where physically challenged people could largely benefit from developments in BCI were identified. This work indicated the necessity for further progress in user–machine adaptable methods, hybrid BCI frameworks, and user's mental condition recognition schemes for developing better EEG devices and enhancing BCI usability. In [10], a profound study on EEG-based BCIs was conducted. The research trends in EEG-based BCIs were surveyed along with exploring BCI paradigms, target applications, kinds of BCI systems, feature types, and classification schemes. This study revealed that the number of BCI research works, particularly EEG-based BCI studies, have continuously increased over the past 5 years. In [6], non-invasive, EEG-dependent BCI systems for motor-impaired people's communication and rehabilitation were investigated. It focused on methods exploited for assisting locked-in people and severely paralyzed individuals to regain communication through three distinct BCI modalities, namely, P300 potentials, slow cortical potentials, and sensorimotor rhythms. This study also explored BCI systems utilized for motor function restoration in patients suffering a chronic stroke and spinal cord impairment. In [11], the BCI significance in multimodal interaction was discussed. The real-time and state-of-the-art multimodal applications of BCI were investigated for illustrating the usefulness of BCI as a modality in multimodal interaction. The technical and experimental challenges involved in multimodal interaction BCI systems were elaborated along with few guidelines for overcoming the encountered challenges such as interference and noise during multimodal blending, distraction during the process, experimental limitations concerning task, environment and the user, and the physical limitations concerning multimodal deployment. This study indicated that through multimodal BCI systems, user experience, task performance, and error handling could be improved. Moreover, the user spectrum could also be broadened using such systems. In [12], a study on BCI was presented. This study described the vital components of BCI and signal acquisition approaches. It explained the electrical signals of BCI such as event-related synchronization (ERS) and event-related potential (ERP). The application areas, usability challenges, and technical hardships encountered during brain signal usage in BCI systems were discussed. Furthermore, the necessary solutions for boosting the BCI system performance were provided. In [13], different classification approaches for EEG-oriented BCIs were reviewed. This study reviewed five categories of classification approaches, namely, nearest neighbor schemes, non-linear Bayesian schemes, neural networks, linear classifiers, and fusions of classifiers. This study revealed that among five categories, fusions of classifiers seemed very effective for contemporaneous BCI experiments. In [14], discriminative spatiospectral attributes of ERPs were studied

for capturing the relevant neural actions from EEG recordings. This study indicated that the highly discriminative frequencies essential for decoding ERPs were found to be in a spectral limit lower than 6.4 Hz. In [15], the ML approach was proposed for EEG signal analysis in real-time. It even discussed the significance of ML schemes for mental condition monitoring and EEG-oriented BCI applications. In [16], several classification schemes were investigated for EEG-BCI systems. Additionally, numerous challenges were identified for further strengthening the EEG categorization performance in BCI.

3. Role of ML in BCI

Recently, BCIs have acquired widespread attraction in human–machine interaction applications [17]. BCI systems perform their tasks through analyzing ERP, EEG signals linked with human motion, expression, and mental condition. EEG signals are generally employed for measuring diverse brain activities. ERPs are electrophysiological responses that arise at a specific period after a specific internal or external event [18]. They occur when exposed to a mental, sensory event or the dereliction of a consistently occurring stimulus. ERPs are employed for determining the subject's brain states in response to distinct stimuli. The ERPs are useful for diverse BCI-related tasks [19]. The EEG and ERP signal features are helpful in the classification of mental states, mental tasks, human emotions, and motor imagery (MI) in BCIs. MI-EEG is a self-controlled EEG that does not involve any external stimulus. In the MI-oriented BCI mechanism, the subject is urged to visualize moving distinct parts of the body for triggering neuronal activities in particular brain regions that are linked with the movements [20]. Generally, during the MI procedure, the event-related synchronization and desynchronization (ERS/ERD) patterns of EEG simultaneously occur. The ERD/ERS patterns generated by visualized or imagined movements are similar to original movements but may vary among diverse subjects. Therefore, every subject requires intensive training. Decoding of MI is a vital segment of the BCI system that translates the subject's intentions to commands which external gadgets can execute. However, the low signal-to-noise ratio (SNR) of EEG recordings makes accurate decoding of recorded EEG of neuronal activities further challenging. Conventional techniques for distinctive feature extraction only concentrate on energy attributes of EEG and highly ignore the further examination of temporal data, which is critical in performance enhancement of MI interpretation. In mental state-oriented BCIs, distinct sorts of mental actions trigger diverse parts of the cortex and provoke distinct EEG signals [21]. These signals aid in classifying mental tasks for realizing a BCI system. In emotion recognition-oriented BCIs, diverse emotions are recognized through analyzing EEG signals [22,23]. However, effective feature extraction and precise classification of ERP and EEG signals linked with mental states, mental tasks, MI, and emotions are serious challenges encountered in BCI systems. EEG-enabled BCI mechanisms possess low SNRs, suffer from weak spatial resolution and are anti-static by nature. These constraints further worsen owing to fluctuating electrode impedances, eye movements, muscular actions and variations in the mental states of users. All these aspects further complicate brain dynamics analysis and BCI-related heterogeneous signal classification. Under these contexts, ML techniques serve as highly promising and productive techniques for performing BCI-related tasks. ML approaches are even helpful in determining the phase of BCI medication for stroke rehabilitation [24]. The effective feature learning and complex pattern recognizing characteristics make ML a more beneficial and supportive technique for signal feature extraction and classification in BCI-related operations [25,26]. Table 1 compares different feature extraction/selection and classification methods in terms of publication year, performance metrics, accuracy level, types of BCI tasks, merits and demerits, and future directions.

Table 1. Comparison of feature extraction and classification techniques.

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[7]	2020	-	Linear regression	-	95%	EEG signal categorization	Showed high accuracy.	Involved numerous mathematical computations.
[15]	2008	CSP	LDA	-	-	EEG signal analysis	Analyzed EEG signals effectively.	Distinct physiological features should be combined for multi-category classification.
[19]	2012	-	SVM	Classification accuracy	90.55%	ERP signal categorization	Effectively categorized ERP signals.	More participants should be considered in the work.
[23]	2019	CSP	Independent decision path integration	Classification accuracy, specificity, sensitivity	70.32%	Mental state categorization	Categorized multicategory mental states.	Accuracy should be improved.
[24]	2018	PCA	SVM	Classification accuracy, specificity, sensitivity	92.5%	BCI therapy stage classification	Assisted in stroke rehabilitation.	More participants should be included in the study.
[27]	2015	EEG coherence-based method	Fisher's linear discriminant	Classification accuracy	73.6%	EEG coherence selection	Effectively discriminated motor and cognitive tasks.	The accuracy level requires improvement.
[28]	2006	ACSP	SVM	Classification accuracy	65.12%	Classification of EEG signals	Extracted most discriminative attributes pertinent to brain states.	Failed to show improvements in EEG signal categorization.
[29]	2003	GA	SVM, LDA, and NNs	Classification accuracy	76%	Classification of EEG signals	Improved feature selection.	Parameter variations and additional databases must be considered.
[30]	2013	WT, GA	KNN, SVM, LDA, Bayesian, MLP and weighted majority voting	Classification accuracy	-	EEG signal categorization	The proposed multiclassifier outperformed the individual classifiers.	-
[31]	2021	PCA	Convolutional autoencoder	Classification accuracy	-	MI classification	Showed promising MI categorization performance compared to contemporary methods.	Performance evaluation using additional and more discriminative subjects is required.
[32]	2020	Spatial filtering	Temporal-spatial CNN	Classification accuracy	65.7%	MI classification	Displayed significant performance enhancement compared with classical strategies.	Required separate training of classification layers and feature extraction layers.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[33]	2020	SSD	CNN	Kappa value, classification accuracy	79.3%	MI classification	Displayed high robustness and classification quality.	DL network structure and layer selection require optimization.
[34]	2005	ARX	LDA	Accuracy	79.1%	EEG feature extraction	ARX method outperformed the AR technique in EEG feature extraction.	Classification performance needs further improvement.
[35]	2019	Riemannian geometry, CSP and PSO	CNN	Classification accuracy	80.44%	EEG signal categorization	Improved EEG signal categorization accuracy for multifarious subjects.	Required data augmentation.
[36]	2013	ICA	NNs and SVM	Classification accuracy	89.8%, 97.1%	EEG signal categorization	Displayed promising feature extraction outputs for classifying EEG signals.	Performance evaluation with additional databases is required.
[37]	2018	WT	SVM	Sensitivity, classification accuracy	>90%	EEG signal categorization	Effectively performed multiclass classification with three distinct subjects.	-
[38]	2019	STFT, Continuous WT	CNN	Classification accuracy, specificity, error percent, kappa value, sensitivity, f1-score,	99.35%	MI classification	Achieved greater accuracy scores in MI categorization than existing strategies.	MI categorization using ResNet, GoogleNet, and VGGNet frameworks must be explored.
[39]	2018	CSP	Neuro-fuzzy scheme	Classification accuracy, the kappa value	91.43%	EEG signal categorization	Outperformed related existing techniques by 4.5%.	Intelligent optimization schemes are required for adjusting classification model parameters.
[40]	2016	mRMR	SVM	Classification accuracy	62.33%	BCI-oriented emotion recognition	Recognized multiple emotion types without employing additional classifiers.	Highly exhaustive setup assessments are required for assessing model parameters affecting the efficacy of the approach.
[41]	2011	WT	SVM	Classification accuracy	88.6%	EEG signal categorization	Exhibited excellent potential and encouraging outcomes towards asynchronous BCI applications.	EEG signals linked with MI should be analyzed.
[42]	2020	PCA, FLD	K-ELM	Classification accuracy	96.54%	MI classification	Achieved higher accuracy score than other approaches.	Experienced certain data loss due to the energy function of compressed input information.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[43]	2017	WT, Lasso regularization, mRMR	GNB, LDA, SVM	Confusion matrix, k-score, classification accuracy	95.47% (GNB), 91.10% (LDA), 92.26% (SVM)	MI classification	Improved two-class MI categorization using only a few feature vectors.	Multiclass MI categorization should be conducted.
[44]	2016	WT, PCA, FFT	SVM, ANN	Classification accuracy	84%	EEG signal categorization	Effectively categorized EEG signals linked with five distinct mental tasks.	EEG signals linked with MI and emotions need to be categorized.
[45]	2016	Kolmogorov complexity	Adaboost ELM	Classification accuracy	79.5%	EEG signal categorization	Improved EEG signal categorization for multi-category samples.	Performance of devised approach in distinct mental task categorization for BCI development should be inspected.
[46]	2016	WT	Probabilistic NB	Classification accuracy	78.33%	Limb movement categorization	Considered both spatial and frequency domain attributes of EEG without immolating the accuracy.	The accuracy level should be upgraded.
[47]	2018	Short-term windowing, symmetrical uncertainty, information gain, oneR, correlation, and evolutionary technique	Bayesian networks, random forest, SVM	Classification accuracy	87%	Mental state categorization	Effectively categorized diverse mental states.	DL models should be exploited for further incrementing the accuracy.
[48]	2006	AR	ELM, SVM, BPNN	Classification accuracy	45.66 ± 7.95% (BPNN), 49.68 ± 8.30% (ELM), 52.07 ± 9.11% (SVM)	Mental task categorization	ELM utilized a shorter training time than BPNN and SVM.	The accuracy score should be intensified.
[49]	2020	CSP	MLP	Mean square error, classification accuracy	97.8214%	EEG signal categorization	Displayed acceptable performance than the rest.	Additional metrics must be employed for assessing the EEG signal categorization performance.
[50]	2020	Wavelet ICA	Fuzzy kernel-SVM	Specificity, accuracy, sensitivity	86.1%	EEG signal categorization	Automatically removed EEG signal artifacts from the raw database.	Further increment in accuracy score is necessary.
[51]	2014	EMD	RBF kernel with SVM	Classification accuracy	100%	EEG signal categorization	Displayed superior EEG signal categorization accuracy.	Additional samples must be incorporated for assessing EEG categorization performance.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[52]	2020	FAWT	Subspace KNN, LDA, SVM, Decision Trees (DT), standard KNN	Accuracy, specificity, kappa value, sensitivity, f1-score	99.33% (Subspace KNN), 81.1% (LDA), 95.72% (SVM), 91.79% (DT), 92.8% (standard KNN)	MI classification	Outperformed the previously available methods in MI categorization with the same dataset.	Parameters employed should be optimized for incrementing the accuracy.
[53]	2019	AR, WT, WPD, class separability feature selection	Ensemble ELM with LDA	Classification accuracy	99.43%	EEG signal categorization	Achieved higher class separability compared to previously existing schemes.	Applicability of ensemble ELM with LDA approach to other BCI-pertinent biomedical signals must be examined.
[54]	2020	Continuous WT	Autoencoder, SVM, logistic regression and MLP	Accuracy, recall, precision, f-score	92.09 ±0.5% (autoencoder), 88.48 ± 0.5% (SVM), 89.25 ± 0.5% (logistic regression), 95.58 ± 0.5% (MLP)	EEG signal categorization	Outperformed the cutting-edge learning frameworks and yielded greater accuracy rates.	Utility of approach in diverse BCI-pertinent tasks must be assayed.
[55]	2009	WT	Fuzzy SVM	Classification accuracy, classification time, minimal misclassification rate,	80.71%	EEG signal categorization	Provided a contemporary means for online EEG signal categorization.	Required huge computational effort and training time.
[56]	2016	PCA	ANN, naive bayes, KNN, LDA, SVM	Classification accuracy	71.80% (ANN), 55.52% (naive bayes), 65.51% (KNN), 62.94% (LDA), 68.56% (SVM)	EEG signal categorization	Efficiently classified EEG signals linked with mental states.	Further increment in accuracy score is needed.
[57]	2016	FFT	MLP-ANN, KNN, SVM, logistic regression	Classification accuracy	66.42% (MLP-ANN), 56.71% (KNN), 68.97% (SVM), 73.03% (logistic regression)	EEG signal categorization	Classified EEG signals linked with MI within a limited period.	Further increment in accuracy levels is required.
[58]	2014	AR, Continuous WT	LDA, KNN, SVM	Classification accuracy	55.92% (LDA), 57.90% (KNN), 82.24% (SVM)	EEG signal categorization	Showed a 12.25% improvement in EEG signal categorization.	Utilized more time for decision formulation.
[59]	2017	WT, PCA	Naive Bayes, KNN, ANN, SVM	Classification accuracy	50.7% (naive bayes), 49.82% (KNN), 55.58% (ANN), 51.82% (SVM)	EEG-dependent emotion classification	Classified EEG-dependent emotions accurately.	Lesser electrodes should be employed for experimentation.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[60]	2014	WT, FFT, PCA, LDA, correlation-dependent feature selection	SVM	Classification accuracy	91.77%	EEG-dependent emotion classification	Provided a promising procedure for visualizing the subject's emotional condition/state.	More subjects and experiments are required for improving the presented model's efficiency.
[61]	2018	FFT	LDA, KNN, SVM	Classification accuracy	95% (LDA), 100% (KNN), 100% (SVM)	EEG signal categorization	Provided a better method for examining EEG signals from diverse human cognitive conditions.	Investigated EEG attributes considering only two mental exercises.
[62]	2012	AR, PSD, Hjorth parameter, LOO, LAR	LDA, SVM	Classification accuracy	74.3% (LDA), 70.5% (SVM)	EEG signal categorization	Did not require tunable parameter for classification and attribute selection phases.	Conducted EEG signal categorization only using an offline approach.
[63]	2012	Bayesian approach with SSO	SVM	Classification accuracy	>95%	MI classification	Outperformed the cutting-edge techniques in MI categorization.	Other features such as cortical potential, readiness potential, or Bereitschafts potential should be incorporated apart from ERS/ERD features.
[64]	2019	Hilbert transform	SVM, Naive bayes, LDA	Classification accuracy, kappa coefficient	82.22% (SVM), 71.64% (naive bayes), 75.52% (LDA)	MI categorization	EEG attributes extracted via Hilbert transform showed the finest performance than previously exploited methods.	Applicability of approach towards diverse BCI-pertinent tasks must be inspected.
[65]	2017	CSP with sparse regression	Weighted naive Bayes	Classification accuracy	85.24%	MI categorization	Improved MI classification performance.	Required large computation time.
[66]	2013	ICA, WT, PCA, AR, interval feature extracting method, fast correlation-dependent filter	Perceptron SVM, linear SVM, random forest, KNN	Classification accuracy	97% (Perceptron SVM), 86% (Linear SVM), 97% (Random Forest), 97% (KNN)	ERP signal categorization	Classified ERP signals accurately.	Electrode information should be coupled with extracted valuable features for further incrementing the accuracy.
[67]	2018	CSP	Multi-kernel ELM	Classification accuracy	87.5 ± 10.5%	EEG signal categorization	Provided an improved solution for devising an MI-oriented BCI.	An increment in accuracy score for undersized samples is necessary.
[68]	2020	CSP	CNNs	Classification accuracy	72.7296% (mental state categorization), 48.0469% (subject independent emotion categorization)	Emotional and mental state categorization	Performed both emotional and mental state categorization using EEG signals.	Emotional and mental state categorization using other bio-medical signals must be investigated.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[69]	2016	-	ANN, KNN	Classification accuracy, sensitivity	98.58% (ANN), 96.06% (KNN)	EEG signal categorization	Displayed fine performance in EEG signal categorization.	Performance must be incremented by combining EEG with diverse biomedical signals.
[70]	2006	PCA	SVM	Classification accuracy	>95%	EEG signal categorization	Lowered training time and substantially incremented speed and accuracy.	-
[71]	2007	-	SVM, KNN and DT	Classification accuracy	64.92% (SVM), 64.63% (KNN), 56.74% (DT)	EEG signal categorization	Ensemble schemes displayed better EEG signal categorization over an individual base classifier.	Issues regarding online evaluation and parameter tuning should be probed.
[72]	2020	-	LSTM	Classification accuracy, recall, f1-score, precision	97.13%	EEG signal categorization	Offered a reliable platform for intelligent visual classification.	Sophisticated techniques are required for distinguishing EEG signals for additional image categories.
[73]	2021	-	KNN	-	-	EEG signal categorization	Displayed highest accuracy scores.	-
[74]	2017	PCA	LDA	Classification accuracy	71 $\pm$10.9%	EEG signal categorization	Extracted both time-frequency and temporal attributes effectively from three distinct brain lobes.	The accuracy level must be incremented.
[75]	2013	-	SVM	Classification accuracy	71.43%	EEG signal categorization	Exhibited good EEG signal categorization performance.	The accuracy score must be upgraded.
[76]	2017	-	ANN, LDA, Bayesian classifier	Classification accuracy	78.84% (ANN), 70.05% (LDA), 65.08% (Bayesian)	EEG signal categorization	Performed better with diminutive training sets and time-variant brain signals.	Further increment in accuracy rate is needed.
[77]	2018	-	Semi-supervised ELM	Classification accuracy	68.12 $\pm$ 1.38%	EEG signal categorization	Provided an efficient and safe approach for categorizing EEG signals.	Additional evaluation measures for examining the risk level of unlabeled information instances should be incorporated.
[78]	2014	-	Multiple kernel-SVM	Classification accuracy	99.20% (2-class), 81.25% (3-class), 76.76% (4-class), 75.25% (5-class)	EEG signal categorization	Effectively executed multi-class categorization of EEG signals linked with mental tasks.	The categorization of EEG signals for BCI-pertinent tasks should be scrutinized.
[79]	2013	-	SVM	-	-	EEG signal analysis	Discussed the role of SVM in EEG signal analysis.	Experiment validation of SVM and its variants in EEG signal categorization should be conducted.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[80]	2016	Band-power scheme	ELM, LDA, SVM	Mutual information, classification accuracy	82.02% (ELM), 77.1% (LDA), 78.18% (SVM)	EEG signal categorization	Displayed greater accuracy level and mutual information.	-
[81]	2015	-	SBLaplace method	Classification accuracy	-	EEG signal categorization	Automatically estimated model parameters without requiring cross-validation.	The performance of EEG signal categorization should be further intensified.
[82]	2020	-	Sparse Bayesian ELM	Classification accuracy	78.5 ± 14.3%	EEG signal categorization	Exhibited good EEG signal categorization accuracy.	More distinctive and high-level attributes should be learned for augmenting the categorization performance.
[83]	2017	STFT	CNN-SA	Classification accuracy, kappa value	90.0%	MI classification	Performed swift classification with few samples and yielded greater performance.	More pooling layers should be incorporated for boosting MI categorization performance.
[84]	2020	-	CNN	Classification accuracy	97.28%	MI classification	Classified raw EEG signals linked with MI without any synthetic feature extraction and preprocessing operations.	A system for real-time EEG wave/signal acquisition should be constructed.
[85]	2020	Stockwell transform	DML	Classification accuracy, recall, precision	64.7%	MI classification	Offered a promising procedure for classifying MI signals with just fewer training samples.	The accuracy level must be further augmented.
[86]	2018	-	CNN	Classification accuracy	86.13%	MI classification	Showed 6–9% of mean improvement in MI categorization accuracy.	Categorization of EEG signals linked with other BCI functions is needed.
[87]	2015	-	Bayesian network	Kappa coefficient	-	MI classification	Displayed excellent multiclass MI categorization performance.	Alternatives for minimizing the time involved for training information collection should be probed.
[88]	2010	-	CNN	Classification accuracy, noise, precision, error, f1 measure, silence, recall	95.5%	ERP signal classification	Provided a contemporary approach for examining brain activities.	The impact of P300 waves on character identification problems should be studied.
[89]	2011	-	LDA	Classification accuracy	-	ERP signal classification	Provided an effective procedure for ERP signal classification.	The accuracy level must be augmented.

Table 1. *Cont.*

Reference	Publication Year	Feature Extraction/Selection Method	Classification Method	Performance Measure	Accuracy Level	BCI Task	Merits	Demerits or Future Directions
[90]	2014	-	Aggregated sparse LDA	Classification accuracy	61.5 ± 19.2%	ERP signal classification	Demonstrated better performance even with inadequate training samples.	Applicability of aggregated sparse LDA for diverse BCI-related functions must be reconnoitered.
[91]	2013	-	STDA	Classification accuracy	80.8 ± 6.32%	ERP signal classification	Displayed superior ERP signal categorization performance with scarce training samples.	Categorization performance must be further augmented.
[92]	2014	-	SVM	Classification accuracy	59.75%, 72.89% and 91.42% for four, three and two mental tasks	Mental task categorization	Effectively executed multiclass mental task categorization.	Involved in an additional procedure for determining the best electrodes and tasks.
[93]	2021	-	CNN	-	-	EEG signal categorization	Displayed good EEG signal categorization performance.	Cortical regions involved with hand unlock/close gestures must be studied.
[94]	2018	CSP	LDA, ELM, SVM	-	-	EEG signal analysis	Described brain activities leading to a substantial rise in hemodynamic response.	Suggested recognition of brain areas and categorization of hemodynamic reactions.
[95]	2013	Spatial filtering	Bayesian network	Classification accuracy	-	EEG signal analysis	Effectively analyzed EEG signals.	The accuracy score should be further enhanced.
[96]	2010	-	Unsupervised LDA	Mean, PMean	-	EEG signal analysis	Outperformed the cutting-edge approaches.	Analysis pertinent to asynchronous BCI tasks is required.
[97]	2019	-	Sparse discriminant analysis	Classification accuracy, time complexity	>60%	P300 detection and categorization	Exhibited good P300 categorization accuracy.	The accuracy score must be further incremented.
[98]	2018	PCA	Weighted SVM	-	-	P300 detection	Performed better in P300 signal recognition.	Training time must be lowered.
[99]	2016	-	SVM	Accuracy	92.5%	P300-based BCI operation	Improved P300 detection performance.	Augmentation of the test set for accuracy up-gradation is required.
[100]	2011	-	Linear SVM, stepwise-LDA, fisher's LDA, Bayesian-LDA, ANN, non-linear SVM	Intensification sequences	-	P300-based BCI operation	Bayesian LDA exhibited remarkable performance.	The impact of the P300 speller on stroke patients should be scrutinized.

4. Methods

4.1. Feature Extraction and Selection

Various feature extraction and selection techniques are employed for BCIs in the literature [25,26]. Commonly exploited techniques explored in this survey include:

4.1.1. Common Spatial Pattern (CSP)

CSP is a feature extraction technique wherein multichannel EEG signals are projected into a subspace [101] and differences between the categories are highlighted while the similarities are reduced. In this scheme, Gaussianity is assumed in addition to considering the time and frequency as known variables. In CSP, a transformed EEG matrix could be obtained by implementing the following steps:

Normalization of the spatial covariance of EEG [101] using Equation (1)

$$C_K = (Y_K\,Y_K{}^T)/Trace\,(Y_K Y_K{}^T) \tag{1}$$

where K represents the categories and *trace*(y) represents the sum of diagonal values of y.

Computation of spatial covariance using Equation (2):

$$C_{ov} = \sum C_K, \quad \forall\ \text{categories} \tag{2}$$

$$C_K = Z_K \lambda_k\,Z_K^T \tag{3}$$

where $\breve{}_K$ represents eigenvalue's diagonal matrix and Z_K represents eigenvector matrix.

Computation of projection matrix using Equation (4):

$$P = Z^T\,U, \tag{4}$$

where U denotes whitening transformation matrix estimated using $U = \sqrt{\lambda}Z_0$. The reduction of original EEG signal into uncorrelated components through projection matrix using Equation (5):

$$W = PY, \tag{5}$$

where W denotes the source component of EEG signals containing specific and common components of diverse tasks.

Estimation of original EEG 'Y' using Equation (6):

$$Y = P^{-1}\,W \tag{6}$$

where columns of P^{-1} denote the spatial patterns.

The CSP method was exploited for extracting highly distinctive EEG features in [28,35,49,67] for EEG signal categorization. An artifact-dropping CSP scheme was utilized in [39] for extracting EEG attributes. This feature extraction approach provided more robust features than the ordinary CSP technique. This approach achieved 91.43% accuracy. The CSP scheme used in [65] for extracting discriminative temporal, frequency, and spatial features utilized a sparse regression technique for selecting crucial features for classifying EEG-MI signals. This methodology is aimed at optimizing the spatial-temporal-frequency patterns for distinctive feature extraction. Through identifying distinctive spatial features on every time-frequency segment, spatial optimization was achieved. Distinctive features contained in several time-frequency segments were selected automatically by this approach. A further selection of vital features was achieved using a sparse regression scheme. Optimization of spatial-temporal-frequency patterns greatly augmented the classification performance.

4.1.2. Principal Component Analysis (PCA)

PCA is a well-known method for dimensionality reduction and feature extraction. In PCA, input data are projected on a *s*-dimension eigenspace of *s* eigenvectors. All data points are projected in the first eigenvector's direction such that resultant variance is at

optimum. PCA executes transformation through optimizing the rate of decrease of data variance. It employs a transformation matrix containing low variance elements. The transformation matrix T_M is represented using Equation (7):

$$T_M = (1/n) \sum_{i-1}^{n} (pi - \mu)(pi - \mu), \tag{7}$$

where *pi* represents the element of the *N* dimension database and *n* represents total elements in the actual database.

$$T_M \cdot Z = \Lambda \cdot Z, \tag{8}$$

where *Z* indicates the matrix with eigenvectors $z_1, z_2, \ldots, z_n$ and Λ denotes eigenvalue diagonal matrix containing elements $\lambda_1, \lambda_2, \ldots, \lambda_n$.

The PCA technique was exploited in [42,44,56,59,66] for extracting signal features. The authors in [31] exploited the PCA method for extracting features. They split the time sequences of different EEG channels into transient blocks and computed connectivity matrices for every block through adaptive sparse representation. They constructed the connectivity patterns and applied PCA to those patterns for learning discriminatory representations. The PCA-based feature extraction led to promising MI categorization performance than existing schemes.

4.1.3. Independent Component Analysis (ICA)

In ICA, data are decomposed into several independent components as per their statistical interrelation. ICA expresses the output/resultant signal *v(t)* with respect to input/source signal *u(t)* as represented by Equation (9):

$$v(t) = f(u(t)) + n(t), \tag{9}$$

where *f* denotes the mixer function and *n(t)* represents the noisy signal.

ICA method was employed for feature extraction in [50,66]. In [36], ERS, ERD, and movement-pertinent cortical potential features were exploited for classifying left/right-hand motion EEG signals. These features were extracted through ICA. The employed ICA technique encouraging feature extraction outputs for classification.

4.1.4. Autoregressive (AR) Method

The AR method is employed for extracting time-domain attributes. It is generally exploited in time series fitting and can characterize EEG signals better. AR coefficients thus provide effective features from EEG signals. Let $\{z_1, z_2, \ldots, z_n\}$ be a time series, now the AR model can be defined using Equation (10):

$$z_t = \Phi_0 + \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \ldots + \Phi_k Z_{t-k} + \epsilon_t, \tag{10}$$

where *k* denotes the order of AR model, ϵ_t denotes Gaussian noise, Φ_0 represents absolute term and $\{\Phi_1, \Phi_2, \ldots, \Phi_k\}$ represents autoregressive coefficients. Using the AR model, the relationship between historical values $\{z_{t-1}, z_{t-2}, \ldots, z_{t-k}\}$ and the observed value z_t can be determined.

The authors in [34,48,62,66] executed feature extraction using AR technique. Highly useful features were extracted using this technique. In the AR method, determining the AR coefficients and order are considered major issues. In [53], AR coefficients were estimated using the Burg technique, and order was determined using the Akaike criterion. The authors in [58], estimated AR coefficients through the forward–backward method. The estimated coefficients were employed as features, which were determined based on AR order, recorded time length, and selected channel. This AR-based feature extraction scheme provided effective features for EEG signal categorization.

4.1.5. Wavelet Packet Decomposition (WPD)

WPD is a feature analysis technique in the frequency-time domain. In WPD, signals are projected onto the space which is spanned through a group of mutually orthogonal wavelet basis functions and are decomposed into high frequency and low-frequency parts. Figure 2 illustrates the steps involved in WPD [53]. The $u(n)$ and $v(n)$ in Figure 2 represent the low-pass and high-pass filters. The signal in WPD is initially decomposed into low-frequency and high-frequency parts. Further, (in the next layer) the signal is again decomposed into low frequency and high-frequency parts. This decomposition process is continued until desired features are obtained. The wavelet packet coefficients of $(y + 1)$ th level and zth sample are obtained using Equations (11) and (12):

$$a_z^{2n,\, y+1} = \sum_p v\,(p - 2z)\, a_p^{n,\, y} \tag{11}$$

$$a_z^{2n+1,\, y+1} = \sum_p u\,(p - 2z)\, a_p^{n,\, y} \tag{12}$$

where $n = 0, 1, 2, \ldots, 2y - 1$. v and u are filters and a denotes wavelet packet coefficient.
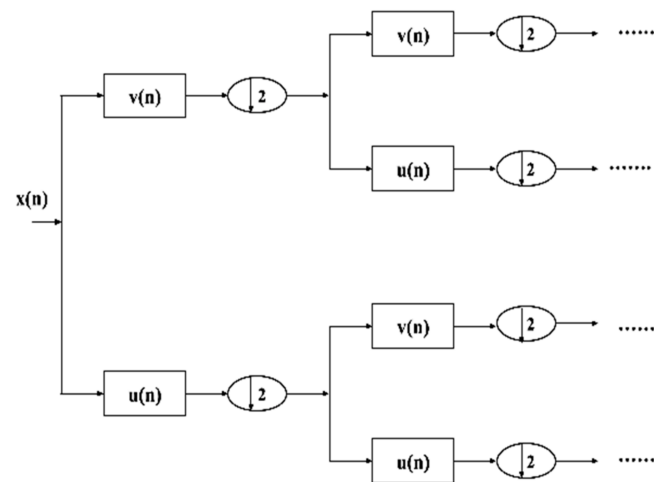
Figure 2. WPD process.

The packet coefficients of the xth band form a 1-dimensional vector $a_x = [a_{x\,1}, a_{x\,2}, \ldots, a_{x\,p}]$.

WPD method was exploited in [53] for extracting EEG signal features. This method provided effective features for classifying EEG signals.

4.1.6. Wavelet Transform (WT)

WT is another popular frequency-time analysis scheme. In WT, functions are obtained through either shifting or scaling a single function called the mother wavelet. On application of WT, signals are decomposed into diverse frequency ranges, which are further categorized into detail and approximation levels. For any signal $c(n)$, the WT is defined using Equation (13):

$$WT(p, q) = |p|^{-1/2} \int_{-\infty}^{\infty} c(n)\, \varphi\left(n - \frac{q}{p}\right) dn \tag{13}$$

where $\varphi(n)$ represents wavelet function, p represents the scale factor, and q represents shift factor.

The two crucial types of WT include Discrete Wavelet Transform (DWT) and Continuous Wavelet Transform (CWT). Figure 3 illustrates the multi-scale decomposition through DWT [53]. The $u(n)$ and $v(n)$ in Figure 3 are the low-pass and high-pass filters. Initially, the

original/input signal $x(n)$ is decomposed into high and low frequencies. In the next layer, the low-frequency part is decomposed further until desired features are acquired.
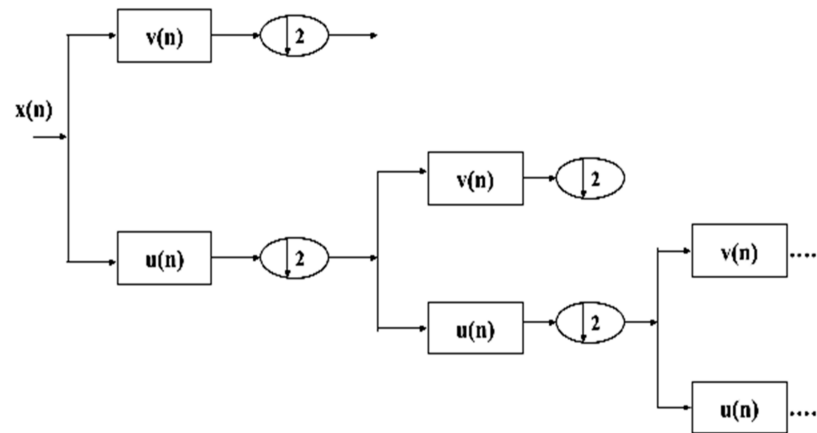

Figure 3. DWT process.

WT was utilized for extracting EEG features in [30,37,41,43,46,54–66]. The original EEG signal was reduced into detail and approximate frequency coefficients. In the initial level, the EEG signal was transformed into low-frequency and high-frequency components of length half that of the original signal. The transformation was applied till desired features were obtained. The adoption of WT in these works yielded useful EEG features for classification. A wavelet ICA method was introduced in [50] by fusing WT and ICA for extracting necessary features required for categorizing EEG signals. This approach provided effective features necessary for categorizing EEG waves. EEG signal attributes in [52] were extracted via a flexible analytic WT (FAWT) scheme. Initially, EEG signals were decomposed into sub-bands by FAWT, and then temporal moment-dependent EEG features were extracted from those sub-bands for MI classification. In [59], EEG features were extracted through the WT technique while the feature dimensions were reduced using PCA for EEG-enabled emotion classification. EEG features for MI classification were extracted through continuous WT and Short-Time Fourier Transform (STFT) in [38]. The continuous WT scheme provided better feature extraction outputs than the STFT method. A hybrid EEG feature extraction scheme was employed in [53]. The required features were extracted through the autoregressive approach, wavelet packet decomposition (WPD), and WT methods. The redundant and irrelevant features were discarded by selecting only the optimal features using a class separability-based feature selection scheme. This hybrid EEG feature extraction scheme provided 99.43% accuracy.

4.1.7. Fast Fourier Transform (FFT)

The time signals are converted to the frequency domain through Fourier analysis. FFT is performed by estimating the discrete Fourier transform (DFT) of the sequence. FFT estimates the transformation through factorizing the DFT matrix into a product of sparse factors. The N-point DFT of a sequence is represented using Equation (14):

$$X_k = \sum_{n=0}^{N-1} x_n\, e^{-i2\pi k/N} \tag{14}$$

where X_k denotes the energy of current frequency and $k = 0, \ldots, N-1$ represents current frequency.

The authors in [44,57,61] exploited FFT for extracting EEG features. It provided desired features for EEG signal categorization. The features required for emotion categorization from EEG information were extracted in [60] using the FFT method along with the WT technique. Both of the methods provided useful features for emotion discrimination. The

employed feature extraction methodology offered a promising procedure for visualizing the subject's emotional condition/state.

4.1.8. Other Techniques

In [29], feature selection for EEG signal categorization was achieved using a genetic algorithm (GA) with SVM. Based on selected features by GA, data were classified using SVM. This approach yielded 76% accuracy. In [32], the filter bank spatial filtering method was exploited for extracting required features from EEG raw signals through retaining only discriminative data and eliminating features that are irrelevant for MI classification. The extracted EEG features considerably enhanced the categorization performance. In [33], a sparse spectrotemporal decomposition (SSD) scheme was exploited for extracting relevant signal attributes necessary for EEG-MI categorization. This feature extraction scheme showed promising outcomes in obtaining desired features. In [35], EEG feature extraction was performed through Riemannian geometry and common spatial features (CSP) methods, and a particle swarm optimization (PSO) scheme was utilized for replacing irrelevant features with the mean values of contiguous features. The relevant features thus selected improved EEG signal categorization accuracy for multifarious subjects. In [40], appropriate features for BCI-oriented emotion recognition were selected through the minimum-redundancy–maximum-relevance (mRMR) method. The features were selected based on two conditions, namely, maximum relevance and minimum redundancy. The mRMR scheme selected the better features by reducing data redundancy between good and bad features and choosing the ones which correlated the most with the categorization variable. The selected features helped in recognizing multiple emotion types. In [45], features of EEG signals captured during resting state, right hand MI and left hand MI were extracted through a Kolmogorov complexity approach. Initially, EEG data were segregated into non-overlapping fragments. Then features were extracted via a Kolmogorov complexity technique for signal categorization. These features improved EEG signal categorization for multi-category samples. In [51], an empirical mode decomposition (EMD) scheme was used for extracting EEG signal attributes. Initially, EEG signals were decomposed into integral mode functions (IMFs). Further characteristic features acquired from IMFs were fed as inputs to the classifier. These features greatly enhanced the EEG signal categorization performance. Feature selection in [43] was executed using lasso regularization and the mRMR technique. The selection of optimal features through these feature selection approaches helped in addressing the inter-subject inconsistency and in enhancing the classifier's performance. EEG signal attributes in [47] were extracted through the short-term windowing scheme for classifying the mental state. Feature selection was carried out using different evaluators such as symmetrical uncertainty, information gain, oneR, correlation, and evolutionary scheme. Among all feature selection evaluators, the evolutionary technique displayed good results in selecting relevant EEG signal attributes. The authors in [62] employed power spectral density (PSD), autoregressive modeling, and Hjorth parameter methods for extracting EEG features. The feature ranking was executed using least angle regression (LAR), and optimal features were selected through a leave-one-out (LOO) criterion. These methods offered valuable features for EEG signal categorization. A Bayesian approach was exploited in [63] for extracting discriminative EEG features. The feature extraction procedure was optimized through a spatiospectral filter optimization (SSO) scheme for the MI classification task. This Bayesian-SSO approach provided improved features for MI categorization and largely helped in attaining >95% accuracy. The authors in [64] exploited the Hilbert transform method for extracting signal features. The proposed Hilbert transform scheme effectively extracted the ERP and band power features for classifying MI tasks. EEG attributes extracted via Hilbert transform showed the finest performance compared to previously exploited methods. In [66], required features were extracted through an interval feature extraction method. The performance of this technique was compared with ICA, WT, and PCA. The proposed interval feature extracting method displayed superior performance than others. The redundant and irrelevant features were

discarded and desired features were selected through a fast correlation-dependent filter technique. In a fast correlation-dependent filter, relevant features are correlated with class variables and not correlated among themselves. Features are selected as per their mutual data with class variable while the features with mutual data below a selected threshold are discarded. The employed feature extraction methodology offered improved features for ERP signal categorization.

4.2. Classification

Different classification methods are exploited in BCI systems [25,26,68]. Commonly exploited techniques explored in this survey include:

4.2.1. K-Nearest Neighbor (KNN)

In KNN, training samples closest to an unobserved point are determined and are assigned to the dominant class. For BCI, nearest neighbors generally are obtained by utilizing a distance metric. In [30], the distance between other samples and target sample was determined using Euclidean distance measure:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(xi - yi)}\, 2, \tag{15}$$

where $d(x,y)$ denotes a distance between x and y samples, n represents a number of features, and x_i, y_i represents the sample's ith feature.

KNN was employed in [56–58,61,69] for categorizing EEG signals. In [73], KNN was used for EEG signal categorization. For classification, Euclidean distance was measured between neighboring signals, and further majority class was allocated to test signal among K neighbors. The KNN method in [73] offered the highest categorization accuracy value with enhanced specificity and sensitivity percentages. In [52], a subspace KNN scheme was employed for MI categorization. The subspace KNN scheme estimated a fresh set of KNN whenever an arbitrary subspace was selected. The majority voting on the test sample's class membership was carried out using aggregating k adjacent neighbors in every selected subspace. This subspace KNN achieved 99.33% accuracy and outperformed existing methods.

4.2.2. Linear Discriminant Analysis (LDA)

LDA seeks to isolate two or more categories of events or objects representing distinct categories. It employs hyperplanes for accomplishing this operation. Separating hyperplane is achieved by searching for a projection that optimizes the distance between the classes. The authors in [30,41,58,74] employed LDA for categorizing EEG signals and in [89] for classifying ERP signals. An aggregated sparse LDA approach was employed in [90] for categorizing ERP signals. For classification, the aggregated sparse LDA learned multiple discriminant vectors by utilizing the conformity between least-squares regression and LDA. This approach performed well and yielded better outputs for single-test ERP classification and outperformed the standard LDA.

4.2.3. Naive Bayes

In naive Bayes, features are assumed to be independent in every class. It forecasts the class C of an arriving instance Z consisting of features $[z_1, \dots, z_n]$ through estimating the highest probability using Equation (16):

$$P(C_i\,/Z) = \left(P(C_i)\prod_j P(z_j/C_j)\right)/P(Z), \tag{16}$$

The naive Bayes technique was employed in [46,47,56,59] for categorizing EEG signals and in [64,65] for MI categorization. The authors in [65], classified EEG-MI signals by assigning a weight for every selected CSP attribute using a weighted naive Bayes method. This approach provided 85.24% accuracy in MI categorization and performed better than

various competing techniques in existing works. EEG-MI signal categorization in [43] was accomplished using the Gaussian Naive Bayes (GNB) technique. The EEG-MI signals were categorized by considering Gaussian distribution and exploiting the naïve Bayes concept. The categorization performance of GNB was compared with two classical classifiers, namely, SVM and LDA. The experimental assessment reported that GNB showed better performance than SVM and LDA as GNB achieved 95.47% accuracy while SVM and LDA achieved 92.26% and 91.10% accuracy. The probabilistic naive Bayes (NB) method was employed in [46] for limb movement classification. It aimed at classifying right and left leg motion directly from EEG signals. Comparison of probabilistic NB with conventional NB indicated that the probabilistic NB performed better than conventional NB in classifying right–left lower limb motion by achieving 78.33% accuracy.

4.2.4. Extreme Learning Machine (ELM)

ELM is based on random initialization of hidden nodes of single-layer hidden feed-forward networks. The only parameters to be learned are the connections between the output layer and the hidden layer [80].

EEG signal categorization was performed in [53,77] using the ELM method. The exploitation of ELM for EEG signal categorization in [80] revealed that it displayed greater accuracy than SVM and LDA schemes. In [48], the ELM technique was exploited for classifying distinct mental tasks. For classification, the weights between hidden and input neurons were chosen and fixed depending on some probability density functions. Then, the weights between output and hidden neurons of the network were determined. Comparison of ELM performance with back propagation neural network (BPNN) and SVM indicated that ELM performed better concerning training time. EEG signals were classified using a multi-kernel ELM in [67]. This approach was developed by integrating polynomial kernel and Gaussian kernels for exploring supplementary data from multiple features for a more powerful categorization of EEG. In multi-kernel ELM, multiple kernels were defined and fused for achieving better EEG signal categorization. This method provided precise EEG classification outputs as compared to similar existing schemes and achieved $87.5 \pm 10.5\%$ accuracy. A sparse Bayesian ELM method was employed for classifying EEG signals in [82]. This approach considered the advantageous characteristics of both Bayesian learning and ELM for EEG signal categorization. The redundant data reduction was achieved by utilizing a relevance identification prior. The ELM structure was optimized by excluding surplus hidden neurons automatically. This method exhibited $78.5 \pm 14.3\%$ accuracy. In [42], EEG-MI signal categorization was accomplished using a kernel extreme learning machine (K-ELM). For classification, the specific kernel function was chosen using a specific kernel technique. The proposed K-ELM outperformed the Bayesian approach in EEG-MI signal categorization. Moreover, it achieved 96.54% accuracy.

4.2.5. Support Vector Machine (SVM)

SVM employs supervised learning for isolating two distinct categories of data. The classes are detected using a discriminant hyperplane. In SVM, the hyperplane that optimizes the distance from the nearest training points is selected. It employs diverse kernels functions such as linear, polynomial, radial basis, etc. for performance optimization.

SVM was employed in [28,29,36,44,58,70,75] for classifying EEG signals and in [60] for emotion categorization. SVM was also exploited for ERP signal categorization in [66]. EEG signals linked with random words and right and left body movements were classified robustly in [37] using the multiclass SVM approach. The employed SVM provided encouraging outputs with 52.78%, 86.12%, and 96.88% sensitivities for EEG signals linked with random word and right, left body movements. The adoption of SVM classifier in [41] led to improved EEG signal categorization than LDA. It provided 88.6% accuracy in EEG signal categorization. The authors in [50] employed a fuzzy kernel-SVM approach for classifying EEG signals. This approach offered 85.2% specificity, 86.1% accuracy, and 77.7% sensitivity. EEG signals were classified by radial basis function (RBF) Kernel with the SVM

approach in [53]. It offered 100% EEG signal categorization accuracy. In [55], EEG signals were classified using a fuzzy SVM scheme. In the proposed fuzzy SVM classifier minor proportion of support vectors were employed for choosing trade-off parameter and kernel parameter along with membership parameter solely depending on training data. This approach classified EEG signals effectively and outperformed the standard SVM technique by achieving 80.71% accuracy. The authors in [78] exploited a multiple kernel-SVM scheme for classifying EEG signals. The RBF kernel and polynomial kernels were combined to form multiple kernels. The proposed multiple kernel-SVM displayed better performance than single kernel SVM by providing mean accuracy of 99.20% and >75% for two-class and multiclass classification. In [79], the suitability of SVM for EEG signal categorization was discussed. It suggested SVM as the efficient classifier for categorizing EEG signals. Distinct ML classifiers such as SVM, naive Bayes, and LDA were exploited in [64] for classifying both right- and left-hand MI functions. In classifying MI functions, different classifiers such as SVM, naïve Bayes, and LDA achieved 82.22%, 71.64%, and 75.52% accuracy. The outcomes showed that the SVM performed better than LDA and naive Bayes schemes. In [92], the classification of diverse mental tasks for the BCI system was performed using the SVM technique. The SVM technique exhibited mean success rates of 59.75%, 72.89%, and 91.42% for classification of four, three and two mental tasks. The distinct emotions were categorized by a multiclass SVM scheme in [40]. This approach outperformed the existing techniques by providing 62.4% and 60.72% accuracy for valence and arousal dimensions when categorizing the emotions.

4.2.6. Neural Networks (NNs)

NNs are improved techniques that are exploited for classification in many BCI applications. In NNs, inputs known as neurons are connected with weights. The neurons with weights are processed using processing units. These units comprise a summation section that is ultimately linked to output [29]. Different types of NNs include Multilayer Perceptron (MLP) networks, Artificial Neural Networks (ANNs), and Convolutional Neural Networks (CNNs).

A. Multilayer Perceptron (MLP)

MLP comprises three layers, namely, input, output, and hidden. In MLP, the second input layer's output is fused to create another layer, and so on, till the problem is suitably classified. It employs a backpropagation method for training NNs.

The authors in [30,49,57] employed MLP for classifying EEG signals. The exploitation of different ML classifiers such as an autoencoder, SVM, logistic regression along MLP in [54] revealed that MLP displayed superior performance in EEG signal categorization. Results signified that the MLP offered $95.58 \pm 0.5\%$ accuracy.

B. Artificial Neural Networks (ANN)

ANNs comprise a group of neurons linked among others arranged in layers. The patterns which codify the tangible problem codification (c) are transmitted through layers and the data are transformed with respective synaptic weights (w). Furthermore, summation of these data is performed by neurons in the following layer for creating another input known as bias. The bias is a threshold that indicates the minimum level required for neuron activation and is denoted by θ. Equation (17) represents the summation function.

$$q = \sum_{i=1}^{N} c_i w_i + \theta \tag{17}$$

The summation result is then evaluated using transfer function $f(q)$ leading to output neuron. These data are then forwarded to other linked neurons, till the final layer is reached where the ANN output is obtained.

The authors in [44,56,69,76] employed ANNs for EEG signal categorization. It was noticed that the ANNs provided superior classification outputs than other compared techniques. In [44], ANNs trained using a classical backpropagation scheme was exploited

for categorizing EEG signals connected with diverse mental tasks such as math, baseline, figure rotation, visual counting, and letter composing. Data recorded while performing these tasks were divided into testing and training sets. One trail was employed for testing and the remaining 9 trails for training. The number of input layer neurons were varied as per the input vector's length. This technique offered 84% accuracy. In [59], ANNs were used for categorizing six distinct emotions, namely, satisfied, pleasant, happy, frustrated, sad, and fear along with different ML schemes such as KNN, naive Bayes, and SVM. The ANN structure employed 6 output and 10 hidden layers for classifying distinct emotional states. The KNN, naïve Bayes, SVM, and ANN achieved 49.82%, 50.7%, 51.82%, and 55.58% accuracies. Results signified that among the exploited ML schemes, ANN displayed good classification performance by providing greater accuracy.

C.   Convolutional Neural Networks (CNN)

In CNN, neurons are organized into three dimensions, namely, height, width, and depth. CNNs comprise a five-layered structure involving input layer, convolutional layer, linear unit layer, pooling, and fully connected layer [38,88]. Considering a CNN architecture comprising L layers, *an* input vector, M maps, b weight vector, ξ neurons, N_e electrodes, N_s signal values and N_p segments of signal values, then the layers 1, 2, 3, and 4 are represented using Equations (18)–(21) as

$$\sigma(1,\ M,\ \xi_1) = \sum_{i=0}^{N_e} a_{ij}\, b(1,M,i) + bias \rightarrow \tag{18}$$

where a_{ij} denotes input vector of layer L_0

$$\sigma(2,\ M,\ \xi_2)\ \rightarrow\ = \sum_{i=0}^{N_s/N_p} a\left(1,M,\xi * \left(N_s/N_p\right)\right) b(2,M,i) + bias \tag{19}$$

$$\sigma(2,\ \xi_3)\ =\ \rightarrow \sum_{i=0}^{M_2} \sum_{k=0}^{N_p} a\,(2,i,k)\, b(4,i,\ k)\ +\ bias \tag{20}$$

$$\sigma(4,\ \xi_4)\ =\ \rightarrow \sum_{i=0}^{\xi_3} a\,(3,i) b(4,i)\ +\ bias \tag{21}$$

CNN technique was used in [33] for MI classification. Using CNN, time-frequency data were extracted from receptive fields. Then, by employing spatial data, feature responses were recalibrated which further ameliorated the MI categorization performance. It was noted from the results that the exploited CNN offered 79.3% accuracy. The authors in [35,38,86] employed CNN for MI categorization. This approach improved accuracy for different BCI subjects. A deep CNN approach was used in [84] for MI classification. This approach effectively classified the raw MI-EEG waves of the right- and left-brain electrodes without any unnatural feature extraction or preprocessing operations. EEG features were learned using a 5-layered CNN, dimension reduction was carried out using a 4-layered max pooling and final classification was achieved using a fully-connected CNN layer. In group-wise classification, this approach exhibited accuracy of 95.76%, 97.28%, 96.01% and 94.80% for 10, 20, 60 and 100 subjects. In [88], convolutional neural networks (CNNs) were employed for classifying ERP signals. This approach outperformed the existing technique under two conditions: (a) with 10 epochs and (b) with 8 electrodes. In this work, seven CNN-based classifiers (i.e., three multi-classifiers and four individual classifiers with distinct feature sets) were proposed. Best results were obtained with multi-classifiers rather than individual classifiers. About 95.5% accuracy was achieved with the multi-classifier solution.

4.2.7. Other Techniques

A neuro-fuzzy classification approach was exploited for categorizing EEG signals in [39]. The EEG non-stationaries were handled using a self-regulatory gaussian fuzzy system. The self-regulatory learning method captured training samples as per the priority and further created, pruned, or upgraded fuzzy rules through scanning existing knowledge content in data patterns for rules' creation. This neuro-fuzzy scheme showed larger

performance improvement than existing techniques. It achieved 91.43% accuracy. A multi-class AdaBoost with ELM technique was exploited in [45] for EEG signal categorization. For classification, the weights between hidden and input neurons were chosen and fixed depending on some probability density functions. Then the weights between output and hidden neurons of the network were determined. Using Adaboost, training samples were fed every time to the weaker learner. Further, hardly categorized samples were allocated higher weights while easily categorized samples were allocated lower weights. After multiple rounds, the acquired poor prediction rules were linearly combined for ultimate classification. This approach displayed better mean categorization accuracy and greatly outperformed classical classification approaches. It provided 79.5% accuracy in EEG signal categorization. The authors in [62] adopted a linear least squares (LLS) classifier for categorizing EEG signals. The proposed LLS classifier displayed better categorization performance than LDA and SVM. A long short-term memory (LSTM) with an ensemble learning approach was exploited in [72] for classifying EEG waves. A classy activation function was introduced into conventional LSTM for maximizing the training procedure and lowering the vanishing gradient effect. The bagging method was then applied to LSTM for optimizing classification performance. The proposed ensemble-LSTM method displayed 97.13% of EEG categorization accuracy and greatly outperformed the classical LSTM method. In [81], the sparse Bayesian (SB) technique with Laplace priors was employed for classifying EEG signals. Utilizing the Laplace prior, the discriminant vectors were learned and all needed parameters were estimated automatically using training data. The results signified that the SBLaplace method outperformed the competing techniques for EEG signal categorization particularly for limited sample size environments. The signals were categorized using a stacked autoencoder (SA), CNN, and a fused SA-CNN scheme in [83]. In the fused SA-CNN scheme, the input EEG features were initially extracted through training convolutional layer filters of CNN. Later, these features were fed to SA for classification. Experimental outputs reported that the fused SA-CNN scheme provided better EEG-MI signal categorization results than the individual SA and CNN techniques. Moreover, the fused SA-CNN scheme offered 90% accuracy and showed 9% improvement in kappa value when compared to individual SA and CNN methods. The authors in [85] performed MI classification through the deep metric learning (DML) method. Convolutional encoders were employed for training a feature space in the training phase wherein spectrogram features of identical subjects are clustered. Another CNN encoder was adopted in the inference phase for comparing the EEG spectrogram's features. The nearest neighbor method was then exploited for classifying the EEG spectrogram category. The proposed DML method exhibited 64.7% accuracy in MI classification. Moreover, this approach performed classification effectively with only a few training samples compared to existing techniques. Mental state categorization through EEG-dependent BCI was presented in [47]. Different classifiers such as SVM, Bayesian networks, and random forest were exploited for classifying diverse mental states such as concentrated, relaxed, and neutral conditions. Among the employed classifiers, random forest showed superior performance and provided 87.16% accuracy. However, it achieved similar accuracy as that of BPNN and SVM. A spatio-temporal discriminant analysis (STDA) was exploited in [91] for classifying ERP signals. The discriminant data between non-target and target classes were optimized through determining projection matrices from temporal and spatial dimensions jointly which further helped in training sample minimization and feature dimensionality reduction. The results signified that the STDA method offered $80.0 \pm 6.32\%$ accuracy.

4.3. Comparative Study

The number of recent papers (i.e., from 2016 to 2021) on BCI-related tasks reviewed in Section 4.3 are depicted in Figure 4. In this survey, a greater number of papers published in 2020 (13 papers) are considered for acquiring data about the latest progress attained in BCIs.
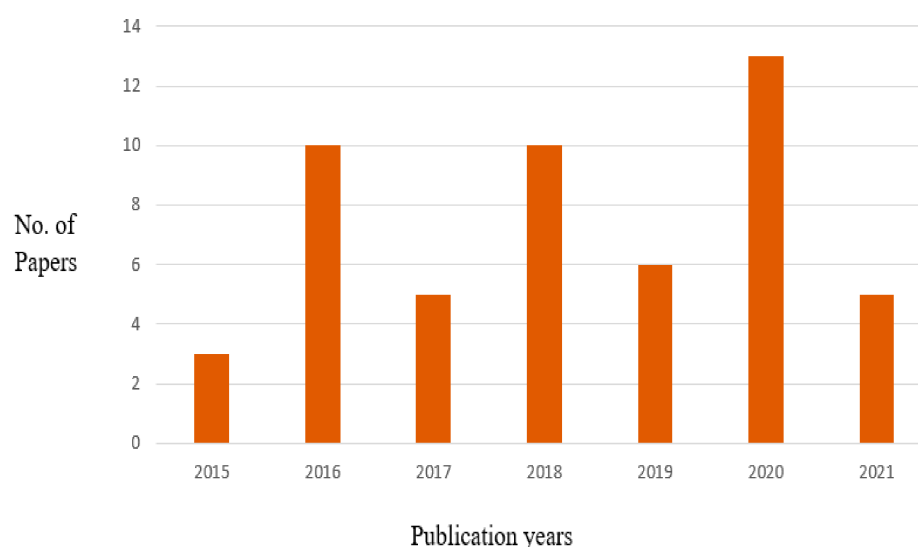
Figure 4. Publication year analysis.

4.4. Findings of the Research

Reviewing existing ML methods for feature extraction, feature selection, and classification in Sections 4.1–4.3, it could be observed that the performance and efficacy of ML methods must be further optimized with reduced training and computational complexities without compromising accuracy for acquiring anticipated and precise outcomes concerning BCI-oriented applications. Better methods for reducing over-fitting issues in neural networks and ameliorating the classification accuracy in distinct BCI databases concerning distinct BCI tasks are needed for satisfying the multifarious demands of BCI users.

5. Conclusions

This review work has presented a profound study of BCI and explained the scope of ML technology in BCI. It has reviewed various BCI-related studies and described the role played by ML towards diverse BCI tasks such as MI classification, emotion classification, and mental state and mental task classification. It has explored diverse feature extraction, selection, and classification schemes exploited in the literature for classification of EEG, ERP signals, mental state, mental task, limb motion, emotion, and MI. It has reviewed and tabulated various ML schemes and other techniques employed by researchers for diverse BCI tasks for providing more knowledge and encouraging improved ML-based BCI applications in the future.

BCI operations are linked with numerous tasks. Therefore, categorization of only EEG/ERP signals linked with a single task such as either MI or emotions or mental states is inadequate for developing contemporary BCI applications. Improved ML frameworks are thus desired for determining EEG, ERP, and other sorts of biomedical signals linked with multiple and discriminative tasks within one framework simultaneously with greater precision level. Moreover, ML models with shorter training time, lesser training samples without immolating classification quality, sensitivity, and computational complexity are strongly needed for existing and looming BCI applications.

References

1. McFarland, D.J.; Wolpaw, J.R. Brain-computer interfaces for communication and control. *Commun. ACM* **2011**, *54*, 60–66. [CrossRef] [PubMed]
2. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain computer interfaces, a review. *Sensors* **2012**, *12*, 1211–1279. [CrossRef] [PubMed]
3. Lv, Z.; Qiao, L.; Wang, Q.; Piccialli, F. Advanced Machine-Learning Methods for Brain-Computer Interfacing. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**, *18*, 1688–1698. [CrossRef] [PubMed]
4. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [CrossRef]
5. Chaudhary, U.; Birbaumer, N.; Ramos-Murguialday, A. Brain–computer interfaces for communication and rehabilitation. *Nat. Rev. Neurol.* **2016**, *12*, 513–525. [CrossRef] [PubMed]
6. Lazarou, I.; Nikolopoulos, S.; Petrantonakis, P.C.; Kompatsiaris, I.; Tsolaki, M. EEG-Based Brain-Computer Interfaces for Communication and Rehabilitation of People with Motor Impairment: A Novel Approach of the 21st Century. *Front. Hum. Neurosci.* **2018**, *12*, 14. [CrossRef] [PubMed]
7. Yi, H. Efficient machine learning algorithm for electroencephalogram modeling in brain–computer interfaces. *Neural Comput. Appl.* **2020**, 1–11. Available online: https://link.springer.com/article/10.1007/s00521-020-04861-3#citeas (accessed on 10 August 2021).
8. Mak, J.N.; Arbel, Y.; Minett, J.W.; McCane, L.M.; Yuksel, B.; Ryan, D.; Erdogmus, D. Optimizing the P300-based brain–computer interface: Current status, limitations and future directions. *J. Neural Eng.* **2011**, *8*, 025003. [CrossRef]
9. Millán, J.d.R.; Rupp, R.; Müller-Putz, G.; Murray-Smith, R.; Giugliemma, C.; Tangermann, M.; Leeb, R. Combining Brain-Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges. *Front. Neurosci.* **2010**, *4*, 161. [CrossRef] [PubMed]
10. Hwang, H.-J.; Kim, S.; Choi, S.; Im, C.-H. EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey. *Int. J. Hum. Comput. Interact.* **2013**, *29*, 814–826. [CrossRef]
11. Gürkök, H.; Nijholt, A. Brain-Computer Interfaces for Multimodal Interaction: A Survey and Principles. *Int. J. Hum. Comput. Interact.* **2012**, *28*, 292–307. [CrossRef]
12. Abdulkader, S.N.; Atia, A.; Mostafa, M.S.M. Brain computer interfacing: Applications and challenges. *Egypt. Inform. J.* **2015**, *16*, 213–230. [CrossRef]
13. Lotte, F.; Congedo, M.; Lécuyer, A.; Lamarche, F.; Arnaldi, B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* **2007**, *4*, 24. [CrossRef] [PubMed]
14. Abibullaev, B.; Zollanvari, A. Learning discriminative spatiospectral features of ERPs for accurate brain–computer interfaces. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 2009–2020. [CrossRef] [PubMed]
15. Müller, K.R.; Tangermann, M.; Dornhege, G.; Krauledat, M.; Curio, G.; Blankertz, B. Machine learning for real-time single-trial EEG-analysis: From brain–computer interfacing to mental state monitoring. *J. Neurosci. Methods* **2008**, *167*, 82–90. [CrossRef] [PubMed]
16. Lotte, F.; Bougrain, L.; Cichocki, A.; Clerc, M.; Congedo, M.; Rakotomamonjy, A.; Yger, F. A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update. *J. Neural Eng.* **2018**, *15*, 031005. [CrossRef]
17. Fouad, M.M.; Amin, K.M.; El-Bendary, N.; Hassanien, A.E. Brain computer interface: A review. *Brain-Comput. Interfaces* **2015**, *74*, 3–30.
18. Ramadan, R.A.; Refat, S.; Elshahed, M.A.; Ali, R.A. Basics of brain computer interface. In *Brain-Computer Interfaces*; Springer: Cham, Germany, 2015; pp. 31–50.
19. Combaz, A.; Chumerin, N.; Manyakov, N.V.; Robben, A.; Suykens, J.A.; Van Hulle, M.M. Towards the detection of error-related potentials and its integration in the context of a P300 speller brain–computer interface. *Neurocomputing* **2012**, *80*, 73–82. [CrossRef]
20. Padfield, N.; Zabalza, J.; Zhao, H.; Masero, V.; Ren, J. EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges. *Sensors* **2019**, *19*, 1423. [CrossRef] [PubMed]
21. Blankertz, B.; Tangermann, M.; Vidaurre, C.; Dickhaus, T.; Sannelli, C.; Popescu, F.; Müller, K.R. Detecting mental states by machine learning techniques: The Berlin brain–computer interface. In *Brain-Computer Interfaces*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 113–135.
22. Torres, E.P.; Torres, E.A.; Hernández-Álvarez, M.; Yoo, S.G. EEG-Based BCI Emotion Recognition: A Survey. *Sensors* **2020**, *20*, 5083. [CrossRef] [PubMed]
23. Jiang, X.; Gu, X.; Xu, K.; Ren, H.; Chen, W. Independent decision path fusion for bimodal asynchronous brain–computer interface to discriminate multiclass mental states. *IEEE Access* **2019**, *7*, 165303–165317. [CrossRef]
24. Mohanty, R.; Sinha, A.M.; Remsik, A.B.; Dodd, K.C.; Young, B.M.; Jacobson, T.; Prabhakaran, V. Machine learning classification to identify the stage of brain-computer interface therapy for stroke rehabilitation using functional connectivity. *Front. Neurosci.* **2018**, *12*, 353. [CrossRef]
25. Bablani, A.; Edla, D.R.; Tripathi, D.; Cheruku, R. Survey on Brain-Computer Interface: An Emerging Computational Intelligence Paradigm. *ACM Comput. Surv.* **2019**, *52*, 1–32. [CrossRef]
26. Chakladar, D.D.; Chakraborty, S. Feature extraction and classification in brain-computer interfacing: Future research issues and challenges. In *Natural Computing for Unsupervised Learning*; Springer: Cham, Germany, 2019; pp. 101–131.
27. Salazar-Varas, R.; Gutiérrez, D. An optimized feature selection and classification method for using electroencephalographic coherence in brain–computer interfaces. *Biomed. Signal Process. Control.* **2015**, *18*, 11–18. [CrossRef]

28. Sun, S.; Zhang, C. Adaptive feature extraction for EEG signal classification. *Med. Biol. Eng. Comput.* **2006**, *44*, 931–935. [CrossRef] [PubMed]

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

30. Ahangi, A.; Karamnejad, M.; Mohammadi, N.; Ebrahimpour, R.; Bagheri, N. Multiple classifier system for EEG signal classification with application to brain–computer interfaces. *Neural Comput. Appl.* **2013**, *23*, 1319–1327. [CrossRef]

31. Mirzaei, S.; Ghasemi, P. EEG motor imagery classification using dynamic connectivity patterns and convolutional autoencoder. *Biomed. Signal Process. Control.* **2021**, *68*, 102584. [CrossRef]

32. Chen, J.; Yu, Z.; Gu, Z.; Li, Y. Deep Temporal-Spatial Feature Learning for Motor Imagery-Based Brain–Computer Interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 2356–2366. [CrossRef]

33. Sun, B.; Zhao, X.; Zhang, H.; Bai, R.; Li, T. EEG Motor Imagery Classification with Sparse Spectrotemporal Decomposition and Deep Learning. *IEEE Trans. Autom. Sci. Eng.* **2020**, *18*, 541–551. [CrossRef]

34. Burke, D.P.; Kelly, S.P.; De Chazal, P.; Reilly, R.B.; Finucane, C. A parametric feature extraction and classification strategy for brain-computer interfacing. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2005**, *13*, 12–17. [CrossRef] [PubMed]

35. Majidov, I.; Whangbo, T. Efficient Classification of Motor Imagery Electroencephalography Signals Using Deep Learning Methods. *Sensors* **2019**, *19*, 1736. [CrossRef]

36. Alomari, M.H.; Samaha, A.; AlKamha, K. Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning. *Int. J. Adv. Comput. Sci. Appl.* **2013**, *4*, 207–212.

37. Atmaji, C.; Putra, A.E.; Tontowi, I.A. Three-Class Classification of EEG Signals Using Support Vector Machine Methods. In Proceedings of the 2018 4th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia, 7–8 August 2018; pp. 1–4.

38. Chaudhary, S.; Taran, S.; Bajaj, V.; Sengur, A. Convolutional Neural Network Based Approach Towards Motor Imagery Tasks EEG Signals Classification. *IEEE Sens. J.* **2019**, *19*, 4494–4500. [CrossRef]

39. Jafarifarmand, A.; Badamchizadeh, M.A.; Khanmohammadi, S.; Nazari, M.A.; Tazehkand, B.M. A New Self-Regulated Neuro-Fuzzy Framework for Classification of EEG Signals in Motor Imagery BCI. *IEEE Trans. Fuzzy Syst.* **2018**, *26*, 1485–1497. [CrossRef]

40. Atkinson, J.; Campos, D. Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Syst. Appl.* **2016**, *47*, 35–41. [CrossRef]

41. Hsu, W.Y. Continuous EEG signal analysis for asynchronous BCI application. *Int. J. Neural Syst.* **2011**, *21*, 335–350. [CrossRef]

42. Venkatachalam, K.; Devipriya, A.; Maniraj, J.; Sivaram, M.; Ambikapathy, A.; Iraj, S.A. A novel method of motor imagery classification using eeg signal. *Artif. Intell. Med.* **2020**, *103*, 101787.

43. Sreeja, S.R.; Rabha, J.; Nagarjuna, K.Y.; Samanta, D.; Mitra, P.; Sarma, M. Motor imagery EEG signal processing and classification using machine learning approach. In Proceedings of the 2017 International Conference on New Trends in Computing Sciences (ICTCS), Amman, Jordan, 11–13 October 2017; pp. 61–66.

44. El Bahy, M.M.; Hosny, M.; Mohamed, W.A.; Ibrahim, S. EEG signal classification using neural network and support vector machine in brain computer interface. In *International Conference on Advanced Intelligent Systems and Informatics*; Springer: Cham, Germany, 2016; pp. 246–256.

45. Gao, L.; Cheng, W.; Zhang, J.; Wang, J. EEG classification for motor imagery and resting state in BCI applications using multi-class Adaboost extreme learning machine. *Rev. Sci. Instrum.* **2016**, *87*, 085110. [PubMed]

46. Rakshit, A.; Khasnobish, A.; Tibarewala, D.N. A Naïve Bayesian approach to lower limb classification from EEG signals. In Proceedings of the 2016 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC), Kolkata, India, 28–30 January 2016; pp. 140–144.

47. Bird, J.J.; Manso, L.J.; Ribeiro, E.P.; Ekart, A.; Faria, D.R. A study on mental state classification using eeg-based brain-machine interface. In Proceedings of the 2018 International Conference on Intelligent Systems (IS), Funchal, Portugal, 25–27 September 2018; pp. 795–800.

48. Liang, N.Y.; Saratchandran, P.; Huang, G.B.; Sundararajan, N. Classification of mental tasks from EEG signals using extreme learning machine. *Int. J. Neural Syst.* **2006**, *16*, 29–38. [CrossRef] [PubMed]

49. Afrakhteh, S.; Mosavi, M.R.; Khishe, M.; Ayatollahi, A. Accurate classification of EEG signals using neural networks trained by hybrid population-physic-based algorithm. *Int. J. Autom. Comput.* **2020**, *17*, 108–122. [CrossRef]

50. Yasoda, K.; Ponmagal, R.S.; Bhuvaneshwari, K.S.; Venkatachalam, K. Automatic detection and classification of EEG artifacts using fuzzy kernel SVM and wavelet ICA (WICA). *Soft Comput.* **2020**, *24*, 16011–16019. [CrossRef]

51. El-Kafrawy, N.M.; Hegazy, D.; Tolba, M.F. Features extraction and classification of eeg signals using empirical mode decomposition and support vector machine. In *International Conference on Advanced Machine Learning Technologies and Applications*; Springer: Cham, Germany, 2014; pp. 189–198.

52. Chaudhary, S.; Taran, S.; Bajaj, V.; Siuly, S. A flexible analytic wavelet transform based approach for motor-imagery tasks classification in BCI applications. *Comput. Methods Programs Biomed.* **2020**, *187*, 105325. [CrossRef] [PubMed]

53. Ren, W.; Han, M. Classification of EEG signals using hybrid feature extraction and ensemble extreme learning machine. *Neural Process. Lett.* **2019**, *50*, 1281–1301. [CrossRef]

54. Ieracitano, C.; Mammone, N.; Hussain, A.; Morabito, F.C. A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Netw.* **2020**, *123*, 176–190. [CrossRef] [PubMed]

55. Xu, Q.; Zhou, H.; Wang, Y.; Huang, J. Fuzzy support vector machine for classification of EEG signals using wavelet-based features. *Med. Eng. Phys.* **2009**, *31*, 858–865. [CrossRef]

56. Anh, N.T.H.; Hoang, T.H.; Thang, V.T.; Bui, T.Q. An artificial neural network approach for electroencephalographic signal classification towards brain-computer interface implementation. In Proceedings of the 2016 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), Hanoi, Vietnam, 7–9 November 2016; pp. 205–210.

57. Ilyas, M.Z.; Saad, P.; Ahmad, M.I.; Ghani, A.R.I. Classification of EEG signals for brain-computer interface applications: Performance comparison. In Proceedings of the 2016 International Conference on Robotics, Automation and Sciences (ICORAS), Melaka, Malaysia, 5–6 November 2016; pp. 1–4.

58. Aydemir, O.; Kayikcioglu, T. Decision tree structure based classification of EEG signals recorded during two dimensional cursor movement imagery. *J. Neurosci. Methods* **2014**, *229*, 68–75. [CrossRef]

59. Mohammadpour, M.; Hashemi, S.M.R.; Houshmand, N. Classification of EEG-based emotion for BCI applications. In Proceedings of the 2017 Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran, 9 April 2017; pp. 127–131.

60. Wang, X.W.; Nie, D.; Lu, B.L. Emotional state classification from EEG data using machine learning approach. *Neurocomputing* **2014**, *129*, 94–106. [CrossRef]

61. Rashid, M.; Sulaiman, N.; Mustafa, M.; Khatun, S.; Bari, B.S. The classification of EEG signal using different machine learning techniques for BCI application. In *International Conference on Robot Intelligence Technology and Applications*; Springer: Singapore, 2018; pp. 207–221.

62. Rodríguez-Bermúdez, G.; García-Laencina, P.J. Automatic and adaptive classification of electroencephalographic signals for brain computer interfaces. *J. Med. Syst.* **2012**, *36*, 51–63.

63. Suk, H.I.; Lee, S.W. A novel Bayesian framework for discriminative feature extraction in brain-computer interfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 286–299. [CrossRef]

64. Shahlaei, F.; Bagh, N.; Zambare, M.S.; Machireddy, R.; Shaligram, A.D. Detection of Event Related Patterns using Hilbert Transform in Brain Computer Interface. In Proceedings of the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Edmonton, AB, Canada, 5–8 May 2019; pp. 1–4.

65. Miao, M.; Zeng, H.; Wang, A.; Zhao, C.; Liu, F. Discriminative spatial-frequency-temporal feature extraction and classification of motor imagery EEG: An sparse regression and Weighted Naïve Bayesian Classifier-based approach. *J. Neurosci. Methods* **2017**, *278*, 13–24. [CrossRef] [PubMed]

66. Kuncheva, L.I.; Rodríguez, J.J. Interval feature extraction for classification of event-related potentials (ERP) in EEG data analysis. *Progress Artif. Intell.* **2013**, *2*, 65–72. [CrossRef]

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

68. Appriou, A.; Cichocki, A.; Lotte, F. Modern Machine-Learning Algorithms: For Classifying Cognitive and Affective States from Electroencephalography Signals. *IEEE Syst. Man Cybern. Mag.* **2020**, *6*, 29–38. [CrossRef]

69. Poorna, S.S.; Baba, P.S.; Ramya, G.L.; Poreddy, P.; Aashritha, L.S.; Nair, G.J.; Renjith, S. Classification of EEG based control using ANN and KNN—A comparison. In Proceedings of the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, India, 15–17 December 2016; pp. 1–6.

70. Thulasidas, M.; Guan, C.; Wu, J. Robust classification of EEG signal for brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2006**, *14*, 24–29. [CrossRef] [PubMed]

71. Sun, S.; Zhang, C.; Zhang, D. An experimental evaluation of ensemble methods for EEG signal classification. *Pattern Recognit. Lett.* **2007**, *28*, 2157–2163. [CrossRef]

72. Zheng, X.; Chen, W.; You, Y.; Jiang, Y.; Li, M.; Zhang, T. Ensemble deep learning for automated visual classification using EEG signals. *Pattern Recognit.* **2020**, *102*, 107147. [CrossRef]

73. Aayesha, Q.M.B.; Afzaal, M.; Qureshi, M.S.; Fayaz, M. Machine learning-based EEG signals classification model for epileptic seizure detection. *Multimed. Tools Appl.* **2021**, *80*, 17849–17877. [CrossRef]

74. Aliakbaryhosseinabadi, S.; Kamavuako, E.N.; Jiang, N.; Farina, D.; Mrachacz-Kersting, N. Classification of EEG signals to identify variations in attention during motor task execution. *J. Neurosci. Methods* **2017**, *284*, 27–34. [CrossRef] [PubMed]

75. Bayram, K.S.; Kızrak, M.A.; Bolat, B. Classification of EEG signals by using support vector machines. In Proceedings of the 2013 IEEE INISTA, Albena, Bulgaria, 19–21 June 2013; pp. 1–3.

76. Vazquez, R.A.; Salazar-Varas, R. Classification of EEG signals using fractal dimension features and artificial neural networks. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017 2017; pp. 1–6.

77. She, Q.; Hu, B.; Gan, H.; Fan, Y.; Nguyen, T.; Potter, T.; Zhang, Y. Safe semi-supervised extreme learning machine for EEG signal classification. *IEEE Access* **2018**, *6*, 49399–49407. [CrossRef]

78. Li, X.; Chen, X.; Yan, Y.; Wei, W.; Wang, Z.J. Classification of EEG signals using a multiple kernel learning support vector machine. *Sensor* **2014**, *14*, 12784–12802. [CrossRef] [PubMed]

79. Bhuvaneswari, P.; Kumar, J.S. Support vector machine technique for EEG signals. *Int. J. Comput. Appl.* **2013**, *63*, 1–5. [CrossRef]

80. Tan, P.; Sa, W.; Yu, L. Applying extreme learning machine to classification of EEG BCI. In Proceedings of the 2016 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Chengdu, China, 19–22 June 2016; pp. 228–232.

81. Zhang, Y.; Zhou, G.; Jin, J.; Zhao, Q.; Wang, X.; Cichocki, A. Sparse Bayesian classification of EEG for brain–computer interface. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 2256–2267. [CrossRef] [PubMed]

82. Jin, Z.; Zhou, G.; Gao, D.; Zhang, Y. EEG classification using sparse Bayesian extreme learning machine for brain–computer interface. *Neural Comput. Appl.* **2020**, *32*, 6601–6609. [CrossRef]

83. Tabar, Y.R.; Halici, U. A novel deep learning approach for classification of EEG motor imagery signals. *J. Neural Eng.* **2017**, *14*, 16003. [CrossRef] [PubMed]

84. Lun, X.; Yu, Z.; Chen, T.; Wang, F.; Hou, Y. A simplified CNN classification method for MI-EEG via the electrode pairs signals. *Front. Hum. Neurosci.* **2020**, *14*, 338. [CrossRef]

85. Alwasiti, H.; Yusoff, M.Z.; Raza, K. Motor imagery classification for brain computer interface using deep metric learning. *IEEE Access* **2020**, *8*, 109949–109963. [CrossRef]

86. Dose, H.; Møller, J.S.; Iversen, H.K.; Puthusserypady, S. An end-to-end deep learning approach to MI-EEG signal classification for BCIs. *Expert Syst. Appl.* **2018**, *114*, 532–542. [CrossRef]

87. He, L.; Hu, D.; Wan, M.; Wen, Y.; Von Deneen, K.M.; Zhou, M. Common Bayesian network for classification of EEG-based multiclass motor imagery BCI. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *46*, 843–854. [CrossRef]

88. Cecotti, H.; Graser, A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 433–445. [CrossRef] [PubMed]

89. Blankertz, B.; Lemm, S.; Treder, M.; Haufe, S.; Müller, K.R. Single-trial analysis and classification of ERP components—a tutorial. *NeuroImage* **2011**, *56*, 814–825. [CrossRef] [PubMed]

90. Zhang, Y.; Zhou, G.; Jin, J.; Zhao, Q.; Wang, X.; Cichocki, A. Aggregation of sparse linear discriminant analyses for event-related potential classification in brain-computer interface. *Int. J. Neural Syst.* **2014**, *24*, 1450003. [CrossRef]

91. Zhang, Y.; Zhou, G.; Zhao, Q.; Jin, J.; Wang, X.; Cichocki, A. Spatial-temporal discriminant analysis for ERP-based brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *21*, 233–243. [CrossRef] [PubMed]

92. Hortal, E.; Iáñez, E.; Úbeda, A.; Planelles, D.; Costa, Á.; Azorín, J.M. Selection of the best mental tasks for a svm-based bci system. In Proceedings of the 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), San Diego, CA, USA, 5–8 October 2014; pp. 1483–1488.

93. Ieracitano, C.; Mammone, N.; Hussain, A.; Morabito, F.C. A novel explainable machine learning approach for EEG-based brain-computer interface systems. *Neural Comput. Appl.* **2021**, 1–14. [CrossRef]

94. Hong, K.S.; Khan, M.J.; Hong, M.J. Feature extraction and classification methods for hybrid fNIRS-EEG brain-computer interfaces. *Front. Hum. Neurosci.* **2018**, *12*, 246. [CrossRef]

95. Zhang, H.; Yang, H.; Guan, C. Bayesian learning for spatial filtering in an EEG-based brain–computer interface. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1049–1060. [CrossRef] [PubMed]

96. Vidaurre, C.; Kawanabe, M.; von Bünau, P.; Blankertz, B.; Müller, K.R. Toward unsupervised adaptation of LDA for brain–computer interfaces. *IEEE Trans. Biomed. Eng.* **2010**, *58*, 587–597. [CrossRef] [PubMed]

97. Wu, Q.; Zhang, Y.; Liu, J.; Sun, J.; Cichocki, A.; Gao, F. Regularized group sparse discriminant analysis for P300-based brain–computer interface. *Int. J. Neural Syst.* **2019**, *29*, 1950002. [CrossRef]

98. Kundu, S.; Ari, S. P300 detection with brain–computer interface application using PCA and ensemble of weighted SVMs. *IETE J. Res.* **2018**, *64*, 406–414. [CrossRef]

99. Bhatnagar, V.; Yede, N.; Keram, R.S.; Chaurasiya, R.K. A modified approach to ensemble of SVM for P300 based brain computer interface. In Proceedings of the 2016 International Conference on Advances in Human Machine Interaction (HMI), Kodigehalli, India, 3–5 March 2016; pp. 1–5.

100. Manyakov, N.V.; Chumerin, N.; Combaz, A.; Van Hulle, M.M. Comparison of classification methods for P300 brain-computer interface on disabled subjects. *Comput. Intell. Neurosci.* **2011**, 519868. [CrossRef] [PubMed]

101. Ramoser, H.; Muller-Gerking, J.; Pfurtscheller, G. Optimal spatial filtering of EEG during imagined hand movement. *IEEE Trans. Rehabil. Engi.* **2000**, *8*, 441–446. [CrossRef] [PubMed]