

# Neural Decoding of EEG Signals with Machine Learning: A Systematic Review

Maham Saeidi <sup>1,\*</sup>, Waldemar Karwowski <sup>1,\*</sup>, Farzad V. Farahani <sup>1,2</sup>, Krzysztof Fiok <sup>1</sup>, Redha Tair <sup>3</sup>, P. A. Hancock <sup>4</sup> and Awad Al-Juaid <sup>5</sup>

- <sup>1</sup> Computational Neuroergonomics Laboratory, Department of Industrial Engineering and Management Systems, University of Central Florida, Orlando, FL 32816, USA; ffaraha2@jhu.edu (F.V.F.); fiok@ucf.edu (K.F.)  
<sup>2</sup> Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21218, USA  
<sup>3</sup> MATIM, Moulin de la Housse, Université de Reims Champagne Ardenne, CEDEX 02, 51687 Reims, France; redha.tair@univ-reims.fr  
<sup>4</sup> Department of Psychology, University of Central Florida, Orlando, FL 32816, USA; peter.hancock@ucf.edu  
<sup>5</sup> Industrial Engineering Department, Taif University, Taif 26571, Saudi Arabia; amjuaid@tu.edu.sa  
\* Correspondence: msaeidi@knights.ucf.edu (M.S.); wkar@ucf.edu (W.K.)

**Abstract:** Electroencephalography (EEG) is a non-invasive technique used to record the brain's evoked and induced electrical activity from the scalp. Artificial intelligence, particularly machine learning (ML) and deep learning (DL) algorithms, are increasingly being applied to EEG data for pattern analysis, group membership classification, and brain-computer interface purposes. This study aimed to systematically review recent advances in ML and DL supervised models for decoding and classifying EEG signals. Moreover, this article provides a comprehensive review of the state-of-the-art techniques used for EEG signal preprocessing and feature extraction. To this end, several academic databases were searched to explore relevant studies from the year 2000 to the present. Our results showed that the application of ML and DL in both mental workload and motor imagery tasks has received substantial attention in recent years. A total of 75% of DL studies applied convolutional neural networks with various learning algorithms, and 36% of ML studies achieved competitive accuracy by using a support vector machine algorithm. Wavelet transform was found to be the most common feature extraction method used for all types of tasks. We further examined the specific feature extraction methods and end classifier recommendations discovered in this systematic review.

**Keywords:** brain signals classification; EEG; deep learning; machine learning; review



**Citation:** Saeidi, M.; Karwowski, W.; Farahani, F.V.; Fiok, K.; Tair, R.; Hancock, P.A.; Al-Juaid, A. Neural Decoding of EEG Signals with Machine Learning: A Systematic Review. *Brain Sci.* **2021**, *11*, 1525. <https://doi.org/10.3390/brainsci11111525>

Academic Editor: Natsue Yoshimura

Received: 1 October 2021

Accepted: 11 November 2021

Published: 18 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The human brain is a complex system containing approximately 100 billion neurons and trillions of synaptic connections [1,2]. The brain's electrical activity became a research focus in the 19th century when Richard Caton recorded brain signals from rabbits [3,4]. Brain recordings were also performed by Hans Berger, the first person to record electroencephalogram (EEG) signals from the human scalp [5]. EEG-based research has since increased significantly, and EEG is now the most commonly used noninvasive tool to study dynamic signatures in the human brain [6,7]. EEG signals measure voltage fluctuations at the scalp and reflect the instantaneous superposition of electric dipoles, primarily from dendritic inputs to large pyramidal cells in the neuropil [8]. Signals traveling in white matter have traditionally been thought to be too fast to superimpose temporally, although recent cable theoretic models [9] and empirical work [10] suggest that white matter may also contribute to brain rhythms measured at the scalp. Classically, the three primary forms of the brain's activity based on EEG signals are brain waves, event-related potential (ERP), and steady-state visual evoked potentials (SSVEPs). Among those, brain waves are most commonly used in EEG signal analysis for different types of tasks. Brain waves have been categorized in terms of five frequency bands: delta, 0.5–4 Hz; theta, 4–8 Hz; alpha, 8–13 Hz;

beta, 13–30 Hz; and gamma, 30–150 Hz [11]. Other classifications of brain signals can be found in previous publications [12,13].

EEG is a low cost, noninvasive neuroimaging technique that provides high temporal resolution recordings of dynamic signatures in the brain; it has therefore become an indispensable tool in a variety of applications, including clinical diagnosis of a range of epileptic seizure types [14,15], brain-computer interface (BCI) systems, sleep analysis, and decoding behavioral activity [6,16,17]. When interpreted carefully [18], classification tools can be used not only for prediction but also to gain neuroscientific knowledge. However, EEG signals are complex, high-dimensional [19] and non-stationary, and have a low signal-to-noise ratio in the temporal domain. Therefore, careful preprocessing is often required to remove artifacts [20], particularly when EEG data are collected concurrently and in the MRI setting [21].

Although EEG has been demonstrated to be a valuable tool for research in various applications, it has several limitations, such as a low signal-to-noise ratio [22,23], nonlinearity and nonstationary properties [24,25], and inter-individual variability [26] which affect analysis and processing performance. To address these limitations, EEG signal processing pipelines are often used. The general EEG classification pipeline includes data preprocessing, initializing the classification procedure, splitting the data set for the classifier, predicting the class of new data, and evaluating the classification model for the test data set [27]. As shown in references [28,29], Riemannian geometry-based classifiers and adaptive classifiers have achieved success in classifying EEG signals. Machine learning (ML) and deep learning (DL) methods have become rapidly growing areas with applications in computational neuroscience, owing to higher levels of neural data analysis efficiency and decoding brain function [30]. In ML and DL, various algorithms are used to simplify processing pipelines and improve the learning process. For instance, supervised ML algorithms first learn on training data. The model and learned parameters are then applied to unseen or new data to predict the class label of the new data [27]. Among different types of classification tasks, binary and multi-label classifications are widely used in clinical studies, and in studies of cognitive function, motor imagery (MI) processing, emotion recognition (ER), and brain disorders, including brain injury, attention disorders, and multiple sclerosis [31–34]. In addition to these recognized areas, various computational neural models have been considered, such as the medial prefrontal cortex (mPFC) and anterior cingulate cortex in modeling of learning predictions and monitoring behavior [35]. A recent computational model, the predicted response-outcome model [36] is based on the hypothesis that mPFC stores predictions of future outcomes. According to this model, Garofalo et al. [35] have proposed that mediofrontal ERP signals of prediction errors are modulated by the likelihood of occurrence during a task. Furthermore, reference [37] has identified a causal role of the ventromedial prefrontal cortex (vmPFC) in the acquisition of fear conditioning in the definition of stimulus-outcome contingencies.

Because of their outstanding robustness and adaptability, several ML and DL models for performing EEG signal classification have been reported. The primary purpose of this study was to review and explore the recent advances over the past two decades in the deployment of supervised ML algorithms as well as DL models for the classification of EEG signals. We investigated the overall trends and ML and DL models used in individual studies and compared classification algorithms based on different tasks. We propose six different categories of tasks on the basis of the studies reviewed herein. We attempted to address defined research questions concerning each of these categories. According to this evaluation, we provide recommendations for choosing a suitable classifier for use in future applications. The remaining sections of this manuscript are organized as follows. In Section 2, we describe the method used to identify and select publications for subsequent inclusion in our synthesis of research studies. In Section 3, we describe general steps to generate an EEG-based BCI system. We review the typical pipeline used to construct EEG classifications, including data acquisition, artifact removal, feature extraction, and classification. In Section 4, we provide a general overview of the literature search,

study characteristics, validity assessments, and the main findings of each of the articles reviewed. In Section 5, we discuss applications of ML and DL techniques for different types of tasks, and provide recommendations regarding the selection of an effective classifier for each task. Finally, Section 6 highlights future perspectives in this field.

## 2. Materials and Methods

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), a standard systematic review and meta-analysis guideline [38], was used in this study. An important component of this systematic review involved the clear definition of research questions to reduce the effects of research expectations. Furthermore, our research method followed the Cochrane Collaboration definitions [39] to minimize the risk of bias.

### 2.1. Research Questions (RQs)

- RQ1: What classification tasks have received the most attention with the introduction of ML and DL algorithms and the use of EEG brain data?
- RQ2: Which feature extraction methods were used for each task to extract appropriate inputs for ML and DL classifiers?
- RQ3: What are the most frequently used ML and DL algorithms for EEG data processing?
- RQ4: Which specific ML and DL models are suitable for classifying EEG data involving different types of tasks?

### 2.2. Search Strategy

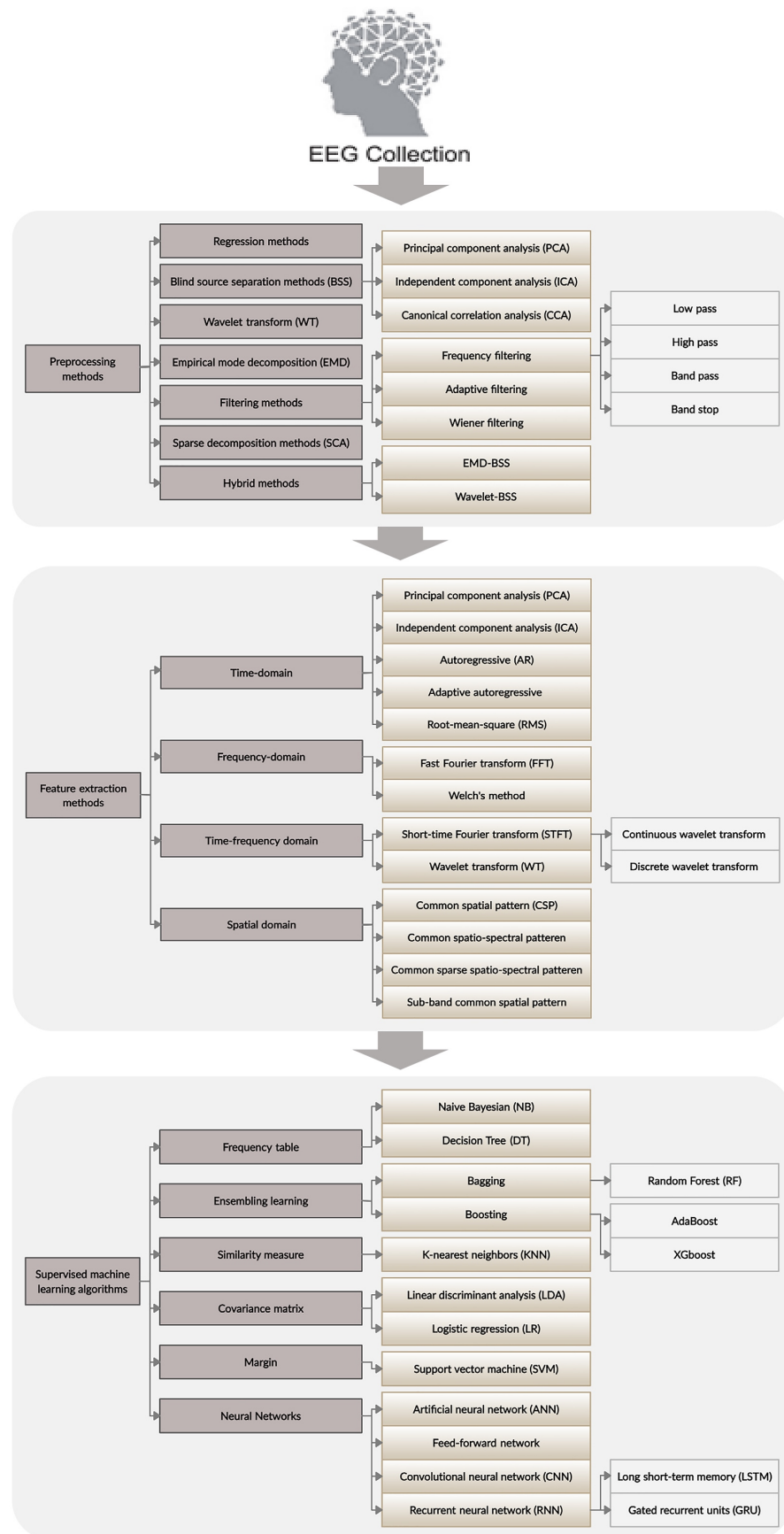
In the exploration phase, we used a search strategy with an organized structure of key terms used to perform comprehensive literature searches in databases. These keywords played a key role in identifying relevant studies and helped us focus on specific publications addressing the research questions. The literature search used multiple academic databases, including Web of Science, IEEE Xplore, Science Direct, arXiv, and PubMed, and the following groups of keywords in the article titles, keywords or abstracts: (“Machine Learning” OR “Deep Learning” OR “Classification” OR “Supervised Learning” OR “Neural Networks”) AND (“Electroencephalography” OR “EEG”).

### 2.3. Criteria for Identification of Studies

After searching the databases, we established a protocol based on specific inclusion and exclusion criteria to identify publications qualified for inclusion in our review. The eligible publications met the following inclusion criteria: (a) publication in English; (b) inclusion of EEG data; (c) application to the classification of brain activities in humans; (d) publication in a peer-reviewed journal; and (e) publication within the past 21 years (2000–2021). The exclusion criteria used during the screening process included: (a) publication in a non-peer-reviewed journal; (b) content types of dissertations, published abstracts, or book chapters; and (c) experimental studies performed in infants. The relevant studies were selected on the basis of the inclusion and exclusion criteria and several data categories were collected including task information (i.e., task type, number of subjects, and number of EEG channels used), databases used, frequency range used for analysis, feature extraction methods, ML/DL algorithms, and their performance (see Table A1 in Appendix A, for details).

## 3. Theoretical Background

Figure 1 illustrates the general pipeline for the construction of a classification model based on EEG data. Processes including data acquisition, preprocessing, feature extraction, and ML and DL models can be differentiated from one another. In the following, we review the existing computational methods for each step.



**Figure 1.** Flow chart of the EEG data analysis pipeline along with taxonomy of existing methods for each step.

### 3.1. EEG Data Acquisition

Typically, EEG recordings were obtained by using an international 10–20 or 10–5 electrode placement system introduced by the American EEG Society. In this method, multiple noninvasive electrodes are placed on the surface of the scalp. Each electrode receives brain signals, which are then amplified and sent to a computer after being converted them into line images representing brain waves. Electrodes used in EEG recording can be categorized as wet electrodes, semi-dry electrodes, and dry electrodes [40]. In wet electrode recording, conductive gel or paste, which is very uncomfortable for participants, must be placed between the electrodes and the skin. Semi-dry electrodes require a small amount of conductive gel, and dry electrodes do not require conductive gel or skin preparation [41,42]. Despite its ease of use and quick setup, EEG recording from dry electrodes yields data with more artifacts than does gel-based EEG recording, thus affecting EEG analysis. In this case, identifying the most sensitive biomarkers for the specific task is essential [43]. In addition, some experimental studies have confirmed that selecting the proper features can improve the performance of ML/DL methods. For instance, Dehais et al. [44] have analyzed the performance of gel-free EEG recording with ERP and power spectral density (PSD) features along with the linear discriminant analysis (LDA) method and found that the performance of LDA with PSD features is much better than that of LDA with ERP. Thus, ERP appears to be more sensitive to noise than PSD features.

### 3.2. Artifacts in EEG Signals and Preprocessing

The preprocessing step facilitates the removal of low-quality data without altering the clean data. This process also fragments the continuous raw signals without changing the data [45]. Artifact removal is an essential preprocessing step in the analysis of EEG signals, because the recordings usually include a significant number of extrinsic artifacts associated with environmental noise and experimental error, as well as intrinsic biological artifacts associated with body function (e.g., eye blinking, movement, respiration, and heartbeat) [46]. Various simple methods can be used to eliminate non-biological artifacts from EEG signals. Because environmental artifacts do not have the same frequency as the EEG signals of interest, they can be eliminated through application of a band filter [46]. Alternatively, standard operating procedures provide proper operational guidance for the data acquisition step and decreasing experimental artifacts [47]. However, the main types of biological noise include ocular artifacts, muscle artifacts, cardiac artifacts, and instrument artifacts, which require the use of filtering and/or computational methods to be removed from EEG data. More in-depth descriptions of these types of artifacts can be found in references [46,48]. According to the literature, extensive research has been conducted in the past decades [46,49] to identify and define efficient methods for both automatic and manual artifact removal.

#### 3.2.1. Regression Methods

A typical approach to artifact removal is the regression method [50], which removes estimated artifacts by determining the amplitude relation of the reference. Hence, signals such as those from an electrocardiogram (ECG) or electrooculogram (EOG) are required to separate artifacts from EEG signals. Although the regression method is based on simple mathematical intuition and therefore is widely used because of the minimal computation required, the dependence of this method on reference channels for ECG and EOG removal is considered a drawback [48].

#### 3.2.2. Blind Source Separation Methods

Blind source separation (BSS) methods are based on the hypothesis that a combination of several distinct original signals results in the signals observed on a multi-channel recording; thus, neither more reference channels nor prior information is required [51]. Three typical methods using the BSS algorithm are principal component analysis (PCA), independent component analysis (ICA), and canonical correlation analysis (CCA). ICA [52,53] is a statistical algorithm that is used for solving the BSS problem and considers random



variables to maximize the independence of the output components via the discovery of a linear transformation. Furthermore, ICA is a powerful tool that reduces dimensions and extracts independent components from original signals. This method has good performance in extracting artifacts such as eye blinks and heartbeats, because they are generated by independent sources and are not associated with particular frequencies. ICA is efficient according to the assumption that original signals are statistically independent of each other and have a non-Gaussian distribution. Moreover, the signal dimension must be greater than the source signal. PCA [54,55] maximizes the variance of the transformed data and depends only on the second-order statistics of covariance [51]. PCA is a well-known method to reduce the dimensionality of the features while protecting their statistical information. The advantage of this method is that it retains the variance of the data set. However, its disadvantage is that if the potentials of drifts and EEG data are similar, PCA cannot extract the appropriate interferences. CCA is widely used in SSVEP-based brain BCI to identify the frequency components of EEG that characterize visual stimulus frequencies [56]. Of note, a comparison of BSS methods is necessary, but that topic is beyond the scope of this review article. Reference [57] provides further information about these methods.

### 3.2.3. Wavelet Transform

Wavelet transform (WT) [58–61] is a spectral estimation method that converts a time-domain signal into a time and frequency domain signal. After decomposition of wavelet transformation on EEG signals and during artifact removal, WT localizes the features and maintains them during the filtering process by defining a threshold for the elimination of noise signals. Whereas WT has good performance in analyzing the components of non-stationary signals, it fails to recognize artifacts that overlap with the spectral properties. Accordingly, new hybrid methods such as wavelet-BSS have been proposed to overcome this drawback [62].

### 3.2.4. Filtering Methods

Various filtering approaches that have been used for EEG artifacts and noise-canceling include frequency filtering, adaptive filtering, and Wiener filtering [63].

#### Frequency Filtering

Frequency filtering is a simple classical separation technique to eliminate artifacts from the desired EEG signals. According to Hu and Zhang [45], four types of frequency filters can be considered: low-pass, high-pass, band-pass, and band-stop filters [45]. However, this method is not efficient if the spectral distributions of artifacts and the EEG components overlap. In the case of overlap, alternative artifact removal techniques are necessary [63].

#### Adaptive Filtering

Adaptive filtering is based on the assumption that the EEG signal of interest and the artifact are uncorrelated. This filter uses a reference signal and produces an estimated signal that is correlated with the artifact; the estimate is then subtracted from the primary signals to yield a noise-free EEG signal [64]. Adaptive filtering uses the least mean squares (LMS) algorithm, which is linear in convergence, to assess the clean signals by upgrading the weight parameter. Another optimization algorithm, the recursive least squares algorithm, is quadratic in convergence and is an extension of LMS [48]. Depending on the type of recursive least squares algorithm, its convergence may be faster than that of the LMS algorithm, but its computational cost is greater. A disadvantage of using adaptive filtering is that providing reference input requires more sensors [64].

#### Wiener Filtering

Wiener filtering is a statistical technique to minimize the mean square error between the signals of interest and the estimated signals by generating a linear time invariant filter [63]. Although Wiener filtering does not require an additional reference, because the

minimization process is applied to estimate the power spectral densities of the EEG signal and artifact signal, the computational process can be complicated. In addition to the listed techniques, many other efficient methods exist, such as CCA, empirical mode decomposition (EMD), and sparse decomposition methods. Furthermore, hybrid methods combining these preprocessing algorithms and other methods such as EMD-BSS, wavelet-BSS, and others, have been used to maximize the efficiency of the algorithm [65,66]. Further details can be found in previous publications [46,48].

### 3.3. Feature Extraction Methods

EEG signals are typically complicated and contain a large amount of information. Thus, the ability to extract the proper features from EEG signals is a critical component of any successful ML and DL algorithm. The feature extraction step aims to transform the data into a low-dimensional space while maintaining the critical information conveyed by the EEG signals [67]. According to the literature, many feature extraction methods have been proposed on the basis of the specific tasks, including time domain, frequency domain, and time-frequency domain as well as spatial information in the signals [68,69]. Among these extraction methods, ICA, PCA, and autoregressive (AR) models are considered time-domain methods. Statistical measures to evaluate parameters such as the mean, standard deviation, variance, root mean square, skewness, kurtosis, relative band energy, and entropy also fall into this category [70]. Fast Fourier transform (FFT) and Welch's method [71] are among the frequency-domain methods used to analyze EEG signals. WT and short time Fourier transform (STFT) are two standard time-frequency domain methods that extract features based on both time and frequency. Findings from recent studies have demonstrated the advantages and drawbacks of each feature extraction method, and have indicated that care must be taken in selecting the appropriate method for a specific type of task [67]. Here, we present a general overview of the most commonly used methods for analyzing EEG signals. More details on the feature extraction methods used according to the specific nature of the task are provided in the sections below.

#### 3.3.1. Principal Component Analysis

PCA is a linear transformation that is widely used to reduce dimensionality. PCA introduces a vector in a lower-dimensional space to reduce signal complexity over time and space [72]. Although PCA is used to separate artifacts from original signals, this transformation can also be used for feature extraction without information loss [70]. PCA creates a set of linear vectors that are not correlated with one another (i.e., principal components) via converting correlated variables from original signals [73]. Although principal components enhance signal similarity and improve the effectiveness of classification of data [74], they are not as interpretable as primary features. In addition, PCA does not work well in analysis of complex data sets [75]. To address these drawbacks, several variations in PCA have been proposed in EEG data analysis such as kernel PCA [76] or sparse PCA [77].

#### 3.3.2. Autoregressive Model

The AR model is a feature extraction method for frequency domain analysis, which has been used to analyze non-stationary signals such as EEG data [73]. AR assumes that real EEG signals can be predicted by the AR process; this prediction can be performed with the order and parameters of the approximation model. The AR model's order is a value from 1 to 12, which effectively indicates the performance of the model. However, selecting an appropriate value for the order of the AR can be challenging, because improper order selection causes miscomputed spectrum estimation, which may increase the computational costs [78].

AR methods such as bilinear AAR, adaptive AR parameters, and multivariate AAR have been widely used in EEG data analysis, thus allowing the AR model parameters to adapt to nonstationary EEG signals [79]. These methods help provide successive parameter estimation and minimize prediction error. For example, AR parameters can be

evaluated in an adaptive AR model by using the Kalman filter, thus improving classification performance up to 83% [74]. Other advantages of the AR model include that it provides appropriate frequency resolution [75] and can be used to estimate the power spectra of shorter segments of EEG data in various applications [80]. However, AR is vulnerable to inappropriate order and parameter selection [75].

### 3.3.3. Fast Fourier Transform

FFT is an effective method for stationary signals. It transforms signals from the time domain to the frequency domain and implements spectral analysis [67]. In this method, features are extracted by using mathematical tools to calculate the PSD. The estimation of PSD for a related band can be computed with FFT, which uses non-parametric methods such as Welch's method [67,81]. Although FFT is commonly used in the data analysis process, and it works effectively for stationary signals, it is not efficient for nonlinearity and nonstationary data such as EEG signals; moreover, the results obtained through this method are not reliable. This shortcoming has motivated researchers to develop novel procedures and methods for the analysis of nonstationary signals, such as the Fourier decomposition method [82], variational mode decomposition (VMD) method [83], and Hilbert-Huang transform (HHT) method [84].

### 3.3.4. Wavelet Transform

WT is a time-frequency transform that considers the features of the EEG signals within a frequency domain and is perfectly localized within the time domain [70]. This method has good performance in spectral analysis of irregular and nonstationary signals within different size windows [85]. One advantage of WT is that it provides accurate frequency information and time information at low and high frequencies, respectively. That is, a narrow window is typically used to evaluate high frequencies, and a wide window is applied to assess low frequencies [68,86]. Thus, WT is suited for transient oscillation in signals, particularly biosignal data, which consist of low-frequency components with long-time periods and high-frequency components with short-time periods [74]. However, WT suffers from Heisenberg uncertainty, which negatively affects its performance [74]. WT evaluates small wavelets within a specific range for a limited duration. The wavelets have oscillating motion starting at zero; these oscillations increase and then decrease to zero [87].

### 3.3.5. Common Spatial Pattern

CSP is a successful feature extraction method in BCI applications, particularly motor imagery tasks [88], that can be used for spatial filtering by using the whole data trail or by splitting trails into time segments. CSP is widely used for binary classification tasks [89]. The objective of CSP is to differentiate between classes by minimizing the variance of one class and maximizing the variance of the other class. This process can be implemented by introducing spatial filters for each class. With this method, EEG signals are transformed into a variance matrix representing the discrimination between classes [73].

The main advantage of using CSP is that it is simple and can be executed rapidly [73]. However, this method has some inherent limitations in extracting optimal features from raw EEG data [88]. To address this issue, several studies have proposed optimal spatial feature selection methods. For example, Jin et al. [88] have developed a novel feature selection method based on an improved objective function by using Dempster-Shafer theory, considering feature distribution. Furthermore, CSP is highly sensitive to artifacts present in the original data set, and changing the positions of electrodes affects the classification accuracy [90]. According to reference [91], several parameters such as the frequency band filter, the time segment, and the subset of CSP filters to be used should be considered to have an effective CSP algorithm. Thus, the performance of the CSP algorithm depends on the subject-specific frequency band. Many approaches have been proposed to address the issue of identifying the optimal frequency band for the CSP algorithms, such as the common spatio-spectral pattern (CSSP), common sparse spectral-spatial pattern (CSSSP), spectrally



weighted common spatial pattern, sub-band common spatial pattern, and discriminant filter bank common spatial pattern [92,93]. CSSP uses a simple time delay embedding with the CSP algorithm, which improves the CSP's performance by optimizing the frequency band at each electrode position [94]. However, the non-stationary EEG data decrease the performance of the CSSP. To overcome this challenge, Cho et al. [95] have included a noise removal term in the Rayleigh coefficient of CSSP and have designed an invariant CSSP algorithm that is both consistent and robust to noise. However, the drawback of the invariant CSSP is that the optimal noise removal value must be determined. The CSSSP [96] has been proposed to enhance the performance of CSSP.

In contrast to the CSSP, which identifies various spectral patterns for each channel, this algorithm searches for a common spectral pattern for all channels [93]. Sub-band common spatial pattern [97] is an other extension of CSP used to filter EEG signals at multiple sub-bands to extract CSP features from each sub-bands, regardless of the associations among features from different sub-bands [93]. To overcome this limitation, a discriminant filter bank common spatial pattern [98] that uses the Fisher ratio of single channel band power values has also been proposed. Further information regarding the application range of the CSP can be obtained in reference [93].

### 3.4. Classification Algorithms

Artificial intelligence includes ML, and DL is a rapidly growing area with applications in the classification process [99]. The objective of classification is to predict the class label of the new data points in various tasks [27]. Classification algorithms can be divided into two categories: conventional classification algorithms and DL algorithms [100]. Conventional classification algorithms represent a precise effort to build classification models by using input data, and applying statistical analysis to classify output values. Most conventional classification algorithms use hand-crafted input features to train the model. This process, called feature creation, has limitations in handling input in high dimensional data sets [100]. DL algorithms rely on representation learning [101] and can accommodate the limitations of conventional classification algorithms through learning features automatically at multiple levels of abstraction [100]. Table 1 presents a brief comparison of conventional classification and DL algorithms [100].

**Table 1.** Brief comparison of conventional classification algorithms and deep learning algorithms.

	Conventional Classification Algorithms	Deep Learning Algorithms
Input features	Hand-crafted	Automatically based on representation learning
Feature selection process	Required	Not required
Model architecture	Based on statistical concepts	Consists of a diverse set of architecture based on sample data
Computational cost	Computational cost is based on the conventional classification models but is lower than that of deep learning algorithms	Computational cost is very high, because hyper parameters must be tuned

#### 3.4.1. Conventional Classification Algorithms

Among different conventional classification algorithms including supervised learning and unsupervised learning, supervised algorithms are the most well-known methods used in EEG data analysis [102]. One of the most commonly used supervised algorithms is artificial neural networks (ANNs). ANNs are computational models [103] that use multi-layered networks of neurons with weighted connections between units, typically followed by a static non-linearity function (e.g., ReLu). During the learning phase, the network can learn by modifying its weights to enhance the performance outcomes in test data classification [104]. Similar examples of well-performed and well-known supervised algorithms include naive Bayes (NB), support vector machine (SVM), k-nearest neighbor (KNN), logistic regression (LR), random forest (RF), and LDA. Each supervised model applies a learning algorithm to generate a more accurate model [105].

NB is a probabilistic classifier that applies Bayes' theorem to classify data on the basis of certain features [106]. It is a simple and effective classifier that needs only small training data sets to estimate the parameters for classification. This advantage makes NB a robust classifier for EEG signals analysis in several types of tasks such as ER [107], seizure detection (SD) [108], and MI [109]. However, NB is based on the assumption that all attributes are independent of one another, and feature vectors have equal effects on the outcome [106].

SVM has been demonstrated to be a useful supervised model based on a statistical learning tool with high generalization. The principle underlying SVM is the separation of two data sets. This separation can be linear or non-linear. In the case of linear separation, SVM uses a discriminant hyperplane to distinguish classes. However, in the case of non-linear separation, SVM uses the kernel function to identify decision boundaries. Compared with that of other supervised algorithms, such as ANNs and KNN, the computational complexity of SVM is low [110,111]. Although the computational complexity of KNN decreases by increasing the k-value, its classification performance also decreases [110,112]. Furthermore, with the advent of DL algorithms, SVM has remained widely used in EEG signal classification, because its computation has a solid mathematical basis. However, the performance of SVM is affected by the kernel function and penalty coefficient parameters; thus, optimizing the parameters introduced into SVM classifiers is essential [113]. Huang et al. [114] have applied a genetic algorithm, and Wang et al. [115] have proposed particle swarm optimization to optimize SVM parameters. According to our investigation, SVM has been widely used in EEG signal classification because of its simplicity and adaptability in solving classification problems such as diagnosis of brain disorders (e.g., SD, and Alzheimer's disease) [116–118].

RF is a tree-based supervised algorithm that constructs an ensemble of decision trees. Each decision tree is generated during the training phase. RF makes predictions from each tree and selects the final decision via a voting method or averaging the results [119] to identify the most commonly used class. The main idea underlying this and related ensemble methods is that a group of weak classifiers can collectively generate a strong classifier to create a successful learning algorithm. However, the overfitting and instability of trees can affect RF model performance, particularly with varying sizes of trees [106]. In contrast to the LR model, which is a probabilistic classification model for both binary and multi-class classification tasks [120], RF works on both discrete and continuous data, thus providing models for classification and regression problems. Furthermore, the parallelization structure of RF results in better performance than that of the other supervised algorithms on large EEG data sets in addressing classification problems [109].

LDA is a linear transformation technique used to identify linear combinations of the variables that most effectively separate the classes [121,122]. LDA is based on the assumption that the density for the data is normally distributed, with equal covariance for both classes. The separating hyperplane is achieved by maximizing the distance between the two classes, while minimizing the distance points within each class [123]. This technique is simple to use and has very low computational requirements. Consequently, LDA has been successfully applied to address classification problems in BCI systems such as MI based BCI [124], P300 speller [125], and multiclass BCI [126]. However, the main limitation of this model is its linear nature which prevents competitive results on nonlinear EEG data [127,128].

### 3.4.2. Deep Learning Algorithms

DL is a new branch of ML that has received widespread attention in EEG classification tasks. Although conventional classification algorithms have been very effective in analyzing massive data sets and understanding the relationship between variables, such algorithms often lead to poor generalization behavior and low classification performance when highly dynamic features are encountered [129]. DL algorithms inspired by neuroscience [130] exploit learning features from raw data without depending completely on preprocess-

ing [131]. Humans can transfer knowledge and memory throughout their lifespan, whereas DL algorithms immediately forget the previous learning after being trained on a new data set [132]. For instance, Borgomaneri et al. [133] have confirmed that fear memory remains in humans after memory reactivation and affects the future learning process. Such a life-long learning task is a large challenge in developing neural network algorithms [134]. Furthermore, DL algorithms apply multiple layers of perceptrons that obtain representation learning [131]. Recent developments in graphics processing unit technology have enabled the development of DL architectures on large data sets. This advantage has significantly improved the performance on large data sets with high-dimensional data [20].

A convolutional neural network (CNN) is a type of deep neural network that has gained attention, particularly in computer vision and neuroimaging [131]. CNN can identify the image of an object by using convolutions within its architecture; including convolutional layers that have parameters to create a feature map; pooling layers that reduce the number of features for computational efficiency; dropout layers that help avoid overfitting by randomly turning off perceptrons; and an output layer that map the learned features into the final decision, such as classification [135,136]. The recent emergence of the CNN algorithm has enabled outstanding performance in several application such as image processing, natural language processing, and classification of EEG recordings, particularly for MI tasks [137–140]. However, CNN performance is highly dependent on hyperparameters such as the number of convolution layers, and the size and number of kernels and pooling windows [137]. Fortunately, CNN architectures have led to automatic optimization of parameters through several iterations [20]; therefore, CNN is very commonly used for addressing classification problems involving large data sets.

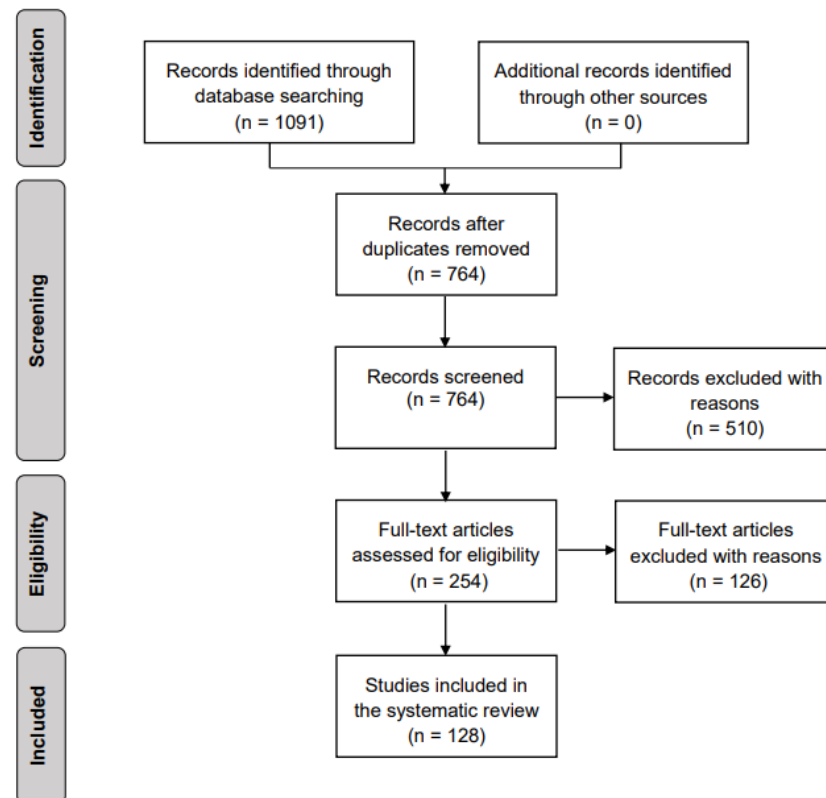
A recurrent neural network (RNN) is a time series-based DL algorithm that uses sequential data and learns from training data, similar to feed-forward and CNN methods. Unlike traditional deep neural networks, which assume independence of the input and output, RNN takes information from inputs continually. Consequently the output of RNN depends on the prior outputs and the current inputs within the sequence [137]. Thus, the architecture of this network includes inbuilt memory cells for storing information from previous output states [141]. The form of RNN architecture has enabled these types of networks to effectively analyze time series data for applications such as speech recognition [142], natural language processing [143], and disease signal identification [144]. The most commonly used RNN variants are long short-term memory (LSTM) [145], LSTM peephole connections [146], gated recurrent units [147], and multiplicative LSTM [148]. The ability of these variants to preserve and retrieve memories is part of the generic structure of these networks. However, CNNs have a different architecture and use filters and pooling layers. According to differences in the internal network structure of CNNs and RNNs, CNNs are effective for analysis of spatial data, because CNN models consider the complete trial as an object and can extract features. In contrast, RNN models are suited for analysis of temporal and sequential data by slicing the trail into several subtrails [137].

## 4. Results

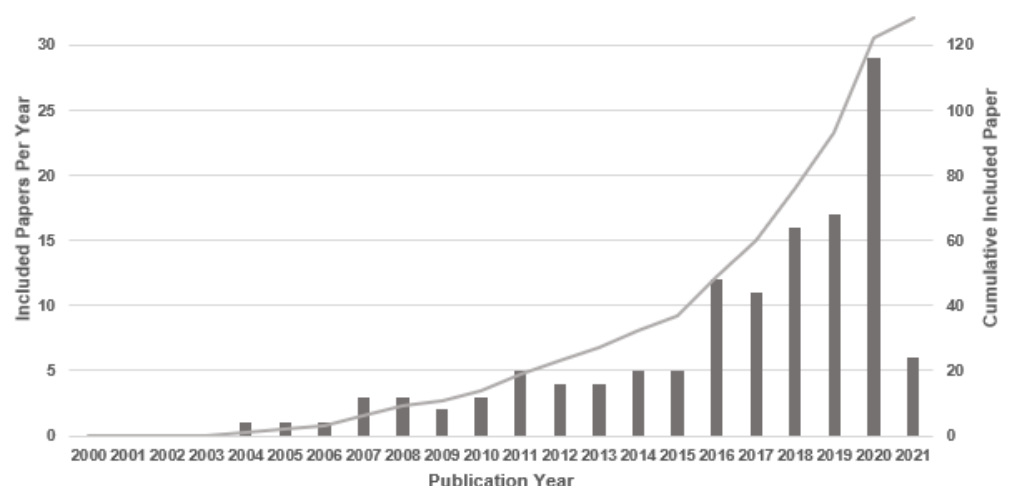
### 4.1. Literature Search

Figure 2 is a flow diagram of the study selection process based on the PRISMA guidelines [38]. As shown in the diagram, we obtained 764 articles from all database searches after the removal of duplicates. To determine which articles were appropriate to consider in this study, we reviewed all abstracts to determine whether the findings met our inclusion criteria. A total of 254 articles remained after this evaluation. The full text of each of these 254 articles was reviewed, and 128 relevant articles met all the aforementioned criteria. The selected studies included 78 articles published in journals, 46 conference and symposium articles, and 4 preprints. To improve understanding of the evolution of research in this domain, we also considered the temporal distribution of these publications (Figure 3). Since 2016, the importance of supervised ML and DL classifications and their role in analyzing EEG data has received increasing attention from the research community;

more than 71% (91 of 128) of the selected studies were published during this period (i.e., 2016–2021). The remaining studies (29%) were published between 2000 and 2015. We anticipate that the number of publications in this domain will continue to grow substantially over the next few years.



**Figure 2.** Flow diagram based on the PRISMA guidelines [38]. The diagram includes the four stages of a PRISMA study: identification, screening, eligibility, and inclusion.

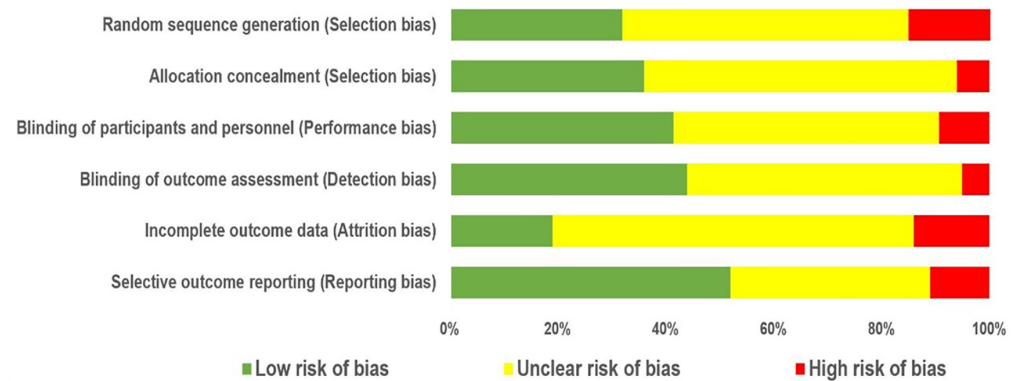


**Figure 3.** Temporal distribution of articles selected for consideration.

#### 4.2. Quality Assessment

To assess the quality of the selected articles, we applied the Cochrane Collaboration tool [39]. According to this tool, the articles were identified as having an (a) low risk of bias, (b) high risk of bias, or (c) unclear risk of bias. The overall quality of the articles was categorized as weak (fewer than three low-risk domains), fair (three low-risk domains),

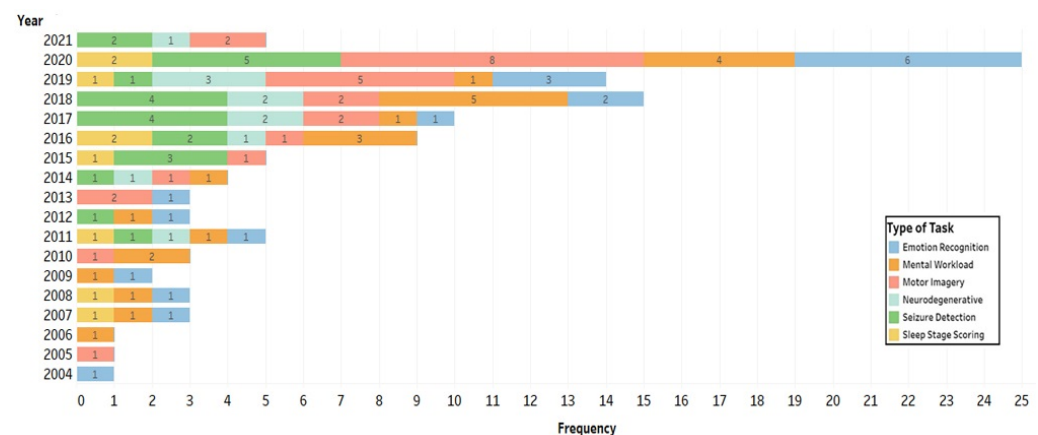
or good (more than three low-risk domains). Of the 128 studies considered in this systematic review, 40 articles were categorized as good, 26 articles were categorized as fair, and 62 articles were categorized as weak (Figure 4).



**Figure 4.** Assessment of the risk of bias in the 128 studies selected for this review by using the Cochrane Collaboration tool.

#### 4.3. Study Characteristics

The tasks presented were in studies that were organized into six groups: ER (15%), mental workload (MWL) (18%), MI (20%), SD (19%), sleep stage scoring (SS) (7%), and diagnosis of neurodegenerative diseases (ND), including Alzheimer's disease (AD), Parkinson's disease (PD), and schizophrenia (SZ) (9%). Other studies (12%) focused on ERP [47,149–152], anxiety and stress [153,154], depression [33,66,155], the detection of alcoholism [156], auditory diseases [157], attention deficit hyperactivity disorder [158], sleep apnea [159], and the classification of creativity [160]. In recent years, the application of supervised ML and DL models in MI and ER has gained significant attention, despite decades-long research studies already in progress in these fields (Figure 5).



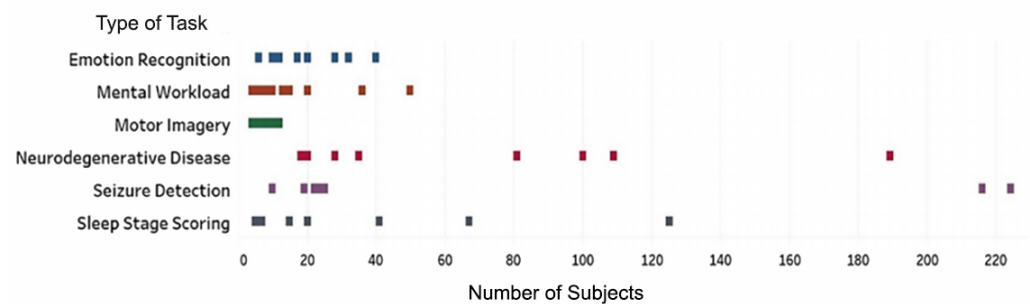
**Figure 5.** Temporal distribution of the number of publications per domain in each year (2004–2021).

In the selected articles, two types of data sets have been used: (1) experimental data sets associated with a specific project or research direction and (2) freely available data sets comprising EEG data associated with various tasks. In the first type of data set, EEG recordings are obtained from participants by using non-invasive electrodes placed on the scalp surface. Of the 128 studies included in this systematic review, 66 studies (52%) generated experimental data sets; the remaining studies used data in publicly maintained databases. Table 2 includes a list of the public data sets frequently used for various tasks in the selected studies. The descriptions include the numbers of participants and the target tasks. Corresponding to the type of data set, the number of participants varied among studies depending on the type of task involved (Figure 6).



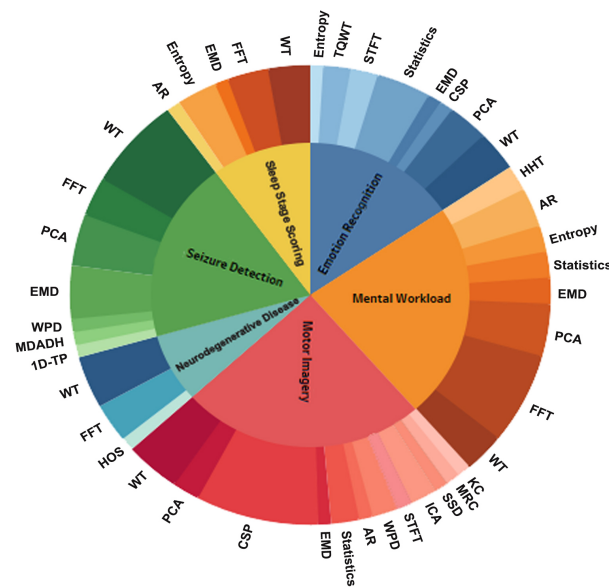
**Table 2.** List of public data sets used for the analysis of EEGs associated with different types of tasks.

Source	Database	Studies Using This Data Set	Number of Subjects	Target Tasks
Koelstra et al. [161]	DEAP	[31,119,162–168]	32	Emotion recognition
Blankertz et al., Leeb et al. [169,170]	BCI Competition	[49,61,104,171–177]	30 subjects in 4 different data sets	Motor imagery
Andrzejak et al. [178]	BONN	[118,179–187]	5	Seizure detection
Moody et al. [188]	CHB-MIT	[189–192]	22	Seizure detection
Goldberger et al. [193]	PhysioNet	[116,194–196]	109	Motor imagery/mental workload
Ihle et al. [197]	European Epilepsy	[198–200]	300	Seizure detection
Kemp et al. [201]	Sleep-EDF	[59,202–204]	197	Sleep stage scoring
Ichimaru et al. [205]	MIT-BIH	[159]	16	Sleep apnea detection
Winterhalder et al. [206]	Freiburg	[207]	21	Seizure detection

**Figure 6.** Number of subjects per task in each study reviewed.

#### 4.4. Which Feature Extraction Methods Were Used for Each Task to Extract Appropriate Inputs for Machine Learning and Deep Learning Classifiers?

When using a feature extraction method, the principal objective is to minimize the loss of essential characteristics embedded in the EEG signals. In addition, feature extraction methods provide conditions that promote the optimal selection of features that are important to the specific classification task. Among the articles reviewed for this study, various methods have been used to extract the features from EEG signals associated with different type of tasks. Figure 7 shows a plot of the feature extraction methods used for each of the six categories: ER, MWL, MI, SD, SS, and ND tasks. The inner circle represents the type of task, and the outer circle represents the utilization rate of the method for each task. The results suggested that EEG signals were commonly analyzed in the time-frequency domain in the reviewed articles. WT was the most common feature extraction method used for all tasks. Our investigation revealed that researchers examining MWL, ND, and SS tasks typically chose FFT; in contrast, PCA was heavily used for feature extraction in studies focused on ER, MWL, and SD tasks. Furthermore, the CSP model was a commonly used feature extraction method in EEG-based BCI systems for MI tasks. In some research studies, the combination of CSP and other methods was used to decode the EEG signals. For example, Wang et al. [208] have examined the performance of CSP with and without discriminative canonical pattern matching method and have found that a combination of CSP and discriminative canonical pattern matching, as compared with the single CSP method, significantly improved the accuracy of feature extraction in pre-movement EEG patterns by 10%. Several other efficient feature extraction methods have been used in studies involving various tasks, such as Higher-Order Spectra (HOS) for the PD detection task [209], one-dimensional ternary patterns (1D-TP) for the SD task [187], and Kolmogorov complexity (Kc) and sparse spectrotemporal decomposition (SSD) for the MI task [210].



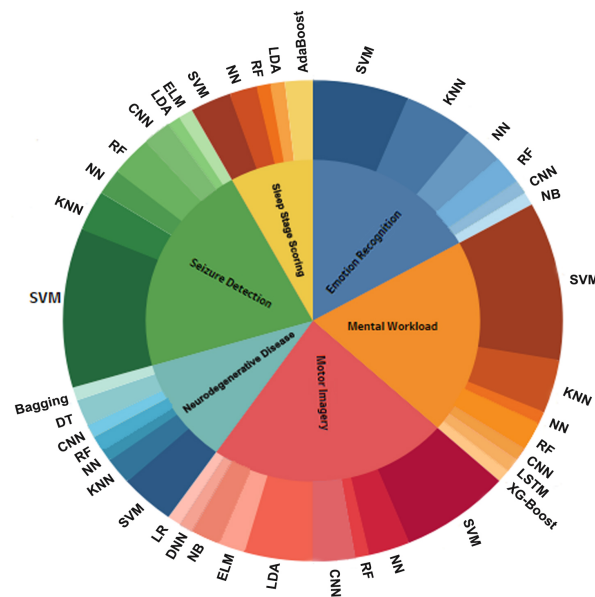
**Figure 7.** Feature extraction methods used in all studies considered in this review. The inner circle represents the type of task, and the outer circle represents the utilization rate of the method for each task. Abbreviations: AR, autoregressive; CSP, common spatial pattern; EMD, empirical mode decomposition; FFT, fast Fourier transform; HHT, Hilbert Huang transform; HOS, higher-order spectra; ICA, independent component analysis; Kc, Kolmogorov complex; MDADH, maximum difference of amplitude distribution histogram; MRC, multiple Riemannian covariance; PCA, principal component analysis; STFT, short-time Fourier transform; SSD, sparse spectrotemporal decomposition; TQWT, tunable Q wavelet transform; WPD, wavelet packet decomposition; WT, wavelet transform; 1D-TP, one-dimensional ternary patterns.

#### 4.5. What Are the Most Frequently Used Machine Learning and Deep Learning Algorithms for EEG Data Processing?

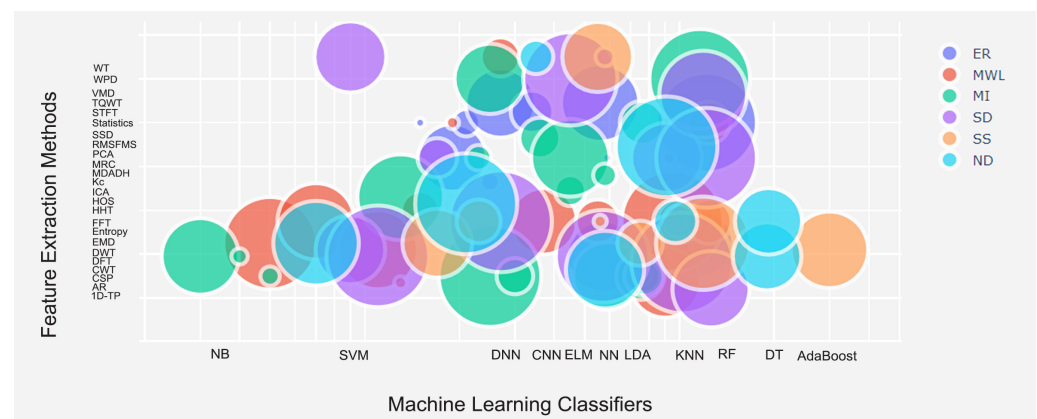
Among the studies performed over the past 21 years that have focused on classifying EEG signals, SVM was the preferred supervised algorithm for nearly all types of tasks; more than 40% of studies achieved the highest accuracy by using this ML classifier. Figure 8 indicates the plot of ML and DL classification algorithms used for each of the six types of tasks: ER, MWL, MI, SD, SS, and ND. The inner circle represents the tasks, and the outer circle represents the utilization rate of the classification algorithms for each task. As confirmed in Figure 8, CNN, KNN, and RF were among the most accurate classifiers after the SVM algorithm.

Assessing the performance of classification models depends on the type of data set (i.e., size and quality) as well as selected feature extraction method. To obtain more intuitive information about our findings, we plotted a bubble chart on a 2-dimensional axis showing the relationships between ML and DL algorithms with feature extraction methods (Figure 9). The size of each bubble shows the performance of classification models for each task, as marked with different colors. As Figure 9 demonstrates, the use of SVM resulted in a wide range of performance in tasks including ER, MWL, and SD, when frequency-domain and time-frequency domain feature extraction methods, such as FFT and WT were chosen. RF had competitive performance in tasks that included ER, MWL, MI, SD, and SS, particularly in references [59,104,119,185,211]. However, in the domain associated with ND tasks, the RF algorithm did not perform as well as in other tasks. According to two recent studies [209,212], the best performance in the ND task classification was achieved when SVM and KNN were applied. More interestingly, none of the MI studies reported high classification accuracy with the KNN classifier; instead, LDA resulted in high performance in only the MI and SD tasks. Furthermore, classification algorithms known as AdaBoost achieved the highest classification accuracy rate for the SS task when the EMD feature

extraction method was chosen. DL algorithms, particularly CNN, showed promising performance in the ER, MWL, and MI tasks.



**Figure 8.** The most efficient machine learning algorithms used for different tasks. The inner circle represents the type of tasks, and the outer circle represents the utilization rate of the supervised machine learning and deep learning classification algorithms used for each task across all studies evaluated. Abbreviations: CNN, convolutional neural network; DNN, deep neural network; DT, decision tree; ELM, extreme learning machine; KNN, K-nearest neighbor; LDA, linear discriminant analysis; LR, logistic regression; LSTM, long short-term memory; NB, naïve Bayes; NN, neural networks; RF, random forest; SVM, support vector machine.



**Figure 9.** Bubble plot of studies according to classification algorithms and feature extraction methods. The size of the bubble indicates the performance of classification models for each task as marked with different colors.

## 5. Discussion

An important EEG signal is the P300 ERP, which has been used to build the P300 speller, a communication tool through which users can type messages by controlling their eye-gazing [213]. Although P300 ERPs are collected through signal acquisition methods, these signals have very low signal-to-noise ratios; thus, the stimuli process must be continued to ensure the improved signal-to-noise ratios. This repetition has drawbacks, such as decreasing typing speed and increasing typing error [213]. In the reviewed literature, several supervised ML and DL algorithms have been applied to classify P300 responses cor-

rectly with a smaller number of repetitions. Among supervised ML algorithms, SVM and LDA have been applied successfully in P300 classification [125,214]. Furthermore, CNN with various architectures has been widely used in the P300 classification task [213,215–217]. However, the details on P300 classification using supervised ML and DL algorithms are beyond the scope of this review. In this section, we provide more comprehensive explanations of the contents of the articles reviewed. These explanations include further discussion of the applications of ML and DL algorithms and the feature extraction methods provided to ML classifiers depending on task types. This section also provides recommendations for selecting a ML and DL classifier according to the selected models in the reviewed articles, according to their observed performance.

### 5.1. Emotion Recognition Task

Emotions play essential roles in the evaluation of human behavior. EEG signals provide a convenient way to analyze an individual's emotional response to external stimuli. Research has focused on classifying and predicting emotion dimensions while participants perform in externally driven activities, including watching video clips, facial pictures, and sequences of images [65,218,219]. In addition to exposure to these external stimuli, expression of the basic innate emotions (i.e., fear, anger, surprise, sadness, happiness, and disgust) causes different behavioral responses in individuals [220,221]. Notably, fearful facial expression is a threatening stimulus resulting in an appropriate organization of defensive responses [222]. Moreover, environmental factors, such as approaching direction [223] have contributed to clarifying the effects of behavioral responses (i.e., defensive responses).

ER research based on EEG signals has recently become a commonly investigated topic (Figure 5). Of the 19 selected articles on this topic, 12 studies (63%) were published in the past 3 years. This field is expected to continue to grow over the next several years. Most studies used the Database for Emotion Analysis Using Physiologic Signals (DEAP), the most commonly used data to evaluate ER tasks and classify emotional states in two dimensions [161]. Table 3 briefly compares the performance of ML and DL algorithms observed in ER studies. As shown in Table 3, the NB classifier did not perform well for this task. Chung and Yoon [31] have used a weighted-log-posterior function for the Bayes classifier and reported an accuracy for valence and arousal classification of 66.6% and 66.4% for two classes and 53.4% and 51.0% for three classes, respectively. Shukla and Chaurasiya [167] have classified emotional states by using emotional dimension in the valence-arousal plane and reported that KNN outperformed other classifiers with an average accuracy of 87.1%. Three studies have analyzed differences in the performance of PCA in extracting desired features with different ML and DL classifiers. An earlier publication [164] has evaluated classification results both with and without a dimensional reduction technique; the results confirmed that KNN with PCA achieves better classification accuracy than other algorithms including SVM, LDA, LR, and DT. Bazgir et al. [163] have performed a similar analysis and decomposed EEG signals into five specific frequency bands by using discrete wavelet transform (DWT) via the db4 mother wavelet function on the selected channels. PCA was then applied to extract spectral features, and SVM with an RBF kernel was used to extract features from ten channels. This classification method achieved 91.3% and 91.1% accuracy for arousal and valence, respectively. Similarly, the high accuracy of ANNs has reported been in a previous study [166] using kernel PCA to extract segment-level features. Two additional studies have reported a high accuracy rate when using the RF algorithm. In one of these studies [162], the authors examined the level of fear based on emotional dimensions. They assessed fear in two- and four-level modes and built classifier models both with and without feature selection. The classification results confirmed that, use of differential entropy features resulted in the highest accuracy (90.07%) of RF classifiers in evaluating fear. Ramzan and Dawn [119] have observed similar results and reported that RF consistently outperformed other algorithms with an accuracy of 98.2% when its statistical features were used. Moreover, Nawaz et al. [168] have evaluated the performance of SVM, KNN, and DT classifiers with different types of features and found

that statistical features performed best in assessing emotional dynamics in the human brain. Specifically, they achieved 77.62%, 78.96%, and 77.6% accuracy for binary classification of valence, arousal, and dominance, respectively, when these features were examined by using the SVM classifier. Qiao et al. [224] have proposed a novel model for multi-subject emotion classification. High-level features obtained by using the DL model and the CNN for feature abstraction resulted in a high accuracy of 87.27%. Overall, in studies on the DEAP data set, CNN can be chosen as a method to achieve competitive performance in ER tasks.

**Table 3.** Comparison of studies conducted on emotion recognition tasks by using the DEAP data set.

Authors	Year	Feature Extraction Method	Classification	Performance (%)
Ramzan and Dawn [119]	2019	Statistics	RF	Accuracy = 98.2
Bazgir et al. [163]	2018	PCA	RBF-SVM	Accuracy = 91.1 (valence) Accuracy = 91.3 (arousal)
Balan et al. [162]	2019	Entropy	RF	Accuracy = 90.07
Qiao et al. [224]	2017	STFT	CNN	Accuracy = 87.27
Shukla and Chaurasiya [167]	2018	DWT	KNN	Accuracy = 87.1 Accuracy = 77.62 (valence)
Nawaz et al. [168]	2020	Statistics	SVM	Accuracy = 78.96 (arousal) Accuracy = 77.6 (dominance)
Doma and Pirouz [164]	2020	PCA	KNN	Accuracy = 74.25
Chung and Yoon [31]	2012	N/A	NB	Accuracy = 66.6 (valence) Accuracy = 66.4 (arousal)
Rozgic et al. [166]	2013	PCA	ANNs	Accuracy > 60

In contrast, the studies that did not use this shared data set [65,218,219,225–230] all achieved higher classification performance by using SVM, KNN, and ANNs algorithms. For example, Seo et al. [227] have compared different ML classifiers to classify boredom and non-boredom in 28 participants on the basis of the historical models of emotion and found that the KNN outperformed both RF and ANNs. Likewise, nearly identical performance has been reported by Heraz et al. [218] in experiments analyzing EEG signals from 17 participants. Murugappan [65] has reported a high accuracy performance for KNN used together with DWT. SVM is also a commonly used ML classifier; Li and Lu [219] have achieved the highest accuracy rate of 93.5% by using this method. In that study, participants were provoked with pictures of facial expressions. The authors applied CSP and linear-SVM and found that the gamma band (approximately 30–100 Hz) was suitable for classifying EEG-based human emotion. Additionally, in prior studies [27,230], tuned Q wavelet transform (TQWT) has been implemented to elicit the relevant features from the sub-bands. In reference [27], the classification of six extracted features and a probabilistic neural network (PNN) resulted in accuracy of 96.16% and 93.88% for classifying emotions in participants diagnosed with PD and healthy control subjects, respectively. Similarly, reference [230] has applied a multiclass least-squares support vector machine and achieved 95.7% accuracy for the classification of four emotions (happiness, fear, sadness, and relaxation). Because the performance of these classifiers differed among these nine publications, further investigation will be needed to identify the most effective ML classification algorithm. However, ANNs, KNN, and SVM methods with different kernels appeared to outperform other algorithms in ER tasks.

### 5.2. Mental Workload Task

Industrial sectors, including transportation, military, and aviation, require individuals to perform multiple tasks simultaneously (i.e., multitasking); operators in these settings must maintain constant vigilance to perform various tasks efficiently and effectively. MWL involves human factors that determine what resources may be required to perform a specific task [211]. Moreover, prior studies have accepted the role of human mental activities (i.e., human behavior) in human cognitive abilities such as working memory [231–233]. In this



respect, Garofalo et al. [234] have indicated that individual differences are highly associated with the human behavior required by the task. Likewise, the influence of action control on execution and inhibition in motor responses in humans, and emotional stimuli in learned action have been discussed in references [235,236], respectively.

In the literature, three main approaches have been used to infer MWL levels: subjective measures, performance-based measures, and physiological measures [32]. Among these, approaches involving physiological methods that interpret MWL by using invasive, semi-invasive, and non-invasive physiological techniques perform relatively better than other measures. EEG data recording is a non-invasive method that provides a superior means to capture brain signals while participants perform complex tasks such as simulated driving or arithmetic calculations [32,237,238]. Overall, our review identified 23 publications that considered the MWL task. Different feature extraction techniques were widely applied in studies such as FFT, PCA, WT, AR, EMD, and HHT. Our investigation revealed that FFT, PCA, WT, and AR were the most widely used and effective methods for feature extraction among the reviewed articles in this domain; these methods were reported as superior 31%, 19%, 15%, and 15% of the time, respectively. For example, FFT has been implemented by researchers in several published studies [81,196,211,239–242] to extract features on the basis of the frequency of the EEG signals; however, another study has confirmed AR as one of the most reliable methods [243]. We also identified several essential feature extraction techniques that were less frequently applied in MWL tasks, including entropy and HHT, which elicited features of both nonlinear and non-stationary signals. Peng et al. [84] have applied HHT and SVM to evaluate attentiveness levels, and Vanitha and Krishnan [244] have used the HHT algorithm to extract EEG features to detect student stress levels. The study in reference [245] used the statistical method known as approximate entropy; more details on this method are provided in reference [246].

Various classifiers have been applied to identify the patterns in mental tasks to enhance classification performance. Our review indicated that SVM was the most commonly used ML algorithm to model MWL tasks. However, other ML methods, including KNN, ANNs, RF, and XGBoost classifiers have been used by others. For example, Dempster-Shafer theory and KNN (DSTKNN) classification methods have been applied in reference [243], in which desired features from EEG signals extracted by using both the AR model and statistical wavelet decomposition were provided to the DSTKNN classifier. The proposed algorithm achieved higher accuracy (93%) than did simple KNN. Likewise, Shah and Ghosh [55] have developed a real-time classification system by using PCA and a simple KNN classification algorithm. Interestingly, the studies in references [196,241] have proposed using FFT to evaluate the PSD on the basis of the time domain features incorporated in different ML models for classification. The results indicated the highest accuracy with KNN, at 99.42% and 90.5%, respectively. In addition to KNN, the application of ANNs [81], RF [211,240], and XGBoost classifiers [32] has been reported to identify individual intelligence quotients, mental work estimation, mental state, and task complexity. DL-based models have been applied in MWL tasks and have shown competitive performance in classification. For example, Jiao et al. [242] have designed a novel CNN architecture with an average result accuracy of 90% in 15 participants. In reference [247], the MWL has been classified by using the KNN classifier, the LSTM classifier, and the CNN + LSTM network. The best performance (61.08%) was achieved for the LSTM classifier. Given the proposed ML and DL algorithms, the FFT model has been implemented in several studies [81,211,240,242] to extract frequency domain features from different EEG bands.

Several studies have indicated that SVM classifiers are among the best techniques. Table 4 illustrates the highest classification accuracy of SVM when used with different types of feature extraction methods. In each study, the number of participants was six to eight, except for one previous study [84] that included 20 participants. After extraction, the combinations of all features were classified. As shown in Table 4, entropy, FFT, and EMD were associated with the highest accuracy when used with the SVM algorithm. The classification

accuracy of entropy used with Immune Feature Weight SVM (IFWSVM), FFT with cubic SVM, and EMD with SVM were 97.5%, 95%, and 94.3%, respectively.

**Table 4.** Classification accuracy of the selected feature extraction methods with the SVM classifier for a mental workload task.

Authors	Year	Feature Extraction Method	Classification	Performance (%)
Guo et al. [245]	2010	Entropy	IFWSVM	Accuracy = 97.5
Rashid et al. [239]	2018	FFT	Cubic SVM	Accuracy = 95
Gupta and Agrawal [248]	2012	EMD	SVM	Accuracy = 94.3
Vanitha and Krishnan [244]	2016	HHT	SVM	Accuracy = 89.07
Wei et al. [249]	2011	PCA	SVR	Accuracy = 85.92
Peng et al. [84]	2020	HHT	SVM	Accuracy = 84.8
Gupta et al. [250]	2020	WT/EMD	Non-linear SVM	Accuracy = 80–100
Hosni et al. [251]	2017	AR	RBF-SVM	Accuracy = 70
Liang et al. [252]	2006	AR	SVM	Accuracy = 67.57

Seven types of ML and DL classification algorithms were used to study MWL tasks: CNN, LSTM, SVM, KNN, ANNs, RF, and XGBoost. Two studies [196,241] have used KNN classifiers with the feature extraction method FFT and compared the performance of KNN versus other algorithms. Both studies have reported that KNN with FFT results in the most accurate performance. Moreover, the studies in references [245,251,253] have reported that IFWSVM and RBF-SVM are more efficient than simple SVM at this task. Guo et al. [245] have proposed the IFWSVM classifier. The immune algorithm was applied on the basis of the assumption that each feature contributes differently to the overall result. Optimal feature weights were extracted with the immune algorithm and used to train the IFWSVM classifier. A previous study [251] has compared the performance of different feature extraction algorithms and found that an AR model-based representation performs better than frequency-based representations. The remaining studies have used a variety of other feature extraction methods. On the basis of our finding, CNN, SVM (particularly IFWSVM and cubic SVM) and KNN classifiers have been selected as good candidate classifiers for tasks associated with the MWL.

### 5.3. Motor Imagery Task

BCI allows for communication between users and external devices by translating brain signals into computer commands [68,254]. Recently, BCI technology has made significant contributions to neuroscience by enabling analysis of brain signals with high temporal resolution [88,254]. The main EEG paradigms used by BCI systems include ERP [255], SSVEP [256], and MI. MI is the mental performance of a movement without any evident muscle activation. However, MI leads to the activation of specific brain areas in a manner similar to that observed in association with actual muscle movement. In MI tasks, EEG signals are recorded while participants are asked to imagine certain muscle movements of their limbs [257]. MI can be used as a component of a treatment plan to promote recovery in patients with limited motor abilities. Because of the complexity of the brain, various methods have been used for feature extraction to discriminate between features of MI tasks. The selected features were provided to classification algorithms that were asked to differentiate between MI tasks including left- and right-handed movements and movements of a foot or a limb. Of the 26 publications focused on MI tasks, the best performance was reported in studies using DL such as CNN, and ML such as SVM, LDA, LR, RF, and NB classifiers.

As reported earlier [68,176], CSP is the most common feature extraction method and has been used extensively in BCI tasks, in agreement with our findings. Of the 26 selected studies focusing on MI tasks, ten (38%) applied this method with various classifiers. For example, Islam Molla et al. [176] have used CSP to extract four sets of features from each sub-band; the selected features were then provided to the SVM classifier, thus resulting in nearly 93% accuracy. Similarly, CSP has been applied in the studies

featured in references [53,171] together with an SVM classifier. Soman et al. [171] have proposed a robust classifier using Twin-SVM and reported that this method enhanced the classification performance of left and right limb movements in the BCI Competition data set by 20%. However, CSP does not always perform perfectly with the SVM algorithm. Earlier studies [172,175,258] have compared the performance of ANNs and LDA with SVM. Jia et al. [175] have compared the results obtained by using backpropagation neural network (BPNN) and SVM algorithms and found that BPNN with CSP consistently outperformed SVM. The accuracy rates achieved with BPNN and SVM classifiers were 91.6% and 88.8%, respectively. Aljalal and Djemal [172] have implemented three classifiers using both CSP and WT. The authors found that LDA was more powerful with the features obtained from CSP and resulted in an accuracy rate of 84.79%; in contrast, SVM together with the WT method, achieved only 82.64% accuracy.

Two additional studies [34,124] have examined the performance of CSP and LDA for feature extraction and classification, respectively. In reference [124], the results indicated an 89.84% classification accuracy with the BCI-Competition IV data set. Similarly, a previous report [34] has indicated a 74.69% classification accuracy with EEG brain signals recorded from eight pianists. In addition to CSP, WT is a powerful tool that can be used to extract desired features in MI tasks. Sreeja et al. [61] have assessed the performance of the NB algorithm with WT and found a much higher classification accuracy than those obtained with either LDA or SVM classifiers. Ines et al. [174] and Maswanganyi et al. [52] have performed similar experiments and analyzed performance differences in ML classifiers by using the BCI Competition data set. Both studies have reported high accuracy when using WT with SVM and WT with NB, respectively.

One group of researchers has implemented three classifiers (SVM, LR, and KNN) to distinguish real from imagined movements by using EEG signals [195]. In this case, the LR algorithm outperformed both SVM and KNN, with overall accuracy rates ranging from 37% to 90%. Behri et al. [104] have compared several algorithms that might be used to differentiate EEG signals from the right foot and the right hand in five study participants. WPD was used to extract the features, and the RF classifier and WPD achieved a maximum accuracy of 98.45% in all participants. Similarly, Attallah et al. [259] have applied four levels of WPD to decompose EEG signals. Four different features were extracted, which were then introduced into the ML algorithm. With the SVM classifier, the highest accuracy of 93.46% and 86% was achieved for the BCI Competition III-IVa data set and the autocalibration and recurrent adaptation data set, respectively. Furthermore, a multi-class Adaboost-ELM algorithm based on features extracted by Kc has been used to classify three states (i.e., left-hand movement, right-hand movement, and resting-state) in ten participants [210]. Moreover, the performance of SVM was compared with that of Fisher linear discriminant, BPNN, and radial basis function neural network (RBF-NN) to achieve optimal performance of classification [260]. Yang et al. [261] have proposed a new framework based on multiple Riemannian covariances and MLP for feature extraction and classification, respectively. Use of this framework to classify MI EEG signals achieved a mean accuracy of 76%.

Several reviewed studies have used convolutional layers for EEG-based MI tasks. Sun et al. [262] have proposed the sparse spectrotemporal decomposition (SSD) algorithm for feature extraction and the CNN classifier with a 1-D convolution layer. Experimental results on BCI Competition IV and Tianjin University (TJU) data sets have achieved 79.3% and 85.7% accuracy, respectively. In reference [263], Liu et al. have considered the original EEG signals and their wavelet power spectrum as model input and reported that a deep CNN architecture based on space-time features and time-frequency features significantly improved the average accuracy performance of four-class MI classification. Dai et al. [264] have designed a hybrid architecture in which a convolutional layer CNN was used to learn network parameters, and the extracted features were then fed into the variational autoencoder (VAE) network.

Overall, we found no comprehensive comparisons of methods used as classifiers for MI tasks. Nevertheless, CNN, SVM, and NB outperformed other ML and DL classifiers

in the articles reviewed in this study. On the basis of the findings, we suggest that CNN, SVM, and NB classifiers together with either CSP or WT extraction methods might be used for MI tasks.

#### 5.4. Seizure Detection Task

Epilepsy is a neurological disorder whose diagnosis with automated SD has recently received significant attention [189], because the features extracted from the EEG signals are advantageous. In this task, EEG signals were recorded occasionally from healthy participants and patients with epileptic symptoms [20]. Seven ML and DL algorithms were used in SD studies: CNN [265,266], SVM [116,118,179,180,182,184,191,192,198,200,207,267], KNN [189,268,269], ANNs [183,199], RF [185,187,190], LDA [186], and ELM [181]. However, among the 24 studies focused on seizures, 12 applied the SVM algorithm with various kernels. Murugappan and Ramakrishnan [118] have used a hierarchical multi-class SVM (H-MSVM) with an ELM kernel to classify SD. Likewise, linear, and Gaussian kernels have been applied for high-dimensional spaces, respectively [198,267]. The RBF kernel is a common function used in the SVM algorithm applied to different tasks, as described by Hamed et al. [179] and Jaiswal and Banka [180]. Furthermore, the most common feature extraction methods reported in the publications selected for our study were WT [118,179,183,184,198,199,267], EMD [116,182,192], PCA [180,189,190,207], and FFT [268,269]. Similarly, the statistical feature extraction method 1D-TP used widely for image processing [270] has been applied [187] to generate the lower and upper features of each signal. Maximum difference of amplitude distribution histogram (MDADH) is a supervised feature selection method based on amplitude distribution histograms generated in both preictal and non-preictal trials. This method has been used by Bandarabadi et al. [200] to select and extract the desired EEG signal features. Further details on amplitude distribution histograms have been provided in previous publications [271,272]. The use of VMD, an extension of EMD, has been proposed by Chakraborty and Mitra [185]. The objective of VMD is to decompose input signals into sub-signals called modes. Because of the difficulties involved in selecting the appropriate number of modes and the associated penalty coefficient, the authors have proposed a kurtosis-based method that can be used to select the optimal number of modes and the penalty coefficient. More details on the VMD method can be found in reference [83].

As documented in Table 5, three shared SD databases have been used in the studies featured in this review: the BONN database, the CHB-MIT database, and the European epilepsy database. Some studies using the BONN database achieved near-perfect classification accuracy, reaching 100% with an SVM classifier [179,180,184] or RF classifier [185], and 97.7% accuracy with CNN [265,266]. In recent studies, residual CNN models using raw EEG signals as input [265,266] have also achieved competitive performance. In reference [180], feature extraction has been performed by using both sub-pattern-based PCA (SpPCA) and cross-sub-pattern correlation-based PCA (SubXPCA) methods; the extracted features were then sent to an SVM with an RBF kernel for further analysis. The feature extraction methods DWT and WPD have been used in references [179,181], respectively. In studies using the CHB-MIT database [190,191], the classification performance has reached 97.12% sensitivity with RF with PCA as a feature extraction method and 96% sensitivity with SVM. In addition, the use of SVM and ANNs classifiers resulted in nearly identical performance in studies using the publicly available European epilepsy database [199,200], in which an average sensitivity value of 73.5% was achieved.

Because we observed only trivial performance differences when comparing studies performed on data from each of the shared databases, we concluded that further research is needed to identify the most effective ML and DL algorithms for seizure detection. Nevertheless, as shown in Table 5, both CNN and SVM are robust algorithms that can be used to detect abnormalities from biomedical signals, and RF and ANNs are the most commonly used classifiers used in the studies included herein that focused on SD tasks.



Thus, we recommend CNN, SVM, ANNs, and RF as good candidate ML and DL classifiers for this type of task.

**Table 5.** Comparison of studies conducted for the seizure detection task by using shared data sets.

Database	Authors	Year	Feature Extraction Method	Classification	Performance (%)
BONN	Hamed et al. [179]	2018	DWT	RBF-SVM	Accuracy = 100
	Savadkoochi et al. [184]	2020	FFT / WT	SVM	Accuracy = 100
	Jaiswal and Banka [180]	2018	PCA	RBF-SVM	Accuracy = 100
	Riaz et al. [182]	2015	EMD	SVM	High performance in the detection of seizures in case 1 and 2
	Ullah et al. [266]	2018	-	1D-CNN	Accuracy = 99.1
	Lu and Triesch [265]	2019	-	Residual CNN	Accuracy = 99
	Chakraborty and Mitra [185]	2021	VMD	RF	Accuracy = 98.7–100
	Ech-Choudany et al. [186]	2021	Dissimilarity-based TFD	LDA	Accuracy = 98
	Mardini et al. [183]	2020	DWT	ANNs	Accuracy = 97.8
	Liu et al. [181]	2017	WPD	ELM	Accuracy = 97.7
CHB-MIT	Murugappan and Ramakrishnan [118]	2016	WT	H-MSVM	Accuracy = 94
	Kaya and Ertugrul [187]	2018	1D-TP	RF	Accuracy > 94
	Pinto-Orellana and Cerqueira [190]	2016	PCA	RF	Sensitivity = 97.1 Specificity = 99.2
	Shoeb et al. [191]	2011	-	SVM	Sensitivity = 96 Specificity = 93
European Epilepsy	Fergus et al. [189]	2015	PCA	KNN	Sensitivity = 94 Specificity = 94
	Usman et al. [192]	2017	EMD	SVM	Sensitivity = 92.2 Specificity = 93.4
	Teixeira et al. [199]	2014	WT	ANNs	Sensitivity = 73.1
	Direito et al. [198]	2017	DWT	Linear-SVM	High performance in a small subset of participants
	Bandarabadi et al. [200]	2015	MDADH	SVM	Sensitivity = 73.98

### 5.5. Sleep Stage Scoring Task

A sleep stage is a period of time in which the sleep process remains constant. Sleep researchers focus on two main stages of sleep: rapid eye movement (REM) and non-rapid eye movement (NREM) [273]. NREM is divided into four stages (stages 1–4), each of which has specific characteristics. EEG signals are recorded during sleep and scored by experts that can classify them into REM or one of the four NREM sleep stages by using various detection methods.

Several ML techniques for automated SS tasks have been applied in the articles included in this review (Table 6). Ebrahimi et al. [274] have identified four sleep stages, extracted features based on WT coefficients, and applied MLP with eight neurons in one hidden layer, achieving 93% accuracy. Similarly, Zoubek et al. [275] have compared the performance between FFT and WT with two different classifiers: KNN and MLP. High accuracy was achieved with MLP with six neurons in one hidden layer when FFT was used as the feature extraction method. In another study [276], the researchers obtained time- and frequency-domain features from polysomnography signals by using dendrogram-based SVM as the classifier. The specificity, sensitivity, and accuracy of the classification of five sleep stages reached 94%, 82%, and 92%, respectively. Correspondingly, dendrogram-based multi-class SVM and WT have been used in one study [202] to classify three sleep states (light sleep and REM, deep sleep, and the awake) with an accuracy of 91.4%. Kuo and Liang [277] have proposed multiscale permutation entropy analysis for sleep scoring tasks along with the AR model and LDA, and have achieved a sensitivity of 89.1% for ten participants. Similarly, Santaji and Desai [59] have compared the performance of three classifiers: RF, SVM, and DT. The RF algorithm, when trained with extracted statistical features outperformed the other two algorithms in classification, with high specificity, sensitivity, and accuracy of 96.35%, 96.12%, and 97.8%, respectively. Delimayanti et al. [204] have applied FFT to elicit high-dimensional features and enhanced the classification performance



of SS tasks by using the SVM algorithm with RBF kernel, and have achieved an average accuracy of 87.84%. Moreover, the performance of the AdaBoost classifier has been analyzed in reference [203,278]. Hassan and Bhuiyan [203] extracted time-frequency features by using the EMD method and compared the performance of different ML classifiers for SS task on the basis of a single channel EEG signal. AdaBoost significantly outperformed the other algorithms, with an accuracy of 92.24%.

**Table 6.** Comparison of studies for the sleep stage scoring task, including sleep stages, feature extraction method, machine learning algorithm, and overall performance.

Authors	Year	Sleep Stages	Feature Extraction Method	Classification	Performance (%)
Santaji and Desai [59]	2020	S1, S2, REM	Entropy	RF	Accuracy = 97.8
Ebrahimi et al. [279]	2008	Awake, S1 and REM, S2, SWS	WT	MLP	Accuracy = 93
Hassan and Bhuiyan [203]	2016	Awake, S1, S2, S3, S4, REM	EMD	AdaBoost	Accuracy = 92.2
Lajnef et al. [276]	2015	Awake, S1, S2, SWS, REM	Entropy	Dendrogram-SVM	Accuracy = 92
Ravan [202]	2019	Awake, LS and REM, DS	WT	Dendrogram-SVM	Accuracy = 91.4
Kuo and Liang [277]	2011	Awake, S1, S2, SWS, REM	Entropy/AR	LDA	Sensitivity = 89.1
Delimayanti et al. [204]	2020	Awake, S1, S2, S3, S4, REM	FFT	RBF-SVM	Accuracy = 87.8
Zoubek et al. [275]	2007	Awake, NREM1, NREM2, SWS, PS	FFT	MLP	Accuracy = 71.6

### 5.6. Neurodegenerative Disease Task

Chronic pain is a brain disease that often occurs in older people [280] with a diagnosis of ND, such as AD, PD, and SZ [281]. According to reference [280], chronic pain in older patients may reduce memory extinction and increase the resilience of pain memory, as discussed by Battaglia et al. [282], who have demonstrated that older individuals have reduced extinction of fear memories. Analysis of EEG is a well-established method that can be used to detect brain abnormalities associated with these diseases. To perform this task, EEG data were recorded for several hours from both healthy participants and patients with these disorders to create a large data set. Analysis of brain signals to diagnose ND has been proposed in only a small number of selected studies. However, ML and DL applications offer new and potentially highly accurate approaches that might be used to diagnose brain abnormalities during early stages of the disease.

In the eight articles on AD selected for this review, the number of participants per study varied between 35 and 189, and the samples of patients with AD, patients with Mild Cognitive Impairment (MCI), and Healthy Controls (HCs) were well balanced. Four of these studies [117,283–285] explored the differences between patients with AD and HCs, whereas the other publications [58,102,274,286] considered all three groups. EEG band-pass filtering is a common strategy used to improve the spectral components of EEG signals; five of the eight publications have reported an EEG bandwidth at or below 40 Hz. The publications included in this review used three EEG feature extraction methods—spectral entropy, FFT, and WT—for the characterization of AD. As reported by Kulkarni [283,284], the combination of spectral entropy and an SVM classifier, compared with that of KNN classifier, resulted in outstanding performance achieving an accuracy of 96%. Two groups have used FFT to extract the EEG features [117,286]. Fison et al. [286] have applied FFT to the EEG signals and compared the outcomes associated with various classifiers, including SVM, DT, and rule-based classifiers. The results revealed that the use of DT to classify MCI versus HCs, AD versus MCI, and AD versus HCs was superior to the use of other classifiers, and achieved an accuracy of 90%, 80%, and 73%, respectively. Likewise, FFT has been used in one study [117] together with SVM to differentiate patients with AD from HCs, with an accuracy of 87%. WT has been developed into discrete and continuous WT feature extraction methods. The combination of DWT with DT has been examined in reference [102]. In addition, in reference [58], continuous wavelet transform (CWT) with MLP achieved the best accuracy, at 95.76% and 86.84% for AD versus HCs and AD versus MCI, respectively. Oltu et al. [274] have proposed an algorithm including three main steps that can be used to elicit the relevant features: DWT to extract EEG sub-bands,

Burg's method to measure the PSD of each sub-band, and a means for determining the amplitude summation of the coherence values for sub-bands. In this study, Bagged Trees outperformed the other classifiers, with a classification accuracy of 96.5%. CNN architecture, used by Morabito et al. [285] with two hidden convolutional layers to extract features of multi-channel EEG signals, has been reported to achieve an 82% accuracy for three-class classification.

Two PD-related articles were considered in the ND category. These studies were performed on experimental data sets with 20 or 18 participants, in references [209,287], respectively. According to the study published in reference [287], FFT was used as a feature extraction method together with a KNN classifier to differentiate between PD patients and HCs, with 88% accuracy. HOS, a powerful method to extract the nonlinear EEG signal features [288], was introduced by Yuvaraj et al. [209]. In that study, the use of RBF-SVM resulted in an accuracy of 99.62%.

Only a very small fraction of the publications included in this review were SZ-related studies. In one study [212], wavelet-based features were elicited from a single channel and classified with the KNN algorithm. Classification accuracy values of 99.21% and 97.2% were obtained with 10-fold and leave-one-subject-out cross-validation methods, respectively. Other researchers have designed RF classifiers by using features extracted based on ERP [289]. With 10-fold cross-validation, the best classification accuracy of 81.10% was achieved.

## 6. Future Directions

Several promising future directions exist in the implementation of EEG signal classification. Most recent research has focused on using DL algorithms that require increasing the amount of data and changing the structure of the model [290]. Although DL models can effectively solve the EEG signal classification tasks, transfer learning strategy from one model to another accelerates training time and yields the best performance results [291]. Another exciting prospect is applying a graph neural network (GNN) framework to consider the brain connectivity network by identifying Regions of Interests [292]. DL models cannot directly work on graph-structured input data because they consider the brain network features as a vector of one dimension [293]. The human brain connectivity represents the brain as a graph with interacting nodes in non-Euclidean space, and existing DL methods generally disregard the interaction and association of brain connectivity networks [294]. GNNs aim to learn graph representations by using a neural network and to pass information via a message-passing algorithm [295,296]. Unlike neural networks, GNNs update the representations of nodes while maintaining the graph topology.

## 7. Conclusions

On the basis of our review of 128 published articles, various supervised machine learning and deep learning algorithms have been widely applied in various tasks, including ER, MWL, MI, SD, SS, and diagnosis of ND. Several metrics can affect the performance of classifiers, including different data sets, preprocessing techniques, and feature extraction methods. We presented an overview of feature extraction methods as part of our findings addressing RQ2. We also introduced the publicly available databases that have frequently been used for each task, and we directly analyzed the classification performance reported in relevant studies. Many of the reviewed studies have compared the performance of different classifiers; CNN, SVM, and KNN were the most frequently used classifiers across all articles reviewed. Although model performance can be attributed to a variety of factors, our findings suggested that SVM and KNN outperformed the other supervised ML classifiers. We also found that CNN and NB had impressive performance in studies focused on MI tasks when either CSP or WT was used as a feature extraction methods. Similarly, the performance of RF was superior to that of the other classifiers in studies focused on ER tasks with the DEAP database. This systematic review provided recommendations for applying supervised machine learning and deep learning algorithms for the neural

decoding of EEG signals in various tasks and experimental protocols. Although each classification algorithm has its own strengths and limitations, these recommendations provide insight into the issues associated with the classification of EEG signals, which might be addressed in future research efforts in this field. Further in-depth studies combining the selection of feature extraction methods and types of classifiers are highly recommended.

**Author Contributions:** Conceptualization, methodology, writing—original draft preparation and revisions: M.S.; conceptualization, funding acquisition, writing—review and revisions, editing, and supervision: W.K.; review and revision, editing, and methodology: F.V.F. and K.F.; review and editing, investigation: R.T., P.A.H. and A.A.-J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applied to this study.

**Informed Consent Statement:** Not applied to this study.

**Data Availability Statement:** Not applied to this study.

**Acknowledgments:** The authors would like to thank Pamela K. Douglas for providing valuable comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

Table A1 summarizes the relevant information from the selected articles, including the task information, database, frequency range, feature extraction method, ML/DL algorithm, and performance.

**Table A1.** Summary of the ML/DL Algorithm-related studies included in this systematic review.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[245]	2010	MWL 7 subj. 6 channels	Own database	Entropy	IFWSVM	Accuracy = 97.5
[176]	2020	MI 30 channels	BCI Competition III and IV	CSP	SVM	Accuracy = 92.2
[239]	2018	MWL 8 subj. 1 channel	Own database	FFT	Cubic SVM KNN LDA	Accuracy = 95
[81]	2014	MWL 50 subj.	Own database	FFT	ANNs	Accuracy = 88.9
[284]	2019	AD 100 subj. 24 channels	Own database	Entropy	SVM KNN	Accuracy = 96
[244]	2016	MWL 6 subj.	Own database	HHT	SVM LDA QDA KNN	Accuracy = 89.07
[253]	2010	MWL 13 subj. 20 channels	Own database	PCA	RBF-SVM	-
[195]	2019	MI 11 subj. 3 channels	PhysioNet database	N/A	LR SVM KNN	Accuracy = 90
[118]	2016	SD 22 subj. 128 channels	BONN database	WT	H-MSVM	Accuracy = 94

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[192]	2017	SD 24 subj. 23 channels	CHB-MIT database	EMD	SVM	Sensitivity = 92.2
[191]	2011	SD 22 subj.	CHB-MIT database	N/A	SVM	Sensitivity = 96
[102]	2018	AD 109 subj. 19 channels	Own database	DWT/FFT	DT J48	Accuracy = 91.7
[243]	2009	MWL 4 subj. 6 channels	Own database	AR/DWT	DSTKNN	Accuracy = 93.04
[279]	2008	SS 7 subj.	Sleep-EDF database	WT	MLP	Accuracy = 93
[248]	2012	MWL 6 subj. 6 channels	Own database	EMD	SVM LDC QDC KNN	Accuracy = 94.3
[104]	2018	MI 5 subj.	BCI Competition III	WPD	RF KNN SVM	Accuracy = 98.45
[225]	2020	ER 40 subj. 14 channels	Own database	TQWT	PNN ELM KNN RF DT	Accuracy = 96.16
[179]	2018	SD 25 subj. 100 channels	BONN database	DWT	RBF-SVM KNN NB	Accuracy = 100
[275]	2007	SS 41 subj. 1 channel	Own database	FFT	MLP KNN	Accuracy = 71.56
[32]	2019	MWL 8 subj. 14 channels	Own database	N/A	XGBoost MLP KNN SVM DT NB	Accuracy = 88
[183]	2020	SD	BONN database	DWT	ANNs SVM KNN NB	Accuracy = 97.82
[194]	2013	MI 6 subj. 8 channels	PhysioNet database	ICA	SVM ANNs	Accuracy = 97.1
[54]	2008	MWL 15 subj. 60 channels	Own database	PCA	SVM	Accuracy = 71.7
[251]	2007	MWL 7 subj. 6 channels	Own database	AR	RBF-SVM	Accuracy = 70
[229]	2020	ER 28 subj. 64 channels	Own database	DWT/EMD	ANNs KNN SVM	Accuracy = 94.3
[297]	2016	ERP 52 subj. 64 channels	Own database	WT	SVM	Accuracy = 87
[151]	2014	ERP 7 subj. 49 channels	Own database	ICA	SVM	Accuracy = 87

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[228]	2004	ER 12 subj. 3 channels	Own database	Statistics	SVM	Accuracy = 66.7
[249]	2011	MWL 6 subj. 32 channels	Own database	PCA	SVR RBF-NN LR	Accuracy = 86.92
[252]	2006	MWL 7 subj.	Own database	AR	SVM ELM	Accuracy = 67.57
[274]	2021	AD 32 subj. 19 channels	Own database	DWT	Bagging DT KNN SVM	Accuracy = 96.5
[286]	2014	AD 100 subj. 19 channels	Own database	FFT	DT J48 SVM	Accuracy = 90
[189]	2015	SD 22 subj.	CHB-MIT database	PCA	KNN	Sensitivity = 93
[218]	2007	ER 17 subj.	Own database	N/A	KNN	Accuracy = 82.27
[61]	2017	MI 5 subj. 30 channels	BCI Competition III	DWT/AR	NB LDA SVM	Accuracy = 95.47
[174]	2013	MI 3 channels	BCI Competition II	WT	SVM LDA MLP	Accuracy = 90
[84]	2020	MWL 20 subj.	Own database	HHT	SVM	Accuracy = 84.8
[289]	2019	SZ 81 subj. 9 channels	Own database	N/A	RF	Accuracy = 81.1
[298]	2019	MI 12 subj. 8 channels	Own database	N/A	ANNs	Accuracy = 90
[212]	2021	SZ 28 subj. 19 channels	Own database	RMSFMS filter	KNN SVM Bagged trees Boosted trees	Accuracy = 99.21
[199]	2014	SD 224 subj. 6 channels	European Epilepsy database	WT	ANNs SVM	Sensitivity = 73.08
[149]	2017	ERP 108 subj.	Own database	TFHA	SVM-RFE	Accuracy = 99
[207]	2012	SD 19 subj. 6 channels	Freiburg database	PCA	SVM	Sensitivity = 85.5
[276]	2015	SS 15 subj. 21 channels	Own database	Entropy	Dendrogram- SVM	Accuracy = 92
[55]	2018	MWL 8 subj. 14 channels	Own database	PCA	KNN SVM LR DT	Accuracy = 70.6
[167]	2018	ER 32 subj. 32 channels	DEAP database	DWT	KNN	Accuracy = 87.1
[261]	2020	MI 9 subj. 22 channels	BCI Competition IV	MRC	MLP	Accuracy = 76



Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[277]	2011	SS 20 subj. 6 channels	Own database	Entropy/AR	LDA	Sensitivity = 89.1
[202]	2019	SS 67 subj. 2 channels	Sleep-EDF database	WT	Dendrogram-SVM	Accuracy = 91.4
[31]	2012	ER 32 subj. 32 channels	DEAP database	N/A	NB	Accuracy = 66.5
[299]	2016	MWL 15 subj. 3 channels	Own database	N/A	SVM KNN ANNs	Accuracy = 95.21
[33]	2018	Depression 23 subj. 19 channels	Own database	HFD/Entropy	MLP LR SVM DT RF NB	Accuracy = 97.56
[58]	2019	AD 189 subj. 19 channels	Own database	CWT	MLP LR SVM	Accuracy = 95.76
[164]	2020	ER 32 subj. 7 channels	DEAP database	PCA	KNN LR DT SVM LDA	Accuracy = 74.25
[124]	2021	MI 23 channels	BCI Competition IV	CSP	LDA	Accuracy = 89.84
[163]	2018	ER 32 subj. 10 channels	DEAP database	PCA	RBF-SVM KNN ANNs	Accuracy = 91.2
[52]	2018	MI 9 subj. 2 channels	BCI Competition IV	DWT	NB KNN LDA	Accuracy = 73
[300]	2010	MI 6 channels	N/A	Statistics	LDA BPNN SVM	Accuracy = 88.6
[162]	2019	ER 32 subj. 32 channels	DEAP database	Entropy	RF RBF-SVM LDA KNN	Accuracy = 90
[301]	2016	MWL 20 subj. 63 channels	Own database	Statistics	SVM	Accuracy = 70
[168]	2020	ER 32 subj. 14 channels	DEAP database	Statistics	SVM KNN DT	Accuracy = 77.6–78.9
[153]	2019	Anxiety 28 subj. 4 channels	Own database	N/A	RF LR MLP	Accuracy = 78.5
[66]	2016	Depression 27 subj. 30 channels	Own database	WT/Statistics	RBF-SVM	Accuracy > 80
[152]	2012	ERP 3 subj. 14 channels	Own database	Statistics	ANNs	Accuracy = 80

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[203]	2016	SS	Sleep-EDF database	EMD	AdaBoost NB LDA ANNs SVM KNN	Accuracy = 92.24
[53]	2019	MI 6 subj. 14 channels	Own database	CSP	SVM	-
[182]	2015	SD 22 subj. 100 channels	BONN database	EMD	SVM	-
[34]	2020	MI 8 subj. 8 channels	Own database	CSP	LDA	Accuracy = 74.69
[159]	2019	Sleep Apnea 16 subj.	MIT-BIH database	HHT	SVM KNN ANNs	Accuracy = 99
[209]	2018	PD 20 subj. 14 channels	Own database	HOS	SVM DT KNN NB PNN	Accuracy = 99.6
[219]	2009	ER 10 subj. 62 channels	Own database	CSP	SVM	Accuracy = 93.5
[226]	2008	ER 10 subj.	Own database	Statistics	SVM ANNs NB	Accuracy = 80
[59]	2020	SS 125 subj.	Sleep-EDF database	Entropy	RF SVM DT	Accuracy = 97.8
[180]	2018	SD 25 subj. 100 channels	BONN database	PCA	RBF-SVM	Accuracy = 100
[185]	2021	SD	BONN database	VMD	RF	Accuracy = 98.7–100
[204]	2020	SS 2 channels	Sleep-EDF database	FFT	RBF-SVM	Accuracy = 87.8
[196]	2020	MWL 36 subj. 19 channels	PhysioBank database	FFT	KNN SVM	Accuracy = 99.4
[154]	2020	Stress 33 subj. 5 channels	Own database	FFT	SVM LR NB KNN DT	Accuracy = 85.2
[165]	2020	ER 32 subj.	DEAP database	STFT	SVM KNN DT	-
[210]	2016	MI 10 subj. 64 channels	Own database	Kolmogorov complexity (Kc)	AdaBoost-ELM	Accuracy = 79.5
[260]	2005	MI 4 subj. 62 channels	Own database	ICA/PCA	SVM ANNs	Accuracy = 77.3
[171]	2015	MI	BCI Competition III and IV	CSP	Twin-SVM	Accuracy = 100

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[150]	2017	ERP 69 subj. 4 channels	Own database	FFT	GB RF RBF-SVM	Accuracy = 74
[287]	2019	PD 18 subj. 128 channels	Own database	FFT	KNN SVM	Accuracy = 88
[258]	2020	MI 5 subj. 64 channels	Own database	CSP	LDA SVM LR NB	Accuracy = 81
[259]	2020	MI	BCI Competition III database (1) Autocalibration and Recurrent Adaptation database (2)	WPD	SVM LDA KNN	Accuracy = 93.46 (1) Accuracy = 86 (2)
[157]	2016	Tinnitus 22 subj. 129 channels	Own database	FFT	SVM	Accuracy = 90.9
[278]	2016	SS 5 subj.	Own database	FFT/DWT	AdaBoost	-
[198]	2017	SD 216 subj. 6 channels	European Epilepsy database	DWT	SVM	-
[156]	2020	Alcoholism detection 2 subj. 64 channels	UCI database	EWD	SVM NB KNN	Accuracy = 98.75
[155]	2013	Depression 90 subj. 19 channels	Own database	FFT	LR LDA KNN	Accuracy = 90
[267]	2020	SD	Own database	DWT	RBF-SVM KNN DT LR	Accuracy = 100
[119]	2019	ER 6 subj. 32 channels	DEAP database	Statistics	RF RBF-SVM KNN NB ANNs DT	Accuracy = 98.2
[117]	2011	AD 32 subj.	Own database	FFT	SVM	Accuracy = 86.97
[116]	2020	SD 24 subj. 22 channels	N/A	EMD	SVM	Accuracy > 90
[191]	2016	SD 22 subj.	CHB-MIT database	PCA	RF	Accuracy = 98.3
[166]	2013	ER 32 subj. 32 channels	DEAP database	PCA	ANNs RBF-SVM	Accuracy > 60
[211]	2018	MWL 10 subj. 14 channels	Own database	FFT/WT	RF SVM MLP	Accuracy = 85–99
[200]	2015	SD 24 subj. 6 channels	European Epilepsy database	MDADH	SVM	-
[230]	2020	ER 20 subj. 24 channels	Own database	TQWT	MC-LSVM	Accuracy = 95.7

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[175]	2019	MI 4 subj. 59 channels	BCI Competition 2008	CSP	BPNN SVM	Accuracy = 91.6
[186]	2021	SD	BONN database	Dissimilarity-based TFD	LDA ANNs SVM	Accuracy = 98
[240]	2018	MWL	OneR database	FFT/Entropy	RF SVM	Accuracy = 87.2
[65]	2011	ER 20 subj. 62 channels	Own database	DWT	KNN	Accuracy = 82.87
[241]	2017	MWL	Own database	FFT	KNN SVM ANNs	Accuracy = 90.5
[173]	2014	MI 7 subj. 22 channels	BCI Competition 2008	EMD	RBF-SVM	Accuracy = 100
[181]	2017	SD 10 subj.	BONN database	WPD	ELM	Accuracy = 97.7
[268]	2017	SD 23 subj.	Own database	FFT	KNN	Sensitivity = 80.9
[187]	2018	SD	BONN database	1D-TP	RF SVM ANNs	Accuracy > 94
[227]	2019	ER 28 subj. 2 channels	Own database	N/A	KNN MLP RF	Accuracy = 86.7
[172]	2017	MI 5 subj.	BCI Competition III	CSP/DWT	LDA SVM ANNs	Accuracy = 84.8
[269]	2020	SD 10 subj.	Own database	DFT	WBCKNN	Accuracy = 99
[160]	2020	Creativity 20 subj. 32 channels	Own database	CSP	QDA SVM	Accuracy = 82
[250]	2020	MWL 7 subj. 6 channels	Keirn and Aunon database	WT/EMD	SVM KNN	Accuracy = 80–100
[184]	2020	SD 10 subj.	BONN database	FT/WT	SVM KNN	Accuracy = 100
[158]	2020	ADHD 97 subj. 19 channels	Own database	PSR	NDC EPNN SVM	Accuracy = 100
[177]	2020	MI 5 subj.	BCI Competition III	PCA	H-KELM	Accuracy = 96.5
[224]	2017	ER 32 subj. 8 channels	DEAP database	STFT	CNN ReLU and Softmax (FC) as activation functions	Accuracy = 87.3
[285]	2016	AD	Own database	WT	MC-DCNN Sigmoid and Softmax (FC)	Accuracy = 82
[302]	2020	MI	BCI Competition IV	CSP	DNN	Accuracy = 83.98

Table A1. Cont.

Article	Year	TaskInformation	Database	Feature Extraction Method	ML/DL Algorithm	Performance
[262]	2020	MI	BCI Competition IV (1) TJU database (2)	SSD	CNN ReLU as activation functions	Accuracy = 79.3 (1) Accuracy = 85.7 (2)
[265]	2019	SD	BONN database (1) Bern-Barcelon database (2)	N/A	Residual-CNN ReLU and Softmax (FC) as activation functions	Accuracy = 99 (1) Accuracy = 92 (2)
[247]	2020	MWL 48 subj.	STEW database	N/A	LSTM CNN + LSTM	Accuracy = 61.08
[263]	2021	MI	BCI Competition IV (1) HGD (2)	N/A	CNN	Accuracy = 81.6 (1) Accuracy = 95.5 (2)
[266]	2018	SD 5 subj.	BONN database	N/A	P-1D-CNN ReLU and Softmax (FC) as activation functions	Accuracy = 99.1
[242]	2018	MWL 15 subj.	Own database	FFT	CNN ReLU and Softmax (FC) as activation functions	Accuracy = 90
[264]	2019	MI	BCI Competition IV (1) Own database (2)	STFT	CNN-VAE	Kappa = 0.564 (1) Kappa = 0.603 (2)

## References

- Herculano-Houzel, S. The human brain in numbers: A linearly scaled-up primate brain. *Front. Hum. Neurosci.* **2009**, *3*, 31. [[CrossRef](#)] [[PubMed](#)]
- Pakkenberg, B.; Pelvig, D.; Marner, L.; Bundgaard, M.J.; Gundersen, H.J.G.; Nyengaard, J.R.; Regeur, L. Aging and the human neocortex. *Exp. Gerontol.* **2003**, *38*, 95–99. [[CrossRef](#)]
- Al-Kadi, M.I.; Reaz, M.B.I.; Ali, M.A.M. Evolution of electroencephalogram signal analysis techniques during anesthesia. *Sensors* **2013**, *13*, 6605–6635. [[CrossRef](#)]
- Molfese, D.L.; Molfese, V.J.; Kelly, S. The use of brain electrophysiology techniques to study language: A basic guide for the beginning consumer of electrophysiology information. *Learn. Disabil. Q.* **2001**, *24*, 177–188. [[CrossRef](#)]
- Kayser, J.; Tenke, C.E. In search of the Rosetta Stone for scalp EEG: Converging on reference-free techniques. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.* **2010**, *121*, 1973. [[CrossRef](#)]
- Ismail, L.E.; Karwowski, W. A graph theory-based modeling of functional brain connectivity based on eeg: A systematic review in the context of neuroergonomics. *IEEE Access* **2020**, *8*, 155103–155135. [[CrossRef](#)]
- Henry, J.C. Electroencephalography: Basic principles, clinical applications, and related fields. *Neurology* **2006**, *67*, 2092. [[CrossRef](#)]
- Buzsaki, G. *Rhythms of the Brain*; Oxford University Press: Oxford, UK, 2006.
- Douglas, P.K.; Douglas, D.B. Reconsidering Spatial Priors In EEG Source Estimation: Does White Matter Contribute to EEG Rhythms? In Proceedings of the 2019 7th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Korea, 18–20 February 2019; pp. 1–12.
- Kumral, D.; Cesnaite, E.; Beyer, F.; Hofmann, S.M.; Hensch, T.; Sander, C.; Hegerl, U.; Haufe, S.; Villringer, A.; Witte, A.V.; et al. Relationship between Regional White Matter Hyperintensities and Alpha Oscillations in Older Adults. *bioRxiv* **2021**. [[CrossRef](#)]
- Siuly, S.; Li, Y.; Zhang, Y. EEG signal analysis and classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *11*, 141–144.
- Cohen, M.X. *Analyzing Neural Time Series Data: Theory and Practice*; MIT Press: Cambridge, MA, USA, 2014.
- Teplan, M. Fundamentals of EEG measurement. *Meas. Sci. Rev.* **2002**, *2*, 1–11.
- Kerr, W.T.; Anderson, A.; Lau, E.P.; Cho, A.Y.; Xia, H.; Bramen, J.; Douglas, P.K.; Braun, E.S.; Stern, J.M.; Cohen, M.S. Automated diagnosis of epilepsy using EEG power spectrum. *Epilepsia* **2012**, *53*, e189–e192. [[CrossRef](#)]
- Cooray, G.K.; Sengupta, B.; Douglas, P.; Englund, M.; Wickstrom, R.; Friston, K. Characterising seizures in anti-NMDA-receptor encephalitis with dynamic causal modelling. *Neuroimage* **2015**, *118*, 508–519. [[CrossRef](#)]
- Grosse-Wentrup, M.; Liefhold, C.; Gramann, K.; Buss, M. Beamforming in noninvasive brain-computer interfaces. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 1209–1219. [[CrossRef](#)]



17. Douglas, P.; Lau, E.; Anderson, A.; Kerr, W.; Head, A.; Wollner, M.A.; Moyer, D.; Durnhofer, M.; Li, W.; Bramen, J.; et al. Single trial decoding of belief decision making from EEG and fMRI data using independent components features. *Front. Hum. Neurosci.* **2013**, *7*, 392. [[CrossRef](#)]
18. Kriegeskorte, N.; Douglas, P.K. Interpreting encoding and decoding models. *Curr. Opin. Neurobiol.* **2019**, *55*, 167–179. [[CrossRef](#)] [[PubMed](#)]
19. Hasenstab, K.; Scheffler, A.; Telesca, D.; Sugar, C.A.; Jeste, S.; DiStefano, C.; Şentürk, D. A multi-dimensional functional principal components analysis of EEG data. *Biometrics* **2017**, *73*, 999–1009. [[CrossRef](#)]
20. Craik, A.; He, Y.; Contreras-Vidal, J.L. Deep learning for electroencephalogram (EEG) classification tasks: A review. *J. Neural Eng.* **2019**, *16*, 031001. [[CrossRef](#)] [[PubMed](#)]
21. Gilles, J.; Meyer, T.; Douglas, P. Leveraging Sparsity: A Low-Rank+ Sparse Decomposition (LR+ SD) Method for Automatic EEG Artifact Removal. *Proc. Int. Work. Sparsity Tech. Med. Imaging* **2014**, *2014*, 80–88.
22. Bigdely-Shamlo, N.; Mullen, T.; Kothe, C.; Su, K.M.; Robbins, K.A. The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* **2015**, *9*, 16. [[CrossRef](#)] [[PubMed](#)]
23. Jas, M.; Engemann, D.A.; Bekhti, Y.; Raimondo, F.; Gramfort, A. Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage* **2017**, *159*, 417–429. [[CrossRef](#)]
24. Cole, S.; Voytek, B. Cycle-by-cycle analysis of neural oscillations. *J. Neurophysiol.* **2019**, *122*, 849–861. [[CrossRef](#)]
25. Gramfort, A.; Strohmeier, D.; Haueisen, J.; Hämäläinen, M.S.; Kowalski, M. Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations. *NeuroImage* **2013**, *70*, 410–422. [[CrossRef](#)] [[PubMed](#)]
26. Clerc, M.; Bougrain, L.; Lotte, F. *Brain-Computer Interfaces 1: Methods and Perspectives*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
27. Sen, P.C.; Hajra, M.; Ghosh, M. Supervised classification algorithms in machine learning: A survey and review. In *Emerging Technology in Modelling and Graphics*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 99–111.
28. Kumar, S.; Yger, F.; Lotte, F. Towards adaptive classification using Riemannian geometry approaches in brain-computer interfaces. In Proceedings of the 2019 7th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Korea, 18–20 February 2019; pp. 1–6.
29. Lotte, F.; Bougrain, L.; Clerc, M. *Electroencephalography (EEG)-Based Brain-Computer Interfaces*; Wiley: Hoboken, NJ, USA, 2015; doi:10.1002/047134608X.W8278. [[CrossRef](#)]
30. Huang, X.; Xiao, J.; Wu, C. Design of Deep Learning Model for Task-Evoked fMRI Data Classification. *Comput. Intell. Neurosci.* **2021**, *2021*, 6660866. [[CrossRef](#)] [[PubMed](#)]
31. Chung, S.Y.; Yoon, H.J. Affective classification using Bayesian classifier and supervised learning. In Proceedings of the 2012 12th International Conference on Control, Automation and Systems, Jeju, Korea, 17–21 October 2012; pp. 1768–1771.
32. Anand, V.; Sreeja, S.; Samanta, D. An automated approach for task evaluation using EEG signals. *arXiv* **2019**, arXiv:1911.02966.
33. Cukic, M.; Pokrajac, D.; Stokic, M.; Radivojevic, V.; Ljubisavljevic, M. EEG machine learning with Higuchi fractal dimension and Sample Entropy as features for successful detection of depression. *arXiv* **2018**, arXiv:1803.05985.
34. Riquelme-Ros, J.V.; Rodríguez-Bermúdez, G.; Rodríguez-Rodríguez, I.; Rodríguez, J.V.; Molina-García-Pardo, J.M. On the better performance of pianists with motor imagery-based brain-computer interface systems. *Sensors* **2020**, *20*, 4452. [[CrossRef](#)]
35. Garofalo, S.; Timmermann, C.; Battaglia, S.; Maier, M.E.; Di Pellegrino, G. Mediofrontal negativity signals unexpected timing of salient outcomes. *J. Cogn. Neurosci.* **2017**, *29*, 718–727. [[CrossRef](#)]
36. Alexander, W.H.; Brown, J.W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **2011**, *14*, 1338–1344. [[CrossRef](#)]
37. Battaglia, S.; Garofalo, S.; di Pellegrino, G.; Starita, F. Revaluing the role of vmPFC in the acquisition of Pavlovian threat conditioning in humans. *J. Neurosci.* **2020**, *40*, 8491–8500. [[CrossRef](#)] [[PubMed](#)]
38. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, P. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)]
39. Higgins, J.P.; Altman, D.G.; Gøtzsche, P.C.; Jüni, P.; Moher, D.; Oxman, A.D.; Savović, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ* **2011**, *343*. [[CrossRef](#)] [[PubMed](#)]
40. Li, G.; Wu, J.; Xia, Y.; He, Q.; Jin, H. Review of semi-dry electrodes for EEG recording. *J. Neural Eng.* **2020**, *17*, 051004. [[CrossRef](#)]
41. Lopez-Gordo, M.A.; Sanchez-Morillo, D.; Valle, F.P. Dry EEG electrodes. *Sensors* **2014**, *14*, 12847–12870. [[CrossRef](#)] [[PubMed](#)]
42. Li, G.; Wang, S.; Li, M.; Duan, Y.Y. Towards real-life EEG applications: Novel superporous hydrogel-based semi-dry EEG electrodes enabling automatically ‘charge–discharge’ electrolyte. *J. Neural Eng.* **2021**, *18*, 046016. [[CrossRef](#)] [[PubMed](#)]
43. Zhou, Y.; Huang, S.; Xu, Z.; Wang, P.; Wu, X.; Zhang, D. Cognitive Workload Recognition Using EEG Signals and Machine Learning: A Review. *IEEE Trans. Cogn. Dev. Syst.* **2021**. [[CrossRef](#)]
44. Dehais, F.; Duprès, A.; Blum, S.; Drougard, N.; Scannella, S.; Roy, R.N.; Lotte, F. Monitoring pilot’s mental workload using ERPs and spectral power with a six-dry-electrode EEG system in real flight conditions. *Sensors* **2019**, *19*, 1324. [[CrossRef](#)]
45. Hu, L.; Zhang, Z. *EEG Signal Processing and Feature Extraction*; Springer: Berlin/Heidelberg, Germany, 2019.
46. Jiang, X.; Bian, G.B.; Tian, Z. Removal of artifacts from EEG signals: A review. *Sensors* **2019**, *19*, 987. [[CrossRef](#)]
47. Johal, P.K.; Jain, N. Artifact removal from EEG: A comparison of techniques. In Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 3–5 March 2016; pp. 2088–2091.
48. Urigüen, J.A.; Garcia-Zapirain, B. EEG artifact removal—state-of-the-art and guidelines. *J. Neural Eng.* **2015**, *12*, 031001. [[CrossRef](#)]

49. Zhang, C.; Tong, L.; Zeng, Y.; Jiang, J.; Bu, H.; Yan, B.; Li, J. Automatic artifact removal from electroencephalogram data based on a priori artifact information. *BioMed Res. Int.* **2015**, *2015*, 720450. [[CrossRef](#)]
50. Klados, M.A.; Papadelis, C.; Braun, C.; Bamidis, P.D. REG-ICA: A hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. *Biomed. Signal Process. Control* **2011**, *6*, 291–300. [[CrossRef](#)]
51. Cashero, Z. Comparison of EEG Preprocessing Methods to Improve the Classification of P300 Trials. Master's Thesis, Colorado State University, Fort Collins, CO, USA, 2011.
52. Maswanganyi, C.; Tu, C.; Owolawi, P.; Du, S. Discrimination of Motor Imagery Task using Wavelet Based EEG Signal Features. In Proceedings of the 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC), Mon Tressor, Mauritius, 6–7 December 2018; pp. 1–4.
53. Zhang, Y.; Li, Q.; Yan, H.; Geng, X. The Research of the feature extraction and classification algorithm Based on EEG signal of motor imagery. In Proceedings of the 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), Jishou, China, 14–15 September 2019; pp. 187–190.
54. Besserve, M.; Philippe, M.; Florence, G.; Laurent, F.; Garnero, L.; Martinerie, J. Prediction of performance level during a cognitive task from ongoing EEG oscillatory activities. *Clin. Neurophysiol.* **2008**, *119*, 897–908. [[CrossRef](#)] [[PubMed](#)]
55. Shah, M.; Ghosh, R. Classification and Prediction of Human Cognitive Skills Using EEG Signals. In Proceedings of the 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), Chennai, India, 22–24 March 2018; pp. 206–212.
56. Bin, G.; Gao, X.; Yan, Z.; Hong, B.; Gao, S. An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. *J. Neural Eng.* **2009**, *6*, 046002. [[CrossRef](#)] [[PubMed](#)]
57. Zhang, Z.; Li, H.; Mandic, D. Blind source separation and artefact cancellation for single channel bioelectrical signal. In Proceedings of the 2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN), San Francisco, CA, USA, 14–17 June 2016; pp. 177–182.
58. Ieracitano, C.; Mammone, N.; Bramanti, A.; Marino, S.; Hussain, A.; Morabito, F.C. A time-frequency based machine learning system for brain states classification via eeg signal processing. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
59. Santaji, S.; Desai, V. Analysis of EEG signal to classify sleep stages using machine learning. *Sleep Vigil.* **2020**, *4*, 145–152. [[CrossRef](#)]
60. Sifuzzaman, M.; Islam, M.R.; Ali, M. *Application of Wavelet Transform and Its Advantages Compared to Fourier Transform*; Vidyasagar University: Midnapore, India, 2009.
61. Sreeja, S.; Rabha, J.; Nagarjuna, K.; 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.
62. Kotte, S.; Dabbakuti, J.K. Methods for removal of artifacts from EEG signal: A review. *J. Phys. Conf. Ser.* **2020**, *1706*, 012093. [[CrossRef](#)]
63. Sweeney, K.T.; Ward, T.E.; McLoone, S.F. Artifact removal in physiological signals—Practices and possibilities. *IEEE Trans. Inf. Technol. Biomed.* **2012**, *16*, 488–500. [[CrossRef](#)]
64. Kher, R.; Gandhi, R. Adaptive filtering based artifact removal from electroencephalogram (EEG) signals. In Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSPP), Tamilnadu, India, 6–8 April 2016; pp. 561–564.
65. Murugappan, M. Human emotion classification using wavelet transform and KNN. In Proceedings of the 2011 International Conference on Pattern Analysis and Intelligence Robotics, Kuala Lumpur, Malaysia, 28–29 June 2011; Volume 1, pp. 148–153.
66. Puk, K.M.; Gandy, K.C.; Wang, S.; Park, H. Pattern classification and analysis of memory processing in depression using EEG signals. In Proceedings of the International Conference on Brain Informatics, Omaha, NE, USA, 13–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 124–137.
67. Al-Fahoum, A.S.; Al-Fraihat, A.A. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *Int. Sch. Res. Not.* **2014**, *2014*, 730218. [[CrossRef](#)]
68. Aggarwal, S.; Chugh, N. Signal processing techniques for motor imagery brain computer interface: A review. *Array* **2019**, *1*, 100003. [[CrossRef](#)]
69. Boostani, R.; Karimzadeh, F.; Nami, M. A comparative review on sleep stage classification methods in patients and healthy individuals. *Comput. Methods Programs Biomed.* **2017**, *140*, 77–91. [[CrossRef](#)]
70. Torres, E.P.; Torres, E.A.; Hernández-Álvarez, M.; Yoo, S.G. EEG-based BCI emotion recognition: A survey. *Sensors* **2020**, *20*, 5083. [[CrossRef](#)]
71. Xiong, Q.; Zhang, X.; Wang, W.F.; Gu, Y. A parallel algorithm framework for feature extraction of EEG signals on MPI. *Comput. Math. Methods Med.* **2020**, *2020*, 9812019. [[CrossRef](#)] [[PubMed](#)]
72. Aboalayon, K.A.I.; Faezipour, M.; Almuhammadi, W.S.; Moslehpour, S. Sleep stage classification using EEG signal analysis: A comprehensive survey and new investigation. *Entropy* **2016**, *18*, 272. [[CrossRef](#)]
73. Lakshmi, M.R.; Prasad, T.; Prakash, D.V.C. Survey on EEG signal processing methods. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2014**, *4*, 84–91.
74. Mane, A.R.; Biradar, S.; Shastri, R. Review paper on feature extraction methods for EEG signal analysis. *Int. J. Emerg. Trend Eng. Basic Sci.* **2015**, *2*, 545–552.
75. Xie, Y.; Oniga, S. A Review of Processing Methods and Classification Algorithm for EEG Signal. *Carpathian J. Electron. Comput. Eng.* **2020**, *12*, 23–29. [[CrossRef](#)]

76. Mirzaei, S.; Ghasemi, P. EEG motor imagery classification using dynamic connectivity patterns and convolutional autoencoder. *Biomed. Signal Process. Control* **2021**, *68*, 102584. [[CrossRef](#)]
77. Rajaguru, H.; Prabhakar, S.K. Sparse PCA and soft decision tree classifiers for epilepsy classification from EEG signals. In Proceedings of the 2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 20–22 April 2017; Volume 1, pp. 581–584.
78. Zhang, Y.; Liu, B.; Ji, X.; Huang, D. Classification of EEG signals based on autoregressive model and wavelet packet decomposition. *Neural Process. Lett.* **2017**, *45*, 365–378. [[CrossRef](#)]
79. Hengstler, S.; Sand, S.; Costa, A.H. Adaptive autoregressive modeling for time-frequency analysis. In Proceedings of the Third International Conference on Information, Communications & Signal Processing (ICICS 2001), Chongqing, China, 19–21 November 2001; pp. 241–244.
80. Shakeel, A.; Tanaka, T.; Kitajo, K. Time-series prediction of the oscillatory phase of EEG signals using the least mean square algorithm-based AR model. *Appl. Sci.* **2020**, *10*, 3616. [[CrossRef](#)]
81. Jahidin, A.H.; Ali, M.M.; Taib, M.N.; Tahir, N.M.; Yassin, I.M.; Lias, S. Classification of intelligence quotient via brainwave sub-band power ratio features and artificial neural network. *Comput. Methods Programs Biomed.* **2014**, *114*, 50–59. [[CrossRef](#)] [[PubMed](#)]
82. Singh, P.; Joshi, S.D.; Patney, R.K.; Saha, K. The Fourier decomposition method for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2017**, *473*, 20160871. [[CrossRef](#)] [[PubMed](#)]
83. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2013**, *62*, 531–544. [[CrossRef](#)]
84. Peng, C.J.; Chen, Y.C.; Chen, C.C.; Chen, S.J.; Cagneau, B.; Chassagne, L. An EEG-based attentiveness recognition system using Hilbert–Huang transform and support vector machine. *J. Med. Biol. Eng.* **2020**, *40*, 230–238. [[CrossRef](#)]
85. Jahankhani, P.; Kodogiannis, V.; Revett, K. EEG signal classification using wavelet feature extraction and neural networks. In Proceedings of the IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06), Sofia, Bulgaria, 3–6 October 2006; pp. 120–124.
86. Wang, Y.; Veluvolu, K.C.; Lee, M. Time-frequency analysis of band-limited EEG with BMFLC and Kalman filter for BCI applications. *J. Neuroeng. Rehabil.* **2013**, *10*, 1–16. [[CrossRef](#)]
87. Rutkowski, G.; Patan, K.; Leśniak, P. Comparison of time-frequency feature extraction methods for EEG signals classification. In Proceedings of the International Conference on Artificial Intelligence and Soft Computing, Zakopane, Poland, 9–13 June 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 320–329.
88. Jin, J.; Xiao, R.; Daly, I.; Miao, Y.; Wang, X.; Cichocki, A. Internal feature selection method of CSP based on L1-norm and dempster-shafer theory. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4814–4825. [[CrossRef](#)] [[PubMed](#)]
89. Fu, R.; Han, M.; Tian, Y.; Shi, P. Improvement motor imagery EEG classification based on sparse common spatial pattern and regularized discriminant analysis. *J. Neurosci. Methods* **2020**, *343*, 108833. [[CrossRef](#)] [[PubMed](#)]
90. Bhalla, A.; Agrawal, R. Relevant feature extraction by combining independent components analysis and common spatial patterns for EEG based motor imagery classification. *Int. J. Eng. Res. Technol.* **2014**, *3*, 246–252.
91. Blankertz, B.; Tomioka, R.; Lemm, S.; Kawanabe, M.; Muller, K.R. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Process. Mag.* **2007**, *25*, 41–56. [[CrossRef](#)]
92. Ang, K.K.; Chin, Z.Y.; Wang, C.; Guan, C.; Zhang, H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Front. Neurosci.* **2012**, *6*, 39. [[CrossRef](#)]
93. Kumar, S.; Sharma, A.; Tsunoda, T. An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information. *BMC Bioinform.* **2017**, *18*, 545. [[CrossRef](#)]
94. Lemm, S.; Blankertz, B.; Curio, G.; Muller, K.R. Spatio-spectral filters for improving the classification of single trial EEG. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1541–1548. [[CrossRef](#)]
95. Cho, H.; Ahn, M.; Ahn, S.; Jun, S.C. Invariant common spatio-spectral patterns. In Proceedings of the 3rd TOBI Workshop, Würzburg, Germany, 20–22 March 2012; pp. 31–32.
96. Dornhege, G.; Blankertz, B.; Krauledat, M.; Losch, F.; Curio, G.; Muller, K.R. Combined optimization of spatial and temporal filters for improving brain-computer interfacing. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2274–2281. [[CrossRef](#)]
97. Novi, Q.; Guan, C.; Dat, T.H.; Xue, P. Sub-band common spatial pattern (SBCSP) for brain-computer interface. In Proceedings of the 2007 3rd International IEEE/EMBS Conference on Neural Engineering, Kohala Coast, HI, USA, 2–5 May 2007; pp. 204–207.
98. Higashi, H.; Tanaka, T. Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification. *IEEE Trans. Biomed. Eng.* **2012**, *60*, 1100–1110. [[CrossRef](#)] [[PubMed](#)]
99. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)] [[PubMed](#)]
100. Wang, Y.; Cang, S.; Yu, H. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst. Appl.* **2019**, *137*, 167–190. [[CrossRef](#)]
101. Bengio, Y.; Courville, A.C.; Vincent, P. Unsupervised feature learning and deep learning: A review and new perspectives. *arXiv* **2012**, arXiv:1206.5538v1.
102. Ficon, G.; Weitschek, E.; Cialini, A.; Felici, G.; Bertolazzi, P.; De Salvo, S.; Bramanti, A.; Bramanti, P.; De Cola, M.C. Combining EEG signal processing with supervised methods for Alzheimer’s patients classification. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 35. [[CrossRef](#)] [[PubMed](#)]



103. Haykin, S.; Network, N. A comprehensive foundation. *Neural Netw.* **2004**, *2*, 41.
104. Behri, M.; Subasi, A.; Qaisar, S.M. Comparison of machine learning methods for two class motor imagery tasks using EEG in brain-computer interface. In Proceedings of the 2018 Advances in Science and Engineering Technology International Conferences (ASET), Abu Dhabi, United Arab Emirates, 6 February–5 April 2018; pp. 1–5.
105. Tan, P.N.; Steinbach, M.; Kumar, V. Classification: Basic concepts, decision trees, and model evaluation. *Introd. Data Min.* **2006**, *1*, 145–205.
106. Hosseini, M.P.; Hosseini, A.; Ahi, K. A Review on machine learning for EEG Signal processing in bioengineering. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 204–218. [[CrossRef](#)]
107. Oktavia, N.Y.; Wibawa, A.D.; Pane, E.S.; Purnomo, M.H. Human emotion classification based on EEG signals using Naïve bayes method. In Proceedings of the 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 21–22 September 2019; pp. 319–324.
108. Lestari, F.P.; Haekal, M.; Edison, R.E.; Fauzy, F.R.; Khotimah, S.N.; Haryanto, F. Epileptic Seizure Detection in EEGs by Using Random Tree Forest, Naïve Bayes and KNN Classification. *J. Phys. Conf. Ser.* **2020**, *1505*, 012055. [[CrossRef](#)]
109. Wang, H.; Zhang, Y. Detection of motor imagery EEG signals employing Naïve Bayes based learning process. *Measurement* **2016**, *86*, 148–158.
110. Palaniappan, R.; Sundaraj, K.; Sundaraj, S. A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals. *BMC Bioinform.* **2014**, *15*, 223. [[CrossRef](#)] [[PubMed](#)]
111. Osowski, S.; Siwek, K.; Markiewicz, T. MLP and SVM networks-a comparative study. In Proceedings of the 6th Nordic Signal Processing Symposium, Espoo, Finland, 9–11 June 2004; pp. 37–40.
112. Beyer, K.; Goldstein, J.; Ramakrishnan, R.; Shaft, U. When is “nearest neighbor” meaningful? In Proceedings of the International Conference on Database Theory, Jerusalem, Israel, 10–12 January 1999; Springer: Berlin/Heidelberg, Germany, 1999; pp. 217–235.
113. Du, S.C.; Huang, D.L.; Wang, H. An adaptive support vector machine-based workpiece surface classification system using high-definition metrology. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 2590–2604. [[CrossRef](#)]
114. Huang, C.L.; Wang, C.J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst. Appl.* **2006**, *31*, 231–240. [[CrossRef](#)]
115. Wang, G.; Qiu, Y.F.; Li, H.X. Temperature forecast based on SVM optimized by PSO algorithm. In Proceedings of the 2010 International Conference on Intelligent Computing and Cognitive Informatics, Kuala Lumpur, Malaysia, 22–23 June 2010; pp. 259–262.
116. Moctezuma, L.A.; Molinas, M. Classification of low-density EEG for epileptic seizures by energy and fractal features based on EMD. *J. Biomed. Res.* **2020**, *34*, 180. [[CrossRef](#)] [[PubMed](#)]
117. Trambaiolli, L.R.; Lorena, A.C.; Fraga, F.J.; Kanda, P.A.; Anghinah, R.; Nitrini, R. Improving Alzheimer’s disease diagnosis with machine learning techniques. *Clin. EEG Neurosci.* **2011**, *42*, 160–165. [[CrossRef](#)]
118. Murugavel, A.M.; Ramakrishnan, S. Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification. *Med. Biol. Eng. Comput.* **2016**, *54*, 149–161. [[CrossRef](#)]
119. Ramzan, M.; Dawn, S. Learning-based classification of valence emotion from electroencephalography. *Int. J. Neurosci.* **2019**, *129*, 1085–1093. [[CrossRef](#)]
120. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; Volume 398.
121. Bandos, T.V.; Bruzzone, L.; Camps-Valls, G. Classification of hyperspectral images with regularized linear discriminant analysis. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 862–873. [[CrossRef](#)]
122. Tharwat, A.; Gaber, T.; Ibrahim, A.; Hassanien, A.E. Linear discriminant analysis: A detailed tutorial. *AI Commun.* **2017**, *30*, 169–190. [[CrossRef](#)]
123. 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*, R1. [[CrossRef](#)]
124. Molla, M.K.I.; Saha, S.K.; Yasmin, S.; Islam, M.R.; Shin, J. Trial Regeneration With Subband Signals for Motor Imagery Classification in BCI Paradigm. *IEEE Access* **2021**, *9*, 7632–7642. [[CrossRef](#)]
125. Bostanov, V. BCI competition 2003-data sets Ib and IIb: Feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1057–1061. [[CrossRef](#)] [[PubMed](#)]
126. Scherer, R.; Muller, G.; Neuper, C.; Graitmann, B.; Pfurtscheller, G. An asynchronously controlled EEG-based virtual keyboard: Improvement of the spelling rate. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 979–984. [[CrossRef](#)]
127. Garcia, G.N.; Ebrahimi, T.; Vesin, J.M. Support vector EEG classification in the Fourier and time-frequency correlation domains. In Proceedings of the First International IEEE EMBS Conference on Neural Engineering, Capri, Italy, 20–22 March 2003; pp. 591–594.
128. Balakrishnama, S.; Ganapathiraju, A.; Picone, J. Linear discriminant analysis for signal processing problems. In Proceedings of the IEEE Southeastcon’99. Technology on the Brink of 2000 (Cat. No. 99CH36300), Lexington, KY, USA, 25–28 March 1999; pp. 78–81.
129. Sakhavi, S.; Guan, C.; Yan, S. Learning temporal information for brain-computer interface using convolutional neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 5619–5629. [[CrossRef](#)] [[PubMed](#)]
130. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]

131. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
132. van de Ven, G.M.; Siegelmann, H.T.; Tolias, A.S. Brain-inspired replay for continual learning with artificial neural networks. *Nat. Commun.* **2020**, *11*, 1–14. [[CrossRef](#)]
133. Borgomaneri, S.; Battaglia, S.; Garofalo, S.; Tortora, F.; Avenanti, A.; di Pellegrino, G. State-dependent TMS over prefrontal cortex disrupts fear-memory reconsolidation and prevents the return of fear. *Curr. Biol.* **2020**, *30*, 3672–3679. [[CrossRef](#)]
134. Parisi, G.I.; Kemker, R.; Part, J.L.; Kanan, C.; Wermter, S. Continual lifelong learning with neural networks: A review. *Neural Netw.* **2019**, *113*, 54–71. [[CrossRef](#)]
135. Valliani, A.A.A.; Ranti, D.; Oermann, E.K. Deep learning and neurology: A systematic review. *Neurol. Ther.* **2019**, *8*, 351–365. [[CrossRef](#)]
136. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [[CrossRef](#)]
137. Al-Saegh, A.; Dawwd, S.A.; Abdul-Jabbar, J.M. Deep learning for motor imagery EEG-based classification: A review. *Biomed. Signal Process. Control* **2021**, *63*, 102172. [[CrossRef](#)]
138. Bhattacharya, S.; Maddikunta, P.K.R.; Pham, Q.V.; Gadekallu, T.R.; Chowdhary, C.L.; Alazab, M.; Piran, M.J. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustain. Cities Soc.* **2021**, *65*, 102589. [[CrossRef](#)] [[PubMed](#)]
139. Fiok, K.; Karwowski, W.; Gutierrez, E.; Saeidi, M.; Aljuaid, A.M.; Davahli, M.R.; Taiar, R.; Marek, T.; Sawyer, B.D. A Study of the Effects of the COVID-19 Pandemic on the Experience of Back Pain Reported on Twitter in the United States: A Natural Language Processing Approach. *Int. J. Environ. Res. Public Health* **2021**, *18*, 4543. [[CrossRef](#)] [[PubMed](#)]
140. Fiok, K.; Karwowski, W.; Gutierrez-Franco, E.; Liciaga, T.; Belmonte, A.; Capobianco, R.; Saeidi, M. Automated Detection of Leadership Qualities Using Textual Data at the Message Level. *IEEE Access* **2021**, *9*, 57141–57148. [[CrossRef](#)]
141. Faust, O.; Hagiwara, Y.; Hong, T.J.; Lih, O.S.; Acharya, U.R. Deep learning for healthcare applications based on physiological signals: A review. *Comput. Methods Programs Biomed.* **2018**, *161*, 1–13. [[CrossRef](#)]
142. Miao, Y.; Gowayyed, M.; Metze, F. EESN: End-to-end speech recognition using deep RNN models and WFST-based decoding. In Proceedings of the 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Scottsdale, AZ, USA, 13–17 December 2015; pp. 167–174.
143. Wiggins, W.F.; Kitamura, F.; Santos, I.; Prevedello, L.M. Natural Language Processing of Radiology Text Reports: Interactive Text Classification. *Radiol. Artif. Intell.* **2021**, *2021*, e210035. [[CrossRef](#)]
144. Tan, J.H.; Hagiwara, Y.; Pang, W.; Lim, I.; Oh, S.L.; Adam, M.; San Tan, R.; Chen, M.; Acharya, U.R. Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. *Comput. Biol. Med.* **2018**, *94*, 19–26. [[CrossRef](#)]
145. Graves, A. Long short-term memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 37–45.
146. Wang, P.; Jiang, A.; Liu, X.; Shang, J.; Zhang, L. LSTM-based EEG classification in motor imagery tasks. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 2086–2095. [[CrossRef](#)] [[PubMed](#)]
147. Roy, S.; Kiral-Kornek, I.; Harrer, S. ChronoNet: A deep recurrent neural network for abnormal EEG identification. In Proceedings of the Conference on Artificial Intelligence in Medicine in Europe, Poznan, Poland, 26–29 June 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 47–56.
148. Krause, B.; Lu, L.; Murray, I.; Renals, S. Multiplicative LSTM for sequence modelling. *arXiv* **2016**, arXiv:1609.07959.
149. Öztoprak, H.; Toyacan, M.; Alp, Y.K.; Arıkan, O.; Doğutepe, E.; Karakaş, S. Machine-based classification of ADHD and nonADHD participants using time/frequency features of event-related neuroelectric activity. *Clin. Neurophysiol.* **2017**, *128*, 2400–2410. [[CrossRef](#)]
150. Papakostas, M.; Tsiakas, K.; Giannakopoulos, T.; Makedon, F. Towards predicting task performance from EEG signals. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 4423–4425.
151. Stewart, A.X.; Nuthmann, A.; Sanguinetti, G. Single-trial classification of EEG in a visual object task using ICA and machine learning. *J. Neurosci. Methods* **2014**, *228*, 1–14. [[CrossRef](#)] [[PubMed](#)]
152. Vo, T.; Gedeon, T.; Tran, D. Modeling the mental differentiation task with EEG. In Proceedings of the International Conference on Neural Information Processing, Doha, Qatar, 12–15 November 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 357–364.
153. Arsalan, A.; Majid, M.; Anwar, S.M. Electroencephalography based machine learning framework for anxiety classification. In Proceedings of the International Conference on Intelligent Technologies and Applications, Bahawalpur, Pakistan, 6–8 November 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 187–197.
154. Saeed, S.M.U.; Anwar, S.M.; Khalid, H.; Majid, M.; Bagci, U. EEG based classification of long-term stress using psychological labeling. *Sensors* **2020**, *20*, 1886. [[CrossRef](#)] [[PubMed](#)]
155. Hosseinifard, B.; Moradi, M.H.; Rostami, R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput. Methods Programs Biomed.* **2013**, *109*, 339–345. [[CrossRef](#)]
156. Anuragi, A.; Sisodia, D.S. Empirical wavelet transform based automated alcoholism detecting using EEG signal features. *Biomed. Signal Process. Control* **2020**, *57*, 101777. [[CrossRef](#)]



157. Li, P.Z.; Li, J.H.; Wang, C.D. A SVM-based EEG signal analysis: An auxiliary therapy for tinnitus. In Proceedings of the International Conference on Brain Inspired Cognitive Systems, Beijing, China, 28–30 November 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 207–219.
158. Kaur, S.; Singh, S.; Arun, P.; Kaur, D.; Bajaj, M. Phase space reconstruction of EEG signals for classification of ADHD and control adults. *Clin. EEG Neurosci.* **2020**, *51*, 102–113. [[CrossRef](#)]
159. Vimala, V.; Ramar, K.; Ettappan, M. An intelligent sleep apnea classification system based on EEG signals. *J. Med. Syst.* **2019**, *43*, 36. [[CrossRef](#)]
160. Stevens Jr, C.E.; Zabelina, D.L. Classifying creativity: Applying machine learning techniques to divergent thinking EEG data. *NeuroImage* **2020**, *219*, 116990. [[CrossRef](#)]
161. Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. Deap: A database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* **2011**, *3*, 18–31. [[CrossRef](#)]
162. Bălan, O.; Moise, G.; Moldoveanu, A.; Leordeanu, M.; Moldoveanu, F. Fear level classification based on emotional dimensions and machine learning techniques. *Sensors* **2019**, *19*, 1738. [[CrossRef](#)] [[PubMed](#)]
163. Bazgir, O.; Mohammadi, Z.; Habibi, S.A.H. Emotion recognition with machine learning using EEG signals. In Proceedings of the 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME), Qom, Iran, 29–30 November 2018; pp. 1–5.
164. Doma, V.; Pirouz, M. A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. *J. Big Data* **2020**, *7*, 1–21. [[CrossRef](#)]
165. Jaswanth, V.; Naren, J. A System for the Study of Emotions with EEG Signals Using Machine Learning and Deep Learning. In *Cognitive Informatics and Soft Computing*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 59–65.
166. Rozgić, V.; Vitaladevuni, S.N.; Prasad, R. Robust EEG emotion classification using segment level decision fusion. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 1286–1290.
167. Shukla, S.; Chaurasiya, R.K. Emotion Analysis Through EEG and Peripheral Physiological Signals Using KNN Classifier. In *International Conference on ISMAC in Computational Vision and Bio-Engineering*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 97–106.
168. Nawaz, R.; Cheah, K.H.; Nisar, H.; Yap, V.V. Comparison of different feature extraction methods for EEG-based emotion recognition. *Biocybern. Biomed. Eng.* **2020**, *40*, 910–926. [[CrossRef](#)]
169. Blankertz, B.; Müller, K.R.; Krusienski, D.J.; Schalk, G.; Wolpaw, J.R.; Schlogl, A.; Pfurtscheller, G.; Millan, J.R.; Schroder, M.; Birbaumer, N. The BCI competition III: Validating alternative approaches to actual BCI problems. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2006**, *14*, 153–159. [[CrossRef](#)]
170. Leeb, R.; Brunner, C.; Müller-Putz, G.; Schlögl, A.; Pfurtscheller, G. *BCI Competition 2008–Graz Data Set B*; Graz University of Technology: Graz, Austria, 2008; pp. 1–6.
171. Soman, S. High performance EEG signal classification using classifiability and the Twin SVM. *Appl. Soft Comput.* **2015**, *30*, 305–318. [[CrossRef](#)]
172. Aljalal, M.; Djemal, R. A Comparative Study of Wavelet and CSP Features Classified Using LDA, SVM and ANN in EEG Based Motor Imagery. In Proceedings of the 2017 9th IEEE-GCC Conference and Exhibition (GCCCE), Manama, Bahrain, 8–11 May 2017; pp. 1–9.
173. 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 Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications, Cairo, Egypt, 28–30 November 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 189–198.
174. Ines, H.; Slim, Y.; Noureddine, E. EEG classification using support vector machine. In Proceedings of the 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13), Hammamet, Tunisia, 18–21 March 2013; pp. 1–4.
175. Jia, H.; Wang, S.; Zheng, D.; Qu, X.; Fan, S. Comparative study of motor imagery classification based on BP-NN and SVM. *J. Eng.* **2019**, *2019*, 8646–8649. [[CrossRef](#)]
176. Molla, M.K.I.; Al Shiam, A.; Islam, M.R.; Tanaka, T. Discriminative feature selection-based motor imagery classification using EEG signal. *IEEE Access* **2020**, *8*, 98255–98265. [[CrossRef](#)]
177. 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.
178. Andrzejak, R.G.; Lehnertz, K.; Mormann, F.; Rieke, C.; David, P.; Elger, C.E. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Phys. Rev. E* **2001**, *64*, 061907. [[CrossRef](#)]
179. Hamad, A.; Hassani, A.E.; Fahmy, A.A.; Houssein, E.H. A hybrid automated detection of epileptic seizures in EEG based on wavelet and machine learning techniques. *arXiv* **2018**, arXiv:1807.10723.
180. Jaiswal, A.K.; Banka, H. Epileptic seizure detection in EEG signal using machine learning techniques. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 81–94. [[CrossRef](#)]
181. Liu, Q.; Zhao, X.; Hou, Z.; Liu, H. Epileptic seizure detection based on the kernel extreme learning machine. *Technol. Health Care* **2017**, *25*, 399–409. [[CrossRef](#)] [[PubMed](#)]

182. Riaz, F.; Hassan, A.; Rehman, S.; Niazi, I.K.; Dremstrup, K. EMD-based temporal and spectral features for the classification of EEG signals using supervised learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *24*, 28–35. [[CrossRef](#)] [[PubMed](#)]
183. Mardini, W.; Yassein, M.M.B.; Al-Rawashdeh, R.; Aljawarneh, S.; Khamayseh, Y.; Meqdadi, O. Enhanced detection of epileptic seizure using EEG signals in combination with machine learning classifiers. *IEEE Access* **2020**, *8*, 24046–24055. [[CrossRef](#)]
184. Savadkoobi, M.; Oladunni, T.; Thompson, L. A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. *Biocybern. Biomed. Eng.* **2020**, *40*, 1328–1341. [[CrossRef](#)]
185. Chakraborty, M.; Mitra, D. Epilepsy seizure detection using kurtosis based VMD's parameters selection and bandwidth features. *Biomed. Signal Process. Control* **2021**, *64*, 102255.
186. Ech-Choudany, Y.; Scida, D.; Assarar, M.; Landré, J.; Bellach, B.; Morain-Nicolier, F. Dissimilarity-based time–frequency distributions as features for epileptic EEG signal classification. *Biomed. Signal Process. Control* **2021**, *64*, 102268. [[CrossRef](#)]
187. Kaya, Y.; Ertuğrul, Ö.F. A stable feature extraction method in classification epileptic EEG signals. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 721–730. [[CrossRef](#)]
188. Moody, G.B.; Mark, R.G.; Goldberger, A.L. PhysioNet: Physiologic signals, time series and related open source software for basic, clinical, and applied research. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 8327–8330.
189. Fergus, P.; Hignett, D.; Hussain, A.; Al-Jumeily, D.; Abdel-Aziz, K. Automatic epileptic seizure detection using scalp EEG and advanced artificial intelligence techniques. *BioMed Res. Int.* **2015**, *2015*, 986736. [[CrossRef](#)]
190. Pinto-Orellana, M.A.; Cerqueira, F.R. Patient-specific epilepsy seizure detection using random forest classification over one-dimension transformed EEG data. In Proceedings of the International Conference on Intelligent Systems Design and Applications, Porto, Portugal, 16–18 December 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 519–528.
191. Shoeb, A.; Kharbouch, A.; Soegaard, J.; Schachter, S.; Gutttag, J. A machine-learning algorithm for detecting seizure termination in scalp EEG. *Epilepsy Behav.* **2011**, *22*, S36–S43. [[CrossRef](#)]
192. Usman, S.M.; Usman, M.; Fong, S. Epileptic seizures prediction using machine learning methods. *Comput. Math. Methods Med.* **2017**, *2017*, 9074759. [[CrossRef](#)]
193. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [[CrossRef](#)]
194. Alomari, M.H.; Samaha, A.; AlKamha, K. Automated classification of L/R hand movement EEG signals using advanced feature extraction and machine learning. *arXiv* **2013**, arXiv:1312.2877.
195. Moreira, J.; Moreira, M.; Pombo, N.; Silva, B.M.; Garcia, N.M. Identification of real and imaginary movements in EEG using machine learning models. In Proceedings of the International Conference on Medical and Biological Engineering, Banja Luka, 16–18 May 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 469–474.
196. Priya, T.H.; Mahalakshmi, P.; Naidu, V.; Srinivas, M. Stress detection from EEG using power ratio. In Proceedings of the 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 24–25 February 2020; pp. 1–6.
197. Ihle, M.; Feldwisch-Drentrup, H.; Teixeira, C.A.; Witon, A.; Schelter, B.; Timmer, J.; Schulze-Bonhage, A. EPILEPSIAE—A European epilepsy database. *Comput. Methods Programs Biomed.* **2012**, *106*, 127–138. [[CrossRef](#)] [[PubMed](#)]
198. Direito, B.; Teixeira, C.A.; Sales, F.; Castelo-Branco, M.; Dourado, A. A realistic seizure prediction study based on multiclass SVM. *Int. J. Neural Syst.* **2017**, *27*, 1750006. [[CrossRef](#)]
199. Teixeira, C.A.; Direito, B.; Bandarabadi, M.; Le Van Quyen, M.; Valderrama, M.; Schelter, B.; Schulze-Bonhage, A.; Navarro, V.; Sales, F.; Dourado, A. Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients. *Comput. Methods Programs Biomed.* **2014**, *114*, 324–336. [[CrossRef](#)] [[PubMed](#)]
200. Bandarabadi, M.; Teixeira, C.A.; Rasekhi, J.; Dourado, A. Epileptic seizure prediction using relative spectral power features. *Clin. Neurophysiol.* **2015**, *126*, 237–248. [[CrossRef](#)] [[PubMed](#)]
201. Kemp, B.; Zwinderman, A.H.; Tuk, B.; Kamphuisen, H.A.; Oberye, J.J. Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG. *IEEE Trans. Biomed. Eng.* **2000**, *47*, 1185–1194. [[CrossRef](#)]
202. Ravan, M. Machine Learning Approach to Measure Sleep Quality using EEG Signals. In Proceedings of the 2019 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 7 December 2019; pp. 1–6.
203. Hassan, A.R.; Bhuiyan, M.I.H. Automatic sleep scoring using statistical features in the EMD domain and ensemble methods. *Biocybern. Biomed. Eng.* **2016**, *36*, 248–255. [[CrossRef](#)]
204. Delimayanti, M.K.; Purnama, B.; Nguyen, N.G.; Faisal, M.R.; Mahmudah, K.R.; Indriani, F.; Kubo, M.; Satou, K. Classification of brainwaves for sleep stages by high-dimensional FFT features from EEG signals. *Appl. Sci.* **2020**, *10*, 1797. [[CrossRef](#)]
205. Ichimaru, Y.; Moody, G. Development of the polysomnographic database on CD-ROM. *Psychiatry Clin. Neurosci.* **1999**, *53*, 175–177. [[CrossRef](#)]
206. Winterhalder, M.; Schelter, B.; Maiwald, T.; Brandt, A.; Schad, A.; Schulze-Bonhage, A.; Timmer, J. Spatio-temporal patient–individual assessment of synchronization changes for epileptic seizure prediction. *Clin. Neurophysiol.* **2006**, *117*, 2399–2413. [[CrossRef](#)]
207. Williamson, J.R.; Bliss, D.W.; Browne, D.W.; Narayanan, J.T. Seizure prediction using EEG spatiotemporal correlation structure. *Epilepsy Behav.* **2012**, *25*, 230–238. [[CrossRef](#)]

208. Wang, K.; Xu, M.; Wang, Y.; Zhang, S.; Chen, L.; Ming, D. Enhance decoding of pre-movement EEG patterns for brain–computer interfaces. *J. Neural Eng.* **2020**, *17*, 016033. [CrossRef] [PubMed]
209. Yuvaraj, R.; Acharya, U.R.; Hagiwara, Y. A novel Parkinson’s Disease Diagnosis Index using higher-order spectra features in EEG signals. *Neural Comput. Appl.* **2018**, *30*, 1225–1235. [CrossRef]
210. 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. Instruments* **2016**, *87*, 085110. [CrossRef] [PubMed]
211. Cheema, B.S.; Samima, S.; Sarma, M.; Samanta, D. Mental workload estimation from EEG signals using machine learning algorithms. In Proceedings of the International Conference on Engineering Psychology and Cognitive Ergonomics, Las Vegas, NV, USA, 15–20 July 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 265–284.
212. Sharma, M.; Acharya, U.R. Automated detection of schizophrenia using optimal wavelet-based  $l_1$  norm features extracted from single-channel EEG. *Cogn. Neurodyn.* **2021**, *15*, 661–674. [CrossRef] [PubMed]
213. Liu, M.; Wu, W.; Gu, Z.; Yu, Z.; Qi, F.; Li, Y. Deep learning based on batch normalization for P300 signal detection. *Neurocomputing* **2018**, *275*, 288–297. [CrossRef]
214. Blankertz, B.; Krusienski, D.; Schalk, G. Documentation Second Wadsworth BCI Dataset (P300 Evoked Potentials) Data Acquired Using BCI2000 P300 Speller Paradigm. BCI Classification Contest November. Available online: [http://www.bbci.de/competition/ii/albany\\_desc/albany\\_desc\\_ii.pdf](http://www.bbci.de/competition/ii/albany_desc/albany_desc_ii.pdf) (accessed on 14 November 2018).
215. Pereira, A.; Padden, D.; Jantz, J.; Lin, K.; Alcaide-Aguirre, R. Cross-Subject EEG Event-Related Potential Classification for Brain-Computer Interfaces Using Residual Networks. Available online: <https://hal.archives-ouvertes.fr/hal-01878227> (accessed on 10 September 2019).
216. Lawhern, V.J.; Solon, A.J.; Waytowich, N.R.; Gordon, S.M.; Hung, C.P.; Lance, B.J. EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *J. Neural Eng.* **2018**, *15*, 056013. [CrossRef]
217. Stober, S.; Cameron, D.J.; Grahn, J.A. Using Convolutional Neural Networks to Recognize Rhythm Stimuli from Electroencephalography Recordings. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 1449–1457.
218. Heraz, A.; Razaki, R.; Frasson, C. Using machine learning to predict learner emotional state from brainwaves. In Proceedings of the Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), Niigata, Japan, 18–20 July 2007; pp. 853–857.
219. Li, M.; Lu, B.L. Emotion classification based on gamma-band EEG. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 1223–1226.
220. Grabowski, K.; Rynkiewicz, A.; Lassalle, A.; Baron-Cohen, S.; Schuller, B.; Cummins, N.; Baird, A.; Podgórska-Bednarz, J.; Pieniżek, A.; Łucka, I. Emotional expression in psychiatric conditions: New technology for clinicians. *Psychiatry Clin. Neurosci.* **2019**, *73*, 50–62. [CrossRef]
221. Borgomaneri, S.; Vitale, F.; Battaglia, S.; Avenanti, A. Early Right Motor Cortex Response to Happy and Fearful Facial Expressions: A TMS Motor-Evoked Potential Study. *Brain Sci.* **2021**, *11*, 1203. [CrossRef]
222. Ellena, G.; Starita, F.; Haggard, P.; Lådavas, E. The spatial logic of fear. *Cognition* **2020**, *203*, 104336. [CrossRef] [PubMed]
223. Candini, M.; Battaglia, S.; Benassi, M.; di Pellegrino, G.; Frassinetti, F. The physiological correlates of interpersonal space. *Sci. Rep.* **2021**, *11*, 1–8.
224. Qiao, R.; Qing, C.; Zhang, T.; Xing, X.; Xu, X. A novel deep-learning based framework for multi-subject emotion recognition. In Proceedings of the 2017 4th International Conference on Information, Cybernetics and Computational Social Systems (ICCS), Dalian, China, 24–26 July 2017; pp. 181–185.
225. Murugappan, M.; Alshuaib, W.; Bourisly, A.K.; Khare, S.K.; Sruthi, S.; Bajaj, V. Tunable Q wavelet transform based emotion classification in Parkinson’s disease using Electroencephalography. *PLoS ONE* **2020**, *15*, e0242014. [CrossRef] [PubMed]
226. Horlings, R.; Dăţcu, D.; Rothkrantz, L.J. Emotion recognition using brain activity. In Proceedings of the 9th International Conference on Computer Systems and Technologies and workshop for PhD Students in Computing, Gabrovo, Bulgaria, 12–13 June 2008.
227. Seo, J.; Laine, T.H.; Sohn, K.A. Machine learning approaches for boredom classification using EEG. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 3831–3846. [CrossRef]
228. Takahashi, K. Remarks on SVM-based emotion recognition from multi-modal bio-potential signals. In Proceedings of the RO-MAN 2004, 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759), Kurashiki, Japan, 22 September 2004; pp. 95–100.
229. Alex, M.; Tariq, U.; Al-Shargie, F.; Mir, H.S.; Al Nashash, H. Discrimination of Genuine and Acted Emotional Expressions Using EEG Signal and Machine Learning. *IEEE Access* **2020**, *8*, 191080–191089. [CrossRef]
230. Khare, S.K.; Bajaj, V.; Sinha, G. Adaptive tunable Q wavelet transform-based emotion identification. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9609–9617. [CrossRef]
231. Xiao, L.; Qiu, J.; Lu, J. A Study of Human Behavior and Mental Workload Based on Neural Network. In Proceedings of the International Conference on Human Aspects of IT for the Aged Population, Toronto, ON, Canada, 17–22 July 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 389–397.
232. Feyen, R.G. Bridging the gap: Exploring interactions between digital human models and cognitive models. In Proceedings of the International Conference on Digital Human Modeling, Beijing, China, 22–27 July 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 382–391.

233. Stone, R.T.; Wei, C.S. Exploring the linkage between facial expression and mental workload for arithmetic tasks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Los Angeles, CA, USA, 2011; Volume 55, pp. 616–619.
234. Garofalo, S.; Battaglia, S.; di Pellegrino, G. Individual differences in working memory capacity and cue-guided behavior in humans. *Sci. Rep.* **2019**, *9*, 1–14.
235. Borgomaneri, S.; Serio, G.; Battaglia, S. Please, don't do it! Ten years of progress of non-invasive brain stimulation in action inhibition. *Cortex* **2020**, *132*, 404–422. [[CrossRef](#)] [[PubMed](#)]
236. Battaglia, S.; Serio, G.; Scarpazza, C.; D'Ausilio, A.; Borgomaneri, S. Frozen in (e) motion: How reactive motor inhibition is influenced by the emotional content of stimuli in healthy and psychiatric populations. *Behav. Res. Ther.* **2021**, *146*, 103963. [[CrossRef](#)]
237. Heine, T.; Lenis, G.; Reichensperger, P.; Beran, T.; Doessel, O.; Deml, B. Electrocardiographic features for the measurement of drivers' mental workload. *Appl. Ergon.* **2017**, *61*, 31–43. [[CrossRef](#)] [[PubMed](#)]
238. So, W.K.; Wong, S.W.; Mak, J.N.; Chan, R.H. An evaluation of mental workload with frontal EEG. *PLoS ONE* **2017**, *12*, e0174949. [[CrossRef](#)]
239. 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 *Proceedings of the International Conference on Robot Intelligence Technology and Applications*, Kuala Lumpur, Malaysia, 16–18 December 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 207–221.
240. Bird, J.J.; Manso, L.J.; Ribeiro, E.P.; Ekárt, 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.
241. Rus, I.D.; Marc, P.; Dinsoreanu, M.; Potolea, R.; Muresan, R.C. Classification of EEG signals in an object recognition task. In *Proceedings of the 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, Romania, 7–9 September 2017; pp. 391–395.
242. Jiao, Z.; Gao, X.; Wang, Y.; Li, J.; Xu, H. Deep convolutional neural networks for mental load classification based on EEG data. *Pattern Recognit.* **2018**, *76*, 582–595. [[CrossRef](#)]
243. Yazdani, A.; Ebrahimi, T.; Hoffmann, U. Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier. In *Proceedings of the 2009 4th International IEEE/EMBS Conference on Neural Engineering*, Antalya, Turkey, 29 April–2 May 2009; pp. 327–330.
244. Vanitha, V.; Krishnan, P. Real Time Stress Detection System Based on EEG Signals. 2017; pp. S271–S275. Available online: <http://www.biomedres.info/biomedical-research/real-time-stress-detectionsystem-based-on-eeg-signals.html> (accessed on 18 April 2019).
245. Guo, L.; Wu, Y.; Zhao, L.; Cao, T.; Yan, W.; Shen, X. Classification of mental task from EEG signals using immune feature weighted support vector machines. *IEEE Trans. Magn.* **2010**, *47*, 866–869. [[CrossRef](#)]
246. Kumar, Y.; Dewal, M.; Anand, R. Features extraction of EEG signals using approximate and sample entropy. In *Proceedings of the 2012 IEEE Students' Conference on Electrical, Electronics and Computer Science*, Bhopal, India, 1–2 March 2012; pp. 1–5.
247. Pandey, V.; Choudhary, D.K.; Verma, V.; Sharma, G.; Singh, R.; Chandra, S. Mental Workload Estimation Using EEG. In *Proceedings of the 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Bangalore, India, 26–27 November 2020; pp. 83–86.
248. Gupta, A.; Agrawal, R. Relevant feature selection from EEG signal for mental task classification. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Kuala Lumpur, Malaysia, 29 May–1 June 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 431–442.
249. Wei, C.S.; Ko, L.W.; Chuang, S.W.; Jung, T.P.; Lin, C.T. EEG-based evaluation system for motion sickness estimation. In *Proceedings of the 2011 5th International IEEE/EMBS Conference on Neural Engineering*, Cancun, Mexico, 27 April–1 May 2011; pp. 100–103.
250. Gupta, A.; Khan, R.U.; Singh, V.K.; Tanveer, M.; Kumar, D.; Chakraborti, A.; Pachori, R.B. A novel approach for classification of mental tasks using multiview ensemble learning (MEL). *Neurocomputing* **2020**, *417*, 558–584. [[CrossRef](#)]
251. Hosni, S.M.; Gadallah, M.E.; Bahgat, S.F.; AbdelWahab, M.S. Classification of EEG signals using different feature extraction techniques for mental-task BCI. In *Proceedings of the 2007 International Conference on Computer Engineering & Systems*, Cairo, Egypt, 27–29 November 2007; pp. 220–226.
252. 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](#)]
253. Ferreira, A.; Almeida, C.; Georgieva, P.; Tomé, A.; Silva, F. Advances in EEG-based biometry. In *Proceedings of the International Conference Image Analysis and Recognition*, Póvoa de Varzin, Portugal, 21–23 June 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 287–295.
254. Tang, J.; Xu, M.; Han, J.; Liu, M.; Dai, T.; Chen, S.; Ming, D. Optimizing SSVEP-based BCI system towards practical high-speed spelling. *Sensors* **2020**, *20*, 4186. [[CrossRef](#)]
255. Xiao, X.; Xu, M.; Jin, J.; Wang, Y.; Jung, T.P.; Ming, D. Discriminative canonical pattern matching for single-trial classification of ERP components. *IEEE Trans. Biomed. Eng.* **2019**, *67*, 2266–2275. [[CrossRef](#)]



256. Ke, Y.; Liu, P.; An, X.; Song, X.; Ming, D. An online SSVEP-BCI system in an optical see-through augmented reality environment. *J. Neural Eng.* **2020**, *17*, 016066. [[CrossRef](#)]
257. Mulder, T. Motor imagery and action observation: Cognitive tools for rehabilitation. *J. Neural Transm.* **2007**, *114*, 1265–1278. [[CrossRef](#)]
258. Gannouni, S.; Belwafi, K.; Aboalsamh, H.; AlSamhan, Z.; Alebdi, B.; Almassad, Y.; Alobaedallah, H. EEG-Based BCI System to Detect Fingers Movements. *Brain Sci.* **2020**, *10*, 965. [[CrossRef](#)]
259. Attallah, O.; Aboharbia, J.; Tamazin, M.; Nasser, A.A. A BCI system based on motor imagery for assisting people with motor deficiencies in the limbs. *Brain Sci.* **2020**, *10*, 864. [[CrossRef](#)]
260. Hung, C.I.; Lee, P.L.; Wu, Y.T.; Chen, L.F.; Yeh, T.C.; Hsieh, J.C. Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers. *Ann. Biomed. Eng.* **2005**, *33*, 1053–1070. [[CrossRef](#)] [[PubMed](#)]
261. Yang, P.; Wang, J.; Zhao, H.; Li, R. Mlp with riemannian covariance for motor imagery based eeg analysis. *IEEE Access* **2020**, *8*, 139974–139982. [[CrossRef](#)]
262. 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](#)]
263. Liu, X.; Lv, L.; Shen, Y.; Xiong, P.; Yang, J.; Liu, J. Multiscale space-time-frequency feature-guided multitask learning CNN for motor imagery EEG classification. *J. Neural Eng.* **2021**, *18*, 026003. [[CrossRef](#)]
264. Dai, M.; Zheng, D.; Na, R.; Wang, S.; Zhang, S. EEG classification of motor imagery using a novel deep learning framework. *Sensors* **2019**, *19*, 551. [[CrossRef](#)] [[PubMed](#)]
265. Lu, D.; Triesch, J. Residual deep convolutional neural network for eeg signal classification in epilepsy. *arXiv* **2019**, arXiv:1903.08100.
266. Ullah, I.; Hussain, M.; Aboalsamh, H. An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Syst. Appl.* **2018**, *107*, 61–71. [[CrossRef](#)]
267. Janghel, R.R.; Verma, A.; Rathore, Y.K. Performance comparison of machine learning techniques for epilepsy classification and detection in EEG signal. In *Data Management, Analytics and Innovation*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 425–438.
268. Birjandtalab, J.; Pouyan, M.B.; Cogan, D.; Nourani, M.; Harvey, J. Automated seizure detection using limited-channel EEG and non-linear dimension reduction. *Comput. Biol. Med.* **2017**, *82*, 49–58. [[CrossRef](#)]
269. Wang, Z.; Na, J.; Zheng, B. An improved knn classifier for epilepsy diagnosis. *IEEE Access* **2020**, *8*, 100022–100030. [[CrossRef](#)]
270. Kaya, Y.; Uyar, M.; Tekin, R.; Yıldırım, S. 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. *Appl. Math. Comput.* **2014**, *243*, 209–219. [[CrossRef](#)]
271. Mormann, F.; Andrzejak, R.G.; Elger, C.E.; Lehnertz, K. Seizure prediction: The long and winding road. *Brain* **2007**, *130*, 314–333. [[CrossRef](#)] [[PubMed](#)]
272. Teixeira, C.; Direito, B.; Feldwisch-Drentrup, H.; Valderrama, M.; Costa, R.; Alvarado-Rojas, C.; Nikolopoulos, S.; Le Van Quyen, M.; Timmer, J.; Schelter, B.; et al. EPILAB: A software package for studies on the prediction of epileptic seizures. *J. Neurosci. Methods* **2011**, *200*, 257–271. [[CrossRef](#)] [[PubMed](#)]
273. Rodríguez-Sotelo, J.L.; Osorio-Forero, A.; Jiménez-Rodríguez, A.; Cuesta-Frau, D.; Cirugeda-Roldán, E.; Peluffo, D. Automatic sleep stages classification using EEG entropy features and unsupervised pattern analysis techniques. *Entropy* **2014**, *16*, 6573–6589. [[CrossRef](#)]
274. Oltu, B.; Akşahin, M.F.; Kibaroglu, S. A novel electroencephalography based approach for Alzheimer’s disease and mild cognitive impairment detection. *Biomed. Signal Process. Control* **2021**, *63*, 102223. [[CrossRef](#)]
275. Zoubek, L.; Charbonnier, S.; Lesecq, S.; Buguet, A.; Chapotot, F. Feature selection for sleep/wake stages classification using data driven methods. *Biomed. Signal Process. Control* **2007**, *2*, 171–179. [[CrossRef](#)]
276. Lajnef, T.; Chaïbi, S.; Ruby, P.; Aguera, P.E.; Eichenlaub, J.B.; Samet, M.; Kachouri, A.; Jerbi, K. Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. *J. Neurosci. Methods* **2015**, *250*, 94–105. [[CrossRef](#)] [[PubMed](#)]
277. Kuo, C.E.; Liang, S.F. Automatic stage scoring of single-channel sleep EEG based on multiscale permutation entropy. In Proceedings of the 2011 IEEE Biomedical Circuits and Systems Conference (BioCAS), San Diego, CA, USA, 10–12 November 2011; pp. 448–451.
278. Yetton, B.D.; Niknazar, M.; Duggan, K.A.; McDevitt, E.A.; Whitehurst, L.N.; Sattari, N.; Mednick, S.C. Automatic detection of rapid eye movements (REMs): A machine learning approach. *J. Neurosci. Methods* **2016**, *259*, 72–82. [[CrossRef](#)]
279. Ebrahimi, F.; Mikaeili, M.; Estrada, E.; Nazeran, H. Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 21–24 August 2008; pp. 1151–1154.
280. Phelps, C.E.; Navratilova, E.; Porreca, F. Cognition in the chronic pain experience: Preclinical insights. *Trends Cogn. Sci.* **2021**, *25*, 365–376. [[CrossRef](#)]
281. Ramachandran, V.S. *Encyclopedia of Human Behavior*; Academic Press: Cambridge, MA, USA, 2012.
282. Battaglia, S.; Garofalo, S.; di Pellegrino, G. Context-dependent extinction of threat memories: Influences of healthy aging. *Sci. Rep.* **2018**, *8*, 1–13. [[CrossRef](#)] [[PubMed](#)]
283. Kulkarni, N.; Bairagi, V. Extracting salient features for EEG-based diagnosis of Alzheimer’s disease using support vector machine classifier. *IETE J. Res.* **2017**, *63*, 11–22. [[CrossRef](#)]



284. Kulkarni, N. EEG Signal Analysis for Mild Alzheimer's Disease Diagnosis by Means of Spectral-and Complexity-Based Features and Machine Learning Techniques. In *Proceedings of the 2nd International Conference on Data Engineering and Communication Technology*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 395–403.
285. Morabito, F.C.; Campolo, M.; Ieracitano, C.; Ebadi, J.M.; Bonanno, L.; Bramanti, A.; Desalvo, S.; Mammone, N.; Bramanti, P. Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings. In *Proceedings of the 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow (RTSI)*, Bologna, Italy, 7–9 September 2016; pp. 1–6.
286. Ficon, G.; Weitschek, E.; Felici, G.; Bertolazzi, P.; De Salvo, S.; Bramanti, P.; De Cola, M.C. Alzheimer's disease patients classification through EEG signals processing. In *Proceedings of the 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, Orlando, FL, USA, 9–12 December 2014; pp. 105–112.
287. Betrouni, N.; Delval, A.; Chaton, L.; Defebvre, L.; Duits, A.; Moonen, A.; Leentjens, A.F.; Dujardin, K. Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results. *Mov. Disord.* **2019**, *34*, 210–217. [[CrossRef](#)]
288. Chua, K.C.; Chandran, V.; Acharya, U.R.; Lim, C.M. Application of higher order statistics/spectra in biomedical signals—A review. *Med. Eng. Phys.* **2010**, *32*, 679–689. [[CrossRef](#)]
289. Zhang, L. EEG signals classification using machine learning for the identification and diagnosis of schizophrenia. In *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 23–27 July 2019; pp. 4521–4524.
290. Jiang, X.; Gao, T. An EEG Emotion Classification System Based on One-Dimension Convolutional Neural Networks and Virtual Reality. In *Proceedings of the International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, Lodz, Poland, 1–3 July 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 194–202.
291. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [[CrossRef](#)]
292. Demir, A.; Koike-Akino, T.; Wang, Y.; Haruna, M.; Erdogmus, D. EEG-GNN: Graph Neural Networks for Classification of Electroencephalogram (EEG) Signals. *arXiv* **2021**, arXiv:2106.09135.
293. Wu, D.; Li, X.; Feng, J. Connectome-based individual prediction of cognitive behaviors via the graph propagation network reveals directed brain network topology. *bioRxiv* **2021**. [[CrossRef](#)]
294. Zhu, S.; Pan, S.; Zhou, C.; Wu, J.; Cao, Y.; Wang, B. Graph geometry interaction learning. *arXiv* **2020**, arXiv:2010.12135.
295. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [[CrossRef](#)]
296. Casas, S.; Gulino, C.; Liao, R.; Urtasun, R. Spagrn: Spatially-aware graph neural networks for relational behavior forecasting from sensor data. In *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 31 May–31 August 2020; pp. 9491–9497.
297. Johannesen, J.K.; Bi, J.; Jiang, R.; Kenney, J.G.; Chen, C.M.A. Machine learning identification of EEG features predicting working memory performance in schizophrenia and healthy adults. *Neuropsychiatr. Electrophysiol.* **2016**, *2*, 1–21. [[CrossRef](#)]
298. Kurkin, S.; Pitsik, E.; Frolov, N. Artificial intelligence systems for classifying eeg responses to imaginary and real movements of operators. In *Saratov Fall Meeting 2018: Computations and Data Analysis: From Nanoscale Tools to Brain Functions*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; Volume 11067, p. 1106709.
299. Islam, S.M.R.; Sajol, A.; Huang, X.; Ou, K.L. Feature extraction and classification of EEG signal for different brain control machine. In *Proceedings of the 2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, Bangladesh, 22–24 September 2016; pp. 1–6.
300. Jian-Feng, H. Comparison of different classifiers for biometric system based on EEG signals. In *Proceedings of the 2010 Second International Conference on Information Technology and Computer Science*, Indianapolis, IN, USA, 30 November–3 December 2010; pp. 288–291.
301. Plewan, T.; Wascher, E.; Falkenstein, M.; Hoffmann, S. Classifying response correctness across different task sets: A machine learning approach. *PLoS ONE* **2016**, *11*, e0152864. [[CrossRef](#)] [[PubMed](#)]
302. Zhao, H.; Zheng, Q.; Ma, K.; Li, H.; Zheng, Y. Deep representation-based domain adaptation for nonstationary EEG classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 535–545. [[CrossRef](#)] [[PubMed](#)]