



Review

# Current Trends, Challenges, and Future Research Directions of Hybrid and Deep Learning Techniques for Motor Imagery Brain–Computer Interface

Emmanouil Lionakis, Konstantinos Karampidis \* and Giorgos Papadourakis

Department of Electrical and Computer Engineering, School of Engineering, Hellenic Mediterranean University, 71004 Heraklion, Crete, Greece; tp3678@edu.hmu.gr (E.L.); papadour@hmu.gr (G.P.)

\* Correspondence: karampidis@hmu.gr

**Abstract:** The field of brain–computer interface (BCI) enables us to establish a pathway between the human brain and computers, with applications in the medical and nonmedical field. Brain computer interfaces can have a significant impact on the way humans interact with machines. In recent years, the surge in computational power has enabled deep learning algorithms to act as a robust avenue for leveraging BCIs. This paper provides an up-to-date review of deep and hybrid deep learning techniques utilized in the field of BCI through motor imagery. It delves into the adoption of deep learning techniques, including convolutional neural networks (CNNs), autoencoders (AEs), and recurrent structures such as long short-term memory (LSTM) networks. Moreover, hybrid approaches, such as combining CNNs with LSTMs or AEs and other techniques, are reviewed for their potential to enhance classification performance. Finally, we address challenges within motor imagery BCIs and highlight further research directions in this emerging field.

**Keywords:** EEG; deep learning; BCI; motor imagery



**Citation:** Lionakis, E.; Karampidis, K.; Papadourakis, G. Current Trends, Challenges, and Future Research Directions of Hybrid and Deep Learning Techniques for Motor Imagery Brain–Computer Interface. *Multimodal Technol. Interact.* **2023**, *7*, 95. <https://doi.org/10.3390/mti7100095>

Academic Editor: Anton Nijholt

Received: 24 August 2023

Revised: 26 September 2023

Accepted: 9 October 2023

Published: 12 October 2023



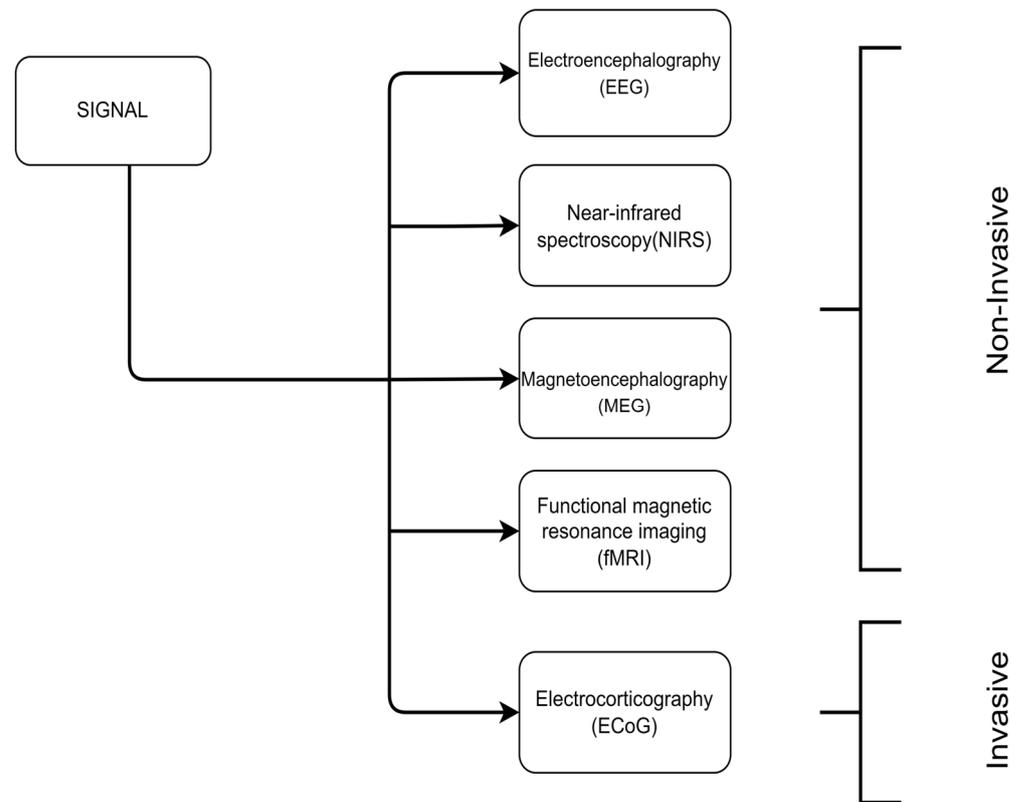
**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Brain–computer interfaces (BCIs) are an emerging field of technology that combines and allows the connection between the brain and a computer or other external devices. BCIs have the potential to revolutionize the way humans interact with machines, opening countless possibilities both in medical and nonmedical domains. In the medical field, it can help people suffering from locked-in syndrome to communicate [1]. Moreover, brain–computer interfaces (BCIs) are displaying potential in the realm of neuroprosthetics, offering the prospect for individuals with limb amputations or paralysis to command robotic limbs or exoskeletons through their brain signals [2]. In epilepsy management, BCIs are researched for real-time seizure detection and intervention, potentially mitigating the impacts of seizures [3]. In the nonmedical domain, an EEG BCI has applications in areas such as gaming where players can play a game using only their thoughts [4] and in fields such as entertainment where users can control drones or other robotic devices [5].

There are several methods and techniques used in BCI research (Figure 1). These methods can be categorized based on their degree of invasiveness as invasive and noninvasive. Invasive methods require physical access to the brain and include ECoG (electrocorticography) [6], which is the process of recording electrical activity in the brain by placing electrodes in direct contact with the cerebral cortex or surface of the brain. Noninvasive techniques do not require any type of surgical operation and can be further discriminated into (a) fMRI (functional magnetic resonance imaging), which measures the brain activity by detecting changes associated with the blood flow [7], (b) MEG (magnetoencephalography) [8], which is the measurement of the magnetic field generated by the electrical activity of neurons, (c) NIRS (near-infrared spectroscopy) [9], a brain imaging method that measures light absorbance to calculate oxyhemoglobin (oxy-HB) and deoxyhemoglobin

(deoxy-HB), which provides an indirect measure of brain activity, particularly in the frontal cortex, and (d) electroencephalography (EEG), [6] which is this paper's subject and refers to the measurement of the electrical activity produced by the brain by placing electrodes to the subject's scalp.

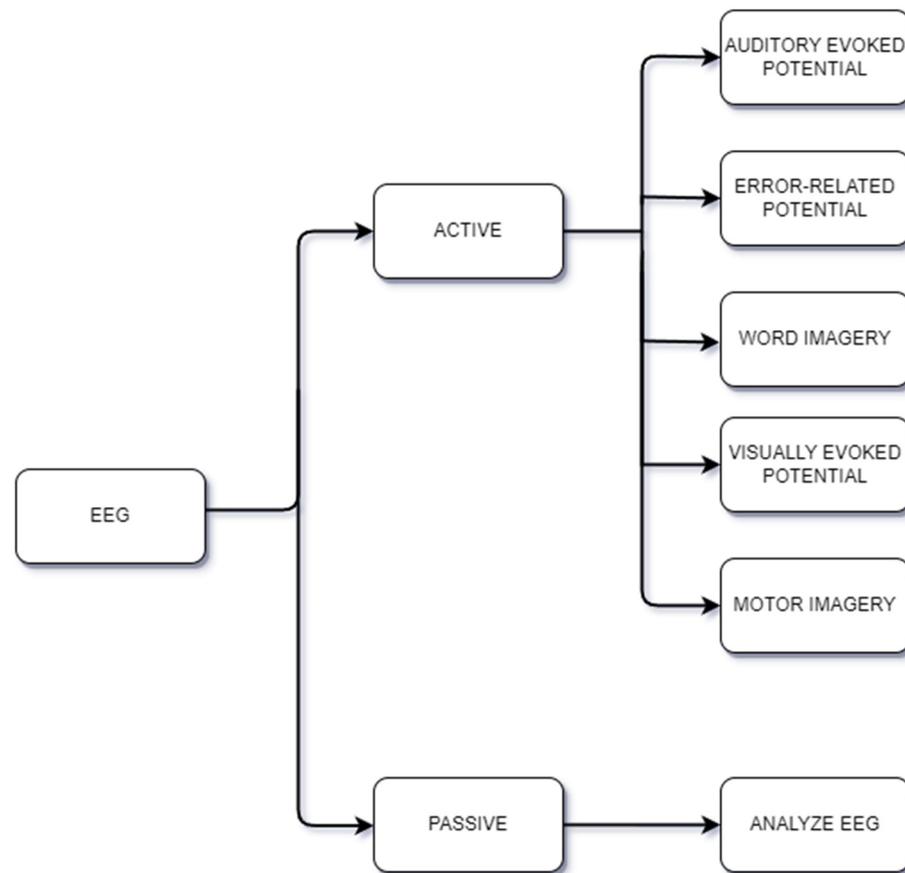


**Figure 1.** Techniques of signal acquisition.

Although there are disadvantages in EEG-based BCIs compared to other techniques [10], such as a low signal-to-noise ratio, a low spatial resolution, and variability in the signal quality due to artifacts, their portability, cost-effectiveness, noninvasiveness and high temporal resolution make them the most spread method for BCI applications.

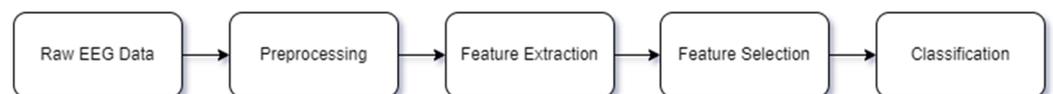
There are a lot of approaches to further categorize EEG BCIs. Authors in [11,12] further categorize BCIs based on the user's activity during the signal acquisition as passive/active. Other approaches are based on the stimuli needed for the activation of the BCI [13], which group noninvasive BCIs into evoked (also referred as exogenous) or spontaneous (also referred as endogenous). Another work [10] categorizes BCIs based on the ability of the system to respond on a specific time-window called synchronous or asynchronous.

In the active/passive approach (Figure 2), the term "Passive" refers to situations where the user is either at rest or involved in a task, and brain activity is being measured. The term "Active" refers to the situation where brain activity is measured while the user is carrying out specific tasks or actions. For example, in motor imagery, the user thinks about performing physical movement without actually performing the movement. On the contrary, in the word imagery paradigm also referred to as imagined speech, the user is required to think about certain words, while visually, auditory, and vibrotactile evoked-potential paradigms involve the measurement of brain activity in response to an external visual, auditory, and vibrotactile stimuli, respectively. Some applications of passive EEG BCIs are measuring the subject's emotion [14], cognitive load [15], and attention [16], while active EEG BCIs have been used to control a cursor on a computer screen [17], moving a robotic arm [18], or even helping patient rehabilitation after a stroke [19].



**Figure 2.** BCI categorization based on user's activity.

A typical pipeline for an EEG motor imagery (MI) BCI application involves several stages (Figure 3). The first stage concerns the signal acquisition, whereas the subject is instructed to imagine a motor movement while their brain activity is measured using EEG. The next stage is signal preprocessing, where several different techniques are used to filter and enhance the quality of the signal. Afterwards, feature extraction is performed, where the processed signal is further analyzed to extract relevant features. These features are then subjected to feature selection, where the most relevant features for the application are identified. Finally, classification is performed using machine learning algorithms to classify the data into different categories, such as left- or right-hand movement, and turn these categories into machine instructions, such as moving a cursor in a 2D space, etc.



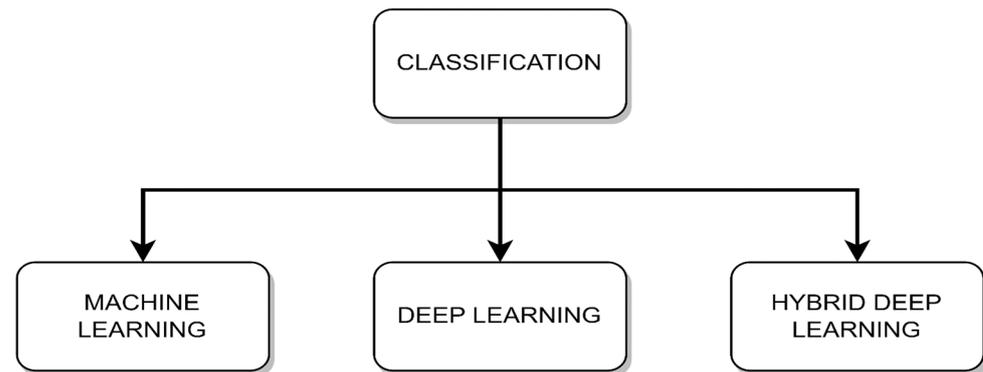
**Figure 3.** Typical EEG BCI pipeline.

The success of this pipeline depends not only on the optimal selection of signal processing techniques, feature extraction methods, feature selection algorithms, and classification models, but also on the user's ability to consistently generate distinguishable brain patterns [20].

Electroencephalogram (EEG) signals, captured noninvasively through electrodes on the scalp, are electrical patterns reflecting the synchronized firing of neurons in the brain. This technique has high temporal resolution, enabling researchers and clinicians to monitor real-time brain activity with millisecond precision. Although EEG signals are considered 1D-signal EEG recordings, they yield time series data, where the x-axis represents time, while the y-axis represents the electrical activity observed at each electrode.

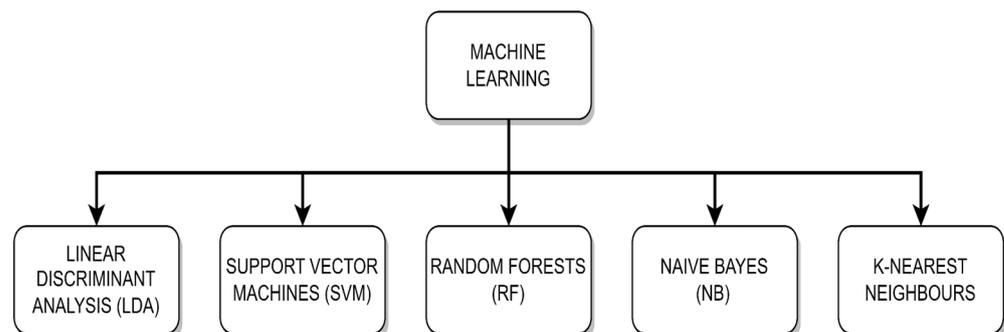
Time series data can undergo different transformations to be reshaped into 2D or 3D formats, with one common approach involving the creation of 2D images by incorporating time–frequency information.

Data classification typically involves one of the following three categories: machine learning, deep learning, and hybrid deep learning as shown in Figure 4. Machine learning provides strong statistical methods for smaller datasets, while deep learning uses neural networks for complex patterns in larger datasets. Hybrid deep learning merges the two previously mentioned methods, extracting benefits from each method to solve problems more effectively.



**Figure 4.** Classification techniques.

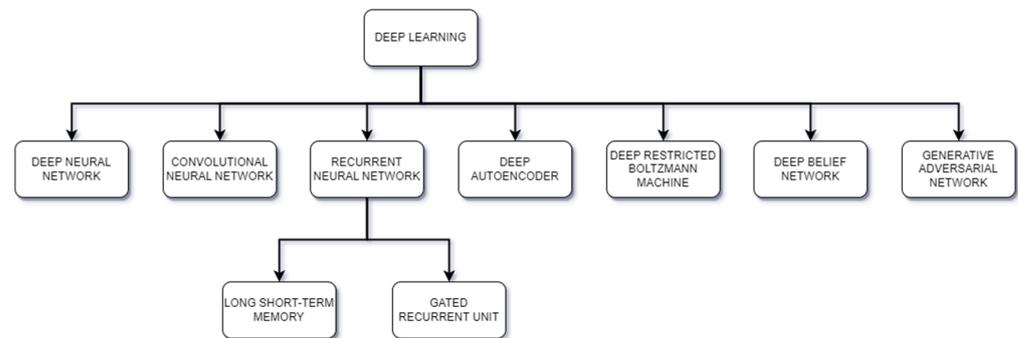
Traditional machine learning is a subfield of artificial intelligence that involves the development of algorithms which can learn from and make predictions or decisions based on data. Unlike deep learning, which requires large datasets and focuses on artificial neural networks, traditional machine learning techniques often work well with smaller datasets and include methods, as shown in Figure 5, such as decision trees, support vector machines, and linear discriminant analysis. These algorithms build models by analyzing input data features and finding patterns or statistical relationships, which can then be used to make predictions on new, unseen data.



**Figure 5.** Machine learning architectures.

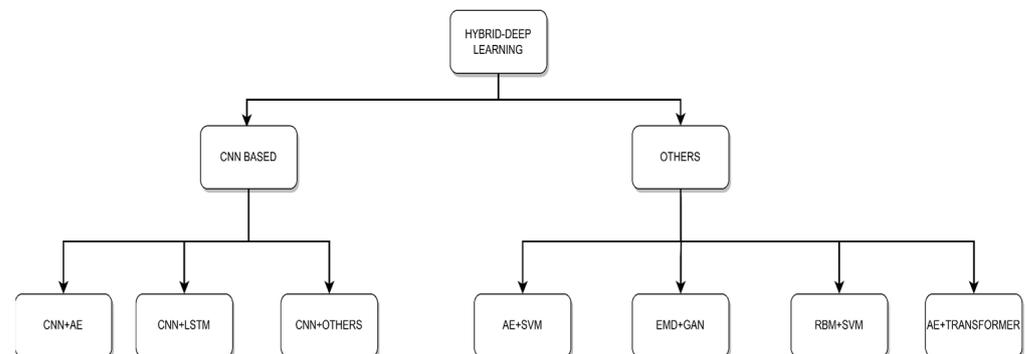
Deep learning has recently become a key component in the development of numerous industries, from healthcare [21] to computer security [22–26]. Deep learning, a subfield of machine learning, excels at managing high-dimensional and complex datasets. Its capacity to represent complex patterns using neural networks with numerous layers has revolutionized a variety of applications, including image recognition [27], natural language processing [28], and the diagnosis of medical conditions [21]. Deep learning has the ability to learn representations directly from the input data, without relying heavily on feature engineering, which is often required in traditional machine learning. Through multiple layers of artificial neural networks, deep learning algorithms can automatically learn hierarchical representations, where initial layers often capture low-level features and

deeper layers capture more complex and abstract features. This capability to automatically extract and learn features from raw data is a major reason why deep learning has been successful in numerous fields. Deep learning can be categorized into two main groups, deep learning and hybrid deep learning; deep learning includes models such as transfer learning, convolutional neural networks, and recurrent neural networks, as shown in Figure 6.



**Figure 6.** Deep learning architectures.

Recent advancements have led to the fusion of deep learning and machine learning techniques, the so-called hybrid deep learning, as shown in Figure 7. Deep learning techniques such as convolutional neural networks typically serve as a feature extractor from raw data, while machine learning methods techniques serves as a classifier. This hybrid approach leverages the strengths of both methods, creating more effective and interpretable models.

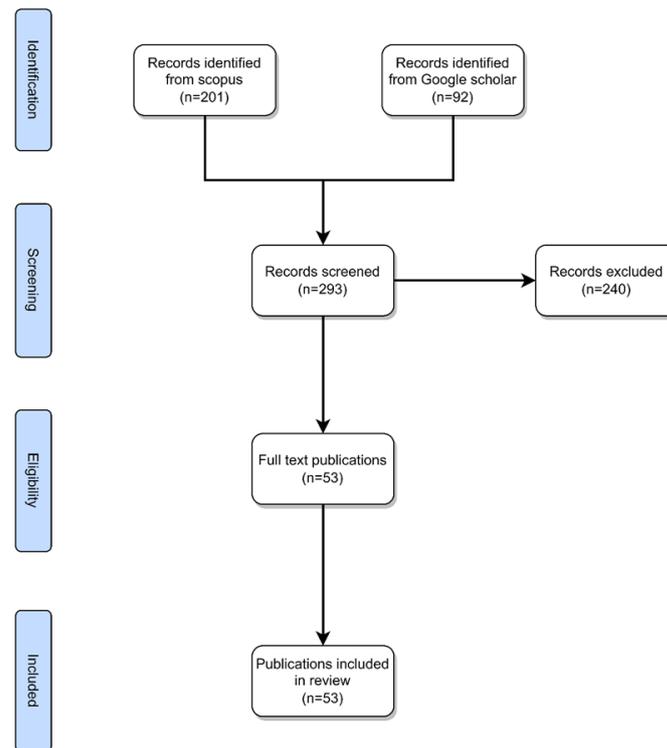


**Figure 7.** Hybrid deep learning architectures.

Several review papers have been published on the subject of EEG BCIs, but they are either outdated [13,29–31] or focus on specific methods [32–38]. For example, the authors in [33–35] provide a significant review but they do not explore the current trends such as deep learning methods. In [36,37], the authors focus only on deep learning methods and in [38], Habashi et al. examine only the proposed methods that involve generative adversarial networks (GANs). Finally, Rajwal et al. [32] focus on the state-of-the-art methods that deploy CNNs.

While some of the aforementioned review papers report state-of-the-art deep learning methods, they overlook the current research trend, i.e., hybrid deep learning techniques. This could be attributed to the fact that both the field itself and the application of deep learning within it are relatively new and rapidly evolving. In this review, our aim is to address this gap. We provide a comprehensive survey of the entire EEG BCI motor imagery system pipeline, focusing on deep learning methods and placing a particular emphasis on these emerging hybrid deep learning methodologies. Furthermore, we present an in-depth evaluation of the most commonly implemented algorithms in the field, assessing their accuracy and computational efficiency. By doing this, we offer an up-to-date and detailed perspective on this exciting and fast-growing area of research.

Scopus and Google Scholar were utilized as the electronic databases to retrieve the articles. This review includes articles related to the following keywords: (1) “EEG” or “Electroencephalogram” or “electroencephalography”, (2) “MI” or “Motor Imagery”, (3) “DL” or “Deep learning”. The Scopus database returned 201 results while Google Scholar returned 92 results adding up to a total of 293 results. From the returned results, we excluded duplicate papers and papers that were not relevant to the review, such as papers that included hybrid methods of signal acquisition, papers that did not utilize deep learning architectures, and papers with similar architectures, methods, and accuracies. After these filters were applied, the number of screened papers was 53 as shown in Figure 8.



**Figure 8.** Protocol for the retrieval of papers.

The rest of this paper is organized as follows, in Section 2, “Datasets”, we present all the publicly available datasets that contains motor imagery tasks and relevant information about the datasets. Section 3, “Deep learning”, includes the deep learning architectures used in the literature for motor imagery brain–computer interfaces. In Section 4, “Hybrid deep learning”, we review the hybrid architectures that contain a combination of machine learning and deep learning architectures to tackle the problem. In Section 5, “Discussion”, we provide our conclusions and insights from the reviewed papers, and we provide directions for further research.

## 2. Datasets

Data play an indispensable role in the classification of EEG signals, as they provide the essential foundation for building and validating algorithms. Utilizing a rich set of EEG data, researchers and engineers can train classifiers to recognize patterns that correspond to different tasks. While private datasets may offer good and relevant data, their accessibility is often restricted to certain institutions or researchers. Public datasets, on the other hand, foster a more inclusive and collaborative approach, making them essential in advancing the field, and they serve as benchmarks in order to make the comparison among the proposed methods more efficient. Table 1 presents the characteristics of available MI public datasets, such as number and type of EEG motor imagery classes, number of electrodes, sampling rate, number of subjects per dataset, number of sessions, and total trials per dataset.

**Table 1.** Publicly available EEG motor imagery datasets.

Datasets	MI CLASSES *	Number of Electrodes	Sampling Rate, Hz	Number of Subjects	Number of Sessions	Number of Trials
BCI Competition II dataset 3b [39]	LH, RH	3	128	3	3	2800
BCI Competition III dataset 4a [40]	RH, RF	118	1000	4	1	1400
BCI Competition IV dataset 2a [41]	LH, RH, BL, T	22	250	9	2	5184
BCI Competition IV dataset 2b [41]	LH, RH	3	250	5	1	6480
EEGMMIDB [42]	BL, LF, RF, BF	64	160	12	1	9100
High Gamma Dataset [43]	LH, RH	128	160	14	13	1000
MIJoint [44]	RH, RE, RS	64	1000	25	7	7500
GigaDB [45]	LH, RH	62	1000	54	2	21,600
MISCP [46]	LH, RH, LL, RL, T	19	200	13	75	4000
Hohyun Cho et al. [47]	LH, RH	64	512	52	1	5500
Brodu N et al. [48]	LH, RH	11	512	1	3	560
Individual imagery [49]	RH, FT	30	256	9	2	1400
Ofner et al. [50]	EF, EE, FS, FP, HO, HC	61	512	15	2	6300

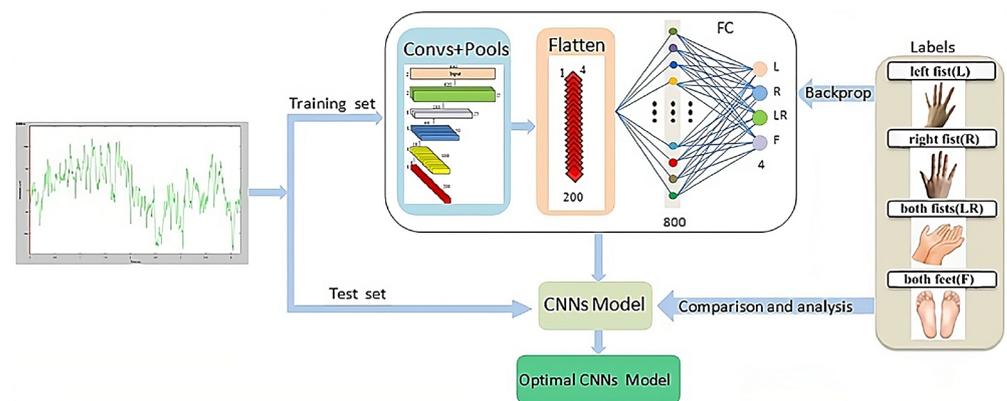
\* LH: left hand, RH: right hand, RL: right leg, LL: left leg, BL: both legs, T: tongue, RE: right elbow, RS: resting state, EF: elbow flexion, EE: elbow extension, FS: forearm supination, FP: forearm pronation, HO: hands open, HC: hands closed, BF: both fists, LF: left fist, RF: right fist.

### 3. Deep Learning

In the following section, we delve into a comprehensive analysis of the most prevalent methods used in the deep learning domain for the interpretation of EEG (electroencephalography) data in the context of brain–computer interface (BCI) with a focus on motor imagery tasks. Our analysis explore a range of approaches from CNNs, which are extensively used due to their proficiency in handling spatial and temporal information, to transfer learning, deep neural networks, and other architectures.

#### 3.1. Convolutional Neural Networks

CNNs mimic the operational principles of the human visual cortex and possess the ability to dynamically comprehend spatial hierarchies in EEG data, recognizing patterns associated with motor imagery tasks through multiple layer transformations [51]. A CNN architecture begins with an input layer that accepts raw or preprocessed EEG data as shown in Figure 9. These data can be represented in various formats, such as time–frequency images, allowing the network to effectively process and analyze the brain signals associated with motor imagery. These data are then convolved using multiple kernels or filters, enabling the network to learn local features. Subsequently, the network employs a pooling layer for dimensionality reduction, refining the comprehension of the information. As the model progresses through these layers, it acquires the capacity to understand increasingly complex features. The final component is a fully connected (dense) classification layer that maps the learned high-level features to the desired output classes, such as different types of motor imagery, effectively acting as a decision-making layer that converts abstract representations into definitive classifications.



**Figure 9.** Typical EEG CNN architecture [52].

Dose et al. proposed a CNN trained on 3 s of segments from EEG signals [53]. The proposed method achieved an accuracy of 80.10%, 69.72%, and 59.71% on two, three, and four MI classes, respectively, on the Physionet dataset.

Miao M et al. proposed a CNN with five layers to classify two motor imagery tasks, right hand and right foot, from the BCI Competition III-IV-a dataset, achieving a 90% accuracy [54].

Zhao et al. proposed a novel CNN with multiple spatial temporal convolution (STC) blocks and fully connected layers [55]. Contrastive learning was used to push the negative samples away and pull the positive samples together. This method achieved an accuracy of 74.10% on BCI III-2a, 73.62% on SMR-BCI, and 69.43% on OpenBMI datasets.

Liu et al. proposed an end-to-end compact multibranch one-dimensional CNN (CMO-CNN) network for decoding MI EEG signals, achieving 83.92% and 87.19% accuracies on the BCI Competition IV-2a and the BCI Competition IV-2b datasets, respectively [56].

Han et al. proposed a parallel CNN (PCNN) to classify motor imagery signals [57]. That method, which achieved an average accuracy of 83.0% on the BCI Competition IV-2b dataset, began by projecting raw EEG signals into a low-dimensional space using a regularized common spatial pattern (RCSP) to enhance class distinctions. Then, the short-time Fourier transform (STFT) collected the mu and beta bands as frequency features, combining them to form 2D images for the PCNN input. The efficacy of the PCNN structure was evaluated against other methods such as stacked autoencoder (SAE), CNN-SAE, and CNN.

Ma et al. proposed an end-to-end, shallow, and lightweight CNN framework, known as Channel-Mixing-ConvNet, aimed at improving the decoding accuracy of the EEG-Motor Raw datasets [58]. Unlike traditional methods, the first block of the network was designed to implicitly stack temporal-spatial convolution layers to learn temporal and spatial EEG features after EEG channels were mixed. This approach integrated the feature extraction capabilities of both layers and enhanced performance. This resulted in a 74.9% accuracy rate on the BCI IV-2a dataset and 95.0% accuracy rate on the High Gamma Dataset (HGD).

Ak et al. performed an EEG data analysis to control a robotic arm. In their work, spectrogram images derived from EEG data were used as input to the GoogLeNet. They tested the system on imagined directional movements—up, down, left, and right—to control the robotic arm [59]. The approach resulted in the robotic arm executing the desired movements with over 90% accuracy, while on their private dataset, they achieved 92.59% accuracy.

Musallam Y et al. proposed the TCNet-Fusion model, which used multiple techniques such as temporal convolutional networks (TCNs), separable convolution, depthwise convolution, and layer fusion [60]. This process created an imagelike representation, which was then fed into the primary TCN. During testing, the model achieved a classification accuracy of 83.73% on the four-class motor imagery of the BCI Competition IV-2a dataset and an accuracy of 94.41% on the High Gamma Dataset.

Zhang et al. proposed a CNN with a 1D convolution on each channel followed by a 2D convolution to extract spatial features based on all 20 channels [61]. Then, to deal with the high computational cost, the idea of pruning was used, which is a technique of reducing the size and complexity of the neural network by removing certain connections or neurons. In the proposed method, a fast recursive algorithm (FRA) was applied to prune redundant parameters in the fully connected layers to reduce computational costs. The proposed architecture achieved an accuracy of 62.7% in the OPENBCI dataset. A similar approach was proposed by Vishnupriya et al. [62] to reduce the complexity of their architecture. The magnitude-based weight pruning was performed on the network, which achieved an accuracy of 84.46% on two MI tasks (left hand, right hand) in Lee et al.'s dataset.

Shajil et al. proposed a CNN architecture to classify four MI tasks, using the common spatial pattern filter on the raw EEG signal, then using the spectrograms extracted from the filtered signals as input into the CNN [63]. The proposed method achieved an accuracy of 86.41% on their private dataset.

Korhan et al. proposed a CNN architecture with five layers [64]. The proposed architecture was compared using only the CNN without any filtering, then with five different filters, and finally, with common spatial patterns followed by the CNN with the last architecture, which achieved the highest accuracy of 93.75% in the BCI Competition III-3a dataset.

Alazrai et al. proposed a CNN network, with the raw signal transformed into the time–frequency domain with the quadratic time–frequency distribution (QTFD), followed by the CNN network to extract and classify the features [65]. The proposed method was tested on their two private datasets, with 11 MI tasks (rest, grasp-related tasks, wrist-related tasks, and finger-related tasks) and obtained accuracies of 73.7% for the able-bodied and 72.8% for the transradial-amputated subjects.

Table 2 summarizes the research articles that utilize CNNs along with the tasks, the datasets used, and their performance.

**Table 2.** Reviewed CNN architectures, datasets and their accuracies.

Authors	Accuracy	Dataset	MI Tasks
Dose et al. [53]	59.71%	EEGMMIDB	LH, RH, RS, BL
Miao M et al. [54]	90%	BCI III-4, private	RH, RF
Zhao R et al. [55]	74.10%, 73.62%, 69.43%	BCI III-2a, SMR-BCI, OpenBMI	RH, RF
Liu X et al. [56]	83.92%, 87.19%	BCI IV-2a, BCI IV-2b	LH, RH, BL, T
Han et al. [57]	83%	BCI IV-2b	LH, RH
Ma et al. [58]	74.9%, 95.0%	BCI IV-2a, HGD	LH, RH, BL, T
Ak et al. [59]	92.59%	Private	U, D, L, R
Musallam et al. [60]	83.73%, 94.41%	BCI IV-2a, HGD	LH, RH, BL, T
Zhang et al. [61]	62.7%	OpenBMI	LH, RH
Vishnupriya et al. [62]	84.46%	Lee et al.	LH, RH
Shajil et al. [63]	86.41%	Private	LH, RH, BH, BL
Korhan et al. [64]	93.75%	BCI III-3a	LH, RH, BL, T
Alazrai et al. [65]	73.7%, 72.8%	Private	RS, SDG, LG, ETG, RDW, EW, FI, FM, FR, FL, FT

LH: left hand, RH: right hand, RL: right leg, BL: both legs, T: tongue, RS: resting state, BF: Both fists, LF: left fist, RF: right fist, U: up, D: down, L: left, R: right, SDG: small-diameter grasp, LG: lateral grasp, ETG: extension-type grasp, RDW: ulnar and radial deviation of the wrist, EW: extension of the wrist, FI: flexion and extension of the index finger, FM: flexion and extension of the middle finger, FR: flexion and extension of the ring finger, FL: flexion and extension of the little finger.

### 3.2. Transfer Learning

In the realm of machine learning, transfer learning offers a valuable approach to improve model performance and efficiency. It involves leveraging the knowledge learned from one task and applying it to a different but related task. The foundation of transfer learning lies in pretrained models, which have been trained on large-scale datasets. This approach is beneficial when considering the computational resources required for training deep learning models from scratch, i.e., time, complexity, and hardware. Transfer learning also offers a noble solution when available training data are not enough to train effectively novel deep learning models. In these situations, where collecting training data is hard or expensive, as in the case of EEG data, a need to develop robust models using available data from diverse domains arises. However, the effectiveness of transfer learning relies on the learned features from the initial task being general and applicable to the target task. By transferring this learned knowledge to a new model and fine-tuning it on a smaller, domain-specific dataset, we can effectively tackle new problems with limited labeled data.

Some popular deep learning models utilized for transfer learning are AlexNet [66], ResNet18 [67], ResNet50, InceptionV3 [68], and ShuffleNet [69]. These models are CNN networks trained on millions of images to classify different classes. For example, AlexNet consists of eight layers with weights, the first five are convolutional layers followed by max-pooling layers, and the last three layers are fully connected layers followed by a softmax layer to provide a probability distribution over the 1000 class labels. Figure 10 shows the architecture of the AlexNet architecture.

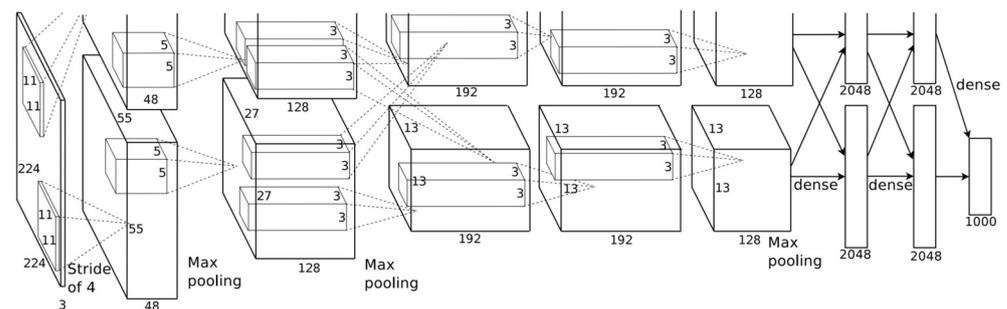
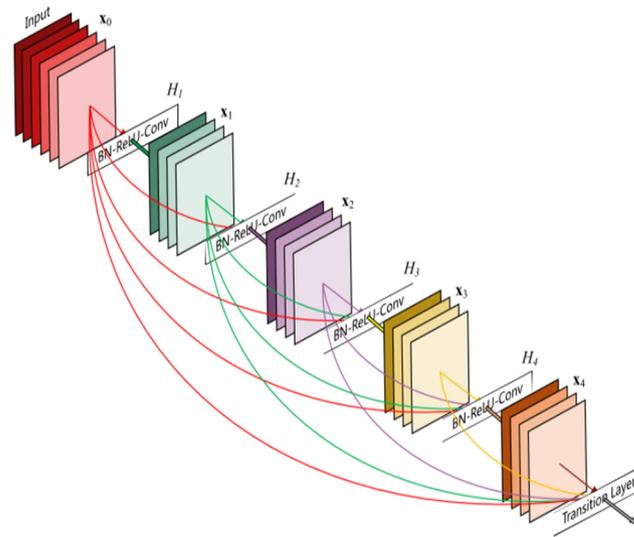


Figure 10. AlexNet architecture [66].

ResNet (residual network) is a deep learning model in which the weight layers learn residual functions with reference to the layer inputs. This method addresses the problem of vanishing gradients in deep neural networks by introducing skip connections, also known as residual blocks. The information can flow directly across multiple layers, making it easier for the network to learn complex features. There are various versions of ResNet utilized for EEG classification, e.g., ResNet34, ResNet50, etc. Figure 11 shows the architecture of ResNet.

While some research papers rely on pretrained models for transfer learning, others take a different approach. Let  $N$  be the total number of subjects in a dataset. Researchers train a custom CNN on  $N-1$  subjects, and afterwards, they use the trained CNN as a base model for transfer learning. That is, they use the remaining  $N$  subject to train the aforementioned CNN, and afterwards, they finetune the whole model. Moreover, some research papers ([70,71]) opt for alternative architectures, such as the one proposed by Schirrneister et al. [43], to facilitate transfer learning in their studies.

Zhang et al. utilized transfer learning to train a hybrid deep neural network (HDNN-TL) which consisted of a convolutional neural network and a long short-term memory model, to decode the spatial and temporal features of the MI signal simultaneously [72]. The classification performance on the BCI Competition IV-2a dataset by the proposed HDNN-TL in terms of kappa value was 0.8 (outperforming the rest of the examined methods).



**Figure 11.** Architecture of the ResNet [67].

Wei et al. [70] proposed a multibranch deep transfer network, the Separate-Common-Separate Network (SCSN) based on splitting the network's feature extractors for individual subjects, and they also explored the possibility of applying maximum mean discrepancy (MMD) to the SCSN (SCSN-MMD). They tested their models on the BCI Competition IV-2a dataset and their own online recorded dataset, which consisted of five subjects (male) and four motor imageries (relaxing, left hand, right hand, and both feet). The results showed that the proposed SCSN achieved an accuracy of (81.8%, 53.2%) and SCSN-MMD achieved an accuracy of (81.8%, 54.8%).

Limpiti et al. used a continuous wavelet transform (CWT) to construct the scalograms from the raw signal, which served as input to five pretrained networks (AlexNet, ResNet18, ResNet50, InceptionV3, and ShuffleNet) [73]. The models were evaluated on the BCI Competition IV-2a dataset. On binary (left hand vs. right hand) and four-class (left hand, right hand, both feet, and tongue) classification, the ResNet18 network achieved the best accuracies at 95.03% and 91.86%.

Wei et al. [74] utilized a CWT to convert the one-dimensional EEG signal into a two-dimensional time–frequency amplitude representation as the input of a pre-trained AlexNet and fine-tuned it to classify two types of MI signals (left hand and right hand). The proposed method achieved a 93.43% accuracy on the BCI Competition II-3 dataset.

Arunabha proposed a multiscale feature-fused CNN (MSFFCNN) efficient transfer learning (TL) and four different variations of the model including subject-specific, subject-independent, and subject-adaptive classification models to exploit the full learning capacity of the classifier [75]. The proposed method achieved a 94.06% accuracy on for four different MI classes (i.e., left hand, right hand, feet, and tongue) on the BCI Competition IV-2a dataset.

Chen et al. proposed a subject-weighted adaptive transfer learning method in conjunction with MLP and CNN classifiers, achieving an accuracy of 96% on their own recorded private dataset [76].

Zhang et al. proposed five schemes for the adaptation of a CNN to two-class motor imagery (left hand, right hand), and after fine-tuning their architecture, they achieved an accuracy of 84.19% on the public GigaDB dataset [71].

Solorzano et al. proposed a method based on transfer learning in neural networks to classify the signals of multiple persons at a time [77]. The resulting neural network classifier achieved a classification accuracy of 73% on the evaluation sessions of four subjects at a time and 74% on three at a time on the BCI Competition IV-2a dataset.

Li et al. proposed a cross-channel specific–mutual feature transfer learning (CCSM-FT) network model with training tricks used to maximize the distinction between the two

kinds of features [78]. The proposed method achieved an 80.26% accuracy on the BCI Competition IV-2a dataset.

A summary of the aforementioned methods can be found in Table 3.

**Table 3.** Reviewed transfer learning architectures and their accuracies.

Authors	Accuracy	Dataset	MI Tasks
Zang et al. [72]	0.8 (kappa)	BCI IV-2a	LH, RH, BL, T
Wei et al. [70]	81.8%, 54.8%	BCI IV-2a, private	LH, RH
Limpiti et al. [73]	95.03%, 91.86%	BCI IV-2a	LH, RH, BL, T
Wei M et al. [74]	93.43%	BCI II-3	LH, RH
Arunabha M. [75]	94.06%	BCI IV-2a	LH, RH, BL, T
Chen et al. [76]	96%	Private	F, B, L, R, S
Zhang et al. [71]	84.19%	GigaDB	LH, RH
Solorzano et al. [77]	74%	BCI IV-2a	LH, RH, BL, T
Li D et al. [78]	80.26%	BCI IV-2a	LH, RH, BL, T

LH: left hand, RH: right hand, RL: right leg, BL: both legs, T: tongue, RS: resting state, BF: Both fists, LF: left fist, RF: right fist, F: forward, B: backward, L: left, R: right, R: rest, S: stop.

### 3.3. Deep Neural Networks

Deep neural networks, a subset of artificial neural networks, have the ability to tackle complex problems. Unlike shallow neural networks that consist of only a few layers, deep neural networks are characterized by their depth, featuring multiple hidden layers between the input and output layers. Each hidden layer progressively extracts higher-level features from the data, allowing the network to learn complex representations and patterns from vast quantities of data.

Suhaimi et al. [79] proposed a deep neural network with four layers each including 50, 30, 15, and 1 node, respectively, achieving a 49.5% classification accuracy in the BCI Competition IV-2b with two MI tasks selected (arm and foot movement).

Cheng et al. proposed a deep neural network which accepted as input multiple sub-bands of the raw signal extracted by a sliding window strategy [80]. Under these sub-bands, diverse spatial-spectral features were extracted and fed into a deep neural network for classification, achieving an accuracy of 71.5% on their private dataset.

Yohanandan et al. proposed a binary classifier (relaxed and right-handed MI tasks) using a deep neural network with the  $\mu$ -rhythm (8–12 Hz frequency) data being fed into the network [81]. The authors used different sliding windows from 1 s to 9 s to determine the highest-accuracy window. An average accuracy of ~83% was achieved on their privately collected dataset from seven human volunteers.

Kumar et al. proposed a deep neural network for the classification of extracted features using a common spatial pattern in the BCI Competition III-4a dataset, achieving an accuracy of ~85% on two MI tasks (right hand and left foot) [82].

Table 4 shows the performance of each one of the aforementioned architectures.

**Table 4.** Reviewed deep neural network architectures and their accuracies.

Authors	Accuracy	Dataset	MI Tasks
Suhaimi et al. [79]	49.5%	BCI Competition 2b	LH, RH
Cheng et al. [80]	71.5%	Private	LH, RH
Yohanandan et al. [81]	83%	Private	RS, RH
Kumar et al. [82]	~85%	BCI Competition III-4a	RH, LF

LH: left hand, RH: right hand, LF: left foot, RS: resting state.

### 3.4. Others

Several alternative methods have been proposed for classifying motor imagery (MI) tasks, aiming at leveraging the potential of different deep learning techniques. Autoencoders [83], which are designed for data reconstruction, have been explored in the context of MI task classification. Autoencoders are neural network architectures that consist of two main phases: encoding and decoding. During the encoding phase, the input signal is passed through a neural network with a progressively reduced number of neurons in each layer until it reaches the bottleneck layer, which has a lower dimensionality compared to the input data. In the decoding phase, the network strives to reconstruct the original signal from this lower-dimensional representation, preserving essential information. This encoding stage in autoencoders enables them to effectively learn compressed representations of input data, such as EEG data, by reducing its dimensionality while retaining significant information. Figure 12 shows an autoencoder used to reconstruct an EEG signal.

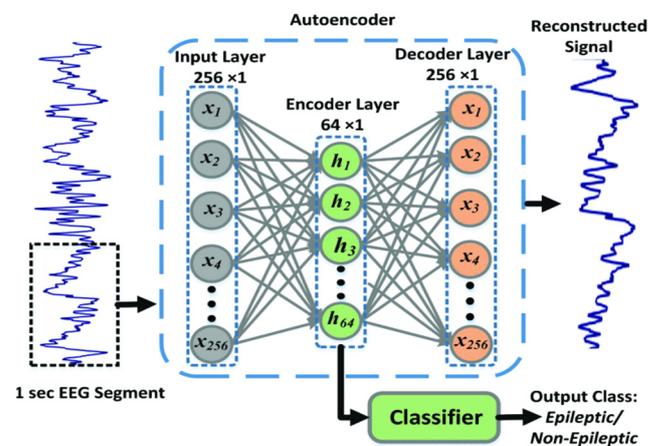


Figure 12. Autoencoder architecture [84].

Autthasan et al. proposed an end-to-end multitask autoencoder and tested it on three datasets, BCI Competition IV-2a, SMR-BCI, and OpenBMI, achieving accuracies of 70.09%, 72.95%, and 66.51%, respectively [85].

Similarly, capsule networks, which introduce a hierarchical structure to capture pose and viewpoint information, have shown promising results in MI task classification [86]. Capsules in capsule networks utilize vector-based representations. This property enables the network to capture hierarchical relationships and spatial dependencies among features. Each capsule comprises a group of neurons, with each neuron's output representing a different property of the same feature, enabling the recognition of the whole entity by first identifying its parts. Ha et al. proposed a capsule network, using the images extracted with the short-time Fourier transform as input to the capsule network [87]. Their proposed method achieved a 77% accuracy on the BCI competition IV-2b dataset (left-hand and right-hand MI tasks).

Long short-term memory (LSTM) networks [88], a type of recurrent neural network, have been utilized to model temporal dependencies in MI data, enabling effective sequence learning for classification. Leon-Urbano et al. proposed an LSTM approach on an MNE python library dataset which consisted of two MI tasks (feet, hands), and after fine-tuning their model, they achieved a 90% accuracy [89]. Saputra et al. also deployed an LSTM network on the BCI Competition IV-2a dataset, achieving an accuracy of 49.65% [90]. Hwang et al. also performed a classification based on an LSTM on the BCI competition IV-2a dataset with a feature extraction based on overlapping band-based FBCSP (filter-bank common spatial pattern), with an accuracy of 97% [91].

Ma et al. proposed a parallel architecture including a temporal LSTM and a spatial bidirectional LSTM [92]. The proposed method was tested on the four MI tasks (moving

both feet, both fists, left fist and right fist) from the EEGMMIDB dataset and achieved an accuracy of 68.20%.

Another proposed method is the restricted Boltzmann machine [93], a type of probabilistic graphical model, leveraging its ability to model joint probability distributions. Xu et al. utilized a restricted Boltzmann machine and a support vector machine (SVM) to classify and recognize deep multiview features [94]. The proposed method achieved an accuracy of 78.50% on the BCI competition IV-2a dataset.

Moreover, metalearning [95] empowers models to acquire the skill of learning on their own, with a limited quantity of data. This is achieved through training the model on a diverse range of tasks, allowing it to leverage the knowledge gained from these tasks when presented with new challenges. Among the various metalearning algorithms, one of the most prominent ones is MAML (model-agnostic metalearning) [95]. MAML trains the model to efficiently update its parameters, facilitating a rapid adaptation to new tasks with minimal updates. Li et al. proposed a metalearning method which learned from the output of other machine learning algorithms [96]. The proposed method achieved an 80% accuracy on the Physionet dataset (on left fist vs. right fist and both fists vs. both feet).

Contrastive learning [97] is a self-supervised learning technique that aims to create meaningful representations by contrasting positive and negative pairs of data. Han et al. proposed the so-called contrastive learning network. The proposed method was tested on the BCI competition IV-2a dataset achieving an accuracy of 79.54% when all the training labels were used [98].

A deep belief network (DBN) [99] is an unsupervised neural network known for its feature extraction from raw data. It uses a two-step training process: unsupervised pretraining with a restricted Boltzmann machine and supervised fine-tuning. Li et al. proposed a deep belief architecture where the time–frequency information from the raw EEG signal was fed into the DBN, which was used for the identification and classification [100]. The proposed method achieved an accuracy of 93.57% on the BCI competition II-3 dataset.

A synopsis of the above-mentioned proposals can be found in Table 5.

**Table 5.** Other reviewed deep learning architectures and their accuracies.

Authors	Accuracy	Dataset	MI Tasks	Architecture
Autthasan et al. [85]	70.09%, 72.95% 66.51%	BCI IV-2a, SMR_BCI, Open BCI	LH, RH, BL, T	Autoencoder
Ha et al. [87]	77%	BCI IV-2b	LH, RH	Capsule network
Urbano et al. [89]	90%	MNE dataset	BF, BH	LSTM
Saputra et al. [90]	49.65%	BCI IV-2a	LH, RH, BL, T	LSTM
Hwang et al. [91]	97%	BCI IV-2a	LH, RH, BL, T	LSTM
Ma et al. [92]	68.20%	EEGMMIDB	LF, RF, BL, BF	LSTM, bi-LSTM
Xu et al. [94]	78.50%	BCI IV-2a	LH, RH, BL, T	Boltzmann Machine
Li et al. [96]	80%	EEGMMIDB	LF, RF, BL, BF	Meta-learning
Han et al. [98]	79.54%	BCI IV-2a	LH, RH, BL, T	Contrastive learning
Li et al. [100]	93.57%	BCI II-3	LH, RH	Deep belief network

LH: left hand, RH: right hand, RL: right leg, BL: both legs, T: tongue, RS: resting state, BF: both fists, LF: left fist, RF: right fist.

#### 4. Hybrid Methods

Hybrid neural networks are a powerful fusion of different types of artificial neural networks, combining the strengths of various architectures to address complex problems

effectively. By integrating components from different neural network types, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and feedforward neural networks (FNNs), hybrid networks can leverage their specialized capabilities to tackle complex problems.

#### 4.1. CNN-Based

##### 4.1.1. CNN and LSTM

Amin et al. proposed a novel deep learning-based lightweight model based on an attention-inception convolutional neural network and long short-term memory [101]. The model was tested on the BCI Competition IV-2a dataset and the High Gamma Dataset, achieving accuracies of 82.8% and 97.1%, respectively.

Khademi et al. proposed a CNN-LSTM hybrid network to extract spatial and temporal sequence features simultaneously and compared it with the same architecture but with pretrained CNN models, ResNet-50 and Inception-v3 [102]. A CWT was used to generate a 2D representation in the time-frequency domain as input for the CNN. The highest accuracy was 86% and achieved with the Inception-v3 hybrid network on the BCI Competition IV-2a dataset.

Echtioui et al. compared a CNN model and a CNN-LSTM model with similar architectures, showing that the addition of the LSTM layer in the model performed worse in terms of accuracy [103]. The CNN-LSTM model achieved a 55.55% accuracy on the BCI Competition IV-2a dataset as opposed to 62.45% for their CNN method.

Li et al. proposed a hybrid network combining a CNN, a DNN, and an LSTM, each as an independent classifier [104]. First, the HNN made a separate prediction for each of the three networks. These predictions were then normalized so that they added up to one. This was done to ensure that the predictions from the three networks could be compared with each other. The maximum prediction from each network was then selected. This gave three values, one for each network. The largest of these three values was then selected as the final prediction. The proposed method achieved an accuracy of 72.22% on the BCI competition IV-2a dataset.

Li et al. [105] proposed a hybrid neural network combining a deep-separation CNN (DSCNN), which is a convolution network divided into two or more convolutions to produce the same output, and a bidirectional LSTM (BLSTM). The proposed model achieved a 98.09% accuracy on the EEGMMIDB dataset, which contains five tasks (eyes closed, open/close left and right fists and open/close both fists, and open/close both feet).

Fadel et al. proposed a hybrid neural network consisting of a deep CNN followed by an LSTM, layer using as input to the model delta [0.5–4 Hz], mu [8–13 Hz], and beta [13–30 Hz] frequency bands [106]. The model was tested on the EEGMMIDB dataset, which includes five different classes (four motor imagery tasks and one resting task), achieved an accuracy of 70.64%. Table 6 shows in brief the hybrid CNN-LSTM architectures.

**Table 6.** Reviewed hybrid CNN-LSTM architectures and their accuracies.

Authors	Accuracy	Dataset	MI Tasks
Amin et al. [101]	82.8%, 97.1%	BCI IV-2a, HGD	LH, RH, BL, T
Khademi et al. [102]	86%	BCI IV-2a	LH, RH, BL, T
Echtioui et al. [103]	55.55%	BCI IV-2a	LH, RH, BL, T
Li et al. [104]	72.22%	BCI IV-2a	LH, RH, BL, T
Li et al. [105]	98.09%	EEGMMIDB	RC, LF, RF, BF, BL
Fadel et al. [106]	70.64%	EEGMMIDB	RC, LF, RF, BF, BL

LH: left hand, RH: right hand, RL: right leg, RS: resting state, BL: Both legs.

#### 4.1.2. CNN and Autoencoders

Tabar et al. proposed a model with a CNN acting as a feature extractor and a stacked autoencoder with six hidden layers acting as a classifier [107]. The proposed architecture achieved an accuracy of 77.6% in the BCI Competition IV-2b dataset.

Dai et al. proposed a model with a CNN acting as a feature extractor but a variation autoencoder which is a directed model that uses learned approximate inference and can be trained purely with gradient-based methods, with seven layers acting as a classifier [108]. The proposed method was tested on the BCI competition IV dataset 2b achieving a kappa value of 0.564.

Hwaidi et al. proposed a network with a variational autoencoder to denoise the signal, followed by a CNN model to extract features and classify the processed signal [109]. The model was tested on the MI tasks of the Physionet dataset MI tasks, left fist, right fist, both fists, and both feet, achieving an accuracy of 98.20%.

A summary of these methods is given on Table 7.

**Table 7.** Reviewed hybrid CNN + AE architectures and their accuracies.

Authors	Accuracy	Dataset	MI Tasks
Tabar et al. [107]	77.6%	BCI IV-2b	LH, RH, BL, T
Dai et al. [108]	0.564 (kappa)	BCI IV-2b	LH, RH
Hwaidi et al. [109]	98.20%	EEGMMIDB	RC, LF, RF, BF, BL

LH: left hand, RH: right hand, RL: right leg, RF: right fist, LF: left fist, RS: resting state, BL: both legs, T: tongue, BF: both fists.

#### 4.1.3. Other CNN Architectures

Gomes et al. proposed a hybrid pretrained CNN (VGG16 and LeNet) and random forests with 100 trees to classify right- and left-hand MI tasks on the BCI Competition IV-2b dataset [110]. Pseudosinogram images were fed to the pretrained model to extract features and then to a random forest classifier. The VGG16 pretrained model slightly outperformed the LeNet with an accuracy of 89.13 and 89.10, respectively, after performing data augmentation.

Ma et al. proposed a CNN that extracted relevant spatial information, which was followed by a transformer layer which captured the long-range dependencies and temporal relationships between different EEG signal segments [111]. The proposed method achieved an accuracy of 83.91% on the BCI Competition IV-2a dataset.

Gao et al. [112] proposed a gated recurrent unit (GRU) and convolutional neural network (CNN), receiving inputs simultaneously, with the CNN extracting frequency and spatial features and the GRU extracting temporal features from the signal, combining their output as an input to the classifier. The proposed method was tested on the BCI competition IV-2a dataset and achieved an accuracy of 80.7%.

Ye et al. proposed a hybrid model with a CNN architecture with three layers with a “ReLU” activation, followed by a gated recurrent unit (GRU) with 64 units to extract the time-dependent relationship [113]. This method achieved an accuracy of 99.40% on the BCI competition IV-2a dataset.

In Table 8, a summary of the discussed architectures is given.

**Table 8.** Other CNN-based architectures reviewed with their accuracies and architecture.

Authors	Accuracy	Dataset	MI Tasks	Architecture
Gomes et al. [110]	89.13%	BCI IV-2b	RH, LH	CNN + RF
Ma et al. [111]	90%	BCI IV-2a	RH, LH, BL, T	CNN + Transformer
Gao et al. [112]	80.7%	BCI IV-2a	RH, LH, BL, T	CNN + GRU
Ye et al. [113]	99.40%	BCI IV-2a	RH, LH, BL, T	CNN + GRU

LH: left hand, RH: right hand, RL: right leg, RS: resting state, BL: both legs, T: tongue.

#### 4.2. Other Methods

Almagor et al. utilized an autoencoder to denoise the signal, comprising of three layers for the encoder and three layers for the decoder, followed by a feature extractor (filter-bank common spatial patterns) and a feature selector (mutual information-based best individual feature), to act as an input to the SVM classifier [114]. The proposed method achieved an accuracy of 59.6% on two MI tasks (right hand, relax) on their privately collected dataset.

Stephe et al. [115] proposed a method which first used an empirical mode decomposition (EMD) to decompose the EEG signals into intrinsic mode functions (IMFs), and then these signals were feed into a GAN which was trained to distinguish between EEG signals for two different motor imagery tasks (right hand, right foot). The proposed method achieved an accuracy of 95.29% on the BCI competition III-4a dataset.

Xu et al. proposed a method where time, frequency, time–frequency, and spatial features were extracted from the raw EEG signal, then fed into a multilayer restricted Boltzmann machine network followed by a support vector machine for the classification [94]. The proposed method achieved an accuracy of 78.50% on the BCI competition IV-2a dataset.

Jiang et al. proposed an autoencoder for dimensionality reduction, in combination with a transformer layer, which consisted of an encoder and a decoder, each with a multihead attention layer and a feedforward layer [116]. The proposed method was tested on the BCI Competition III-3 dataset and achieved an accuracy of 91.30% for left-hand and right-hand MI tasks.

A short summary of the above papers is presented in Table 9.

**Table 9.** Other reviewed architectures and their accuracies and architecture.

Authors	Accuracy	Dataset	MI Tasks	Architecture
Almagor et al. [114]	59.6%	Private	RH, RS	AE + SVM
Stephe et al. [115]	95.29%	BCI III-4a	RH, RF	EMD + GAN
Xu et al. [94]	78.50%	BCI IV-2a	RH, LH, BL, T	RBM + SVM
Jiang et al. [116]	91.30%	BCI III-3	LH, RH	AE + transformer

LH: left hand, RH: right hand, RS: resting state, BL: Both legs, T: tongue, RF: right foot.

#### 5. Discussion

This literature review provides a comprehensive overview of the flourishing research landscape within the realm of brain–computer interfaces (BCIs) utilizing electroencephalograms (EEGs), with a specific focus on the motor imagery paradigm. Notably, the integration of deep learning methodologies, particularly convolutional neural networks (CNNs), has emerged as a prominent and successful approach, yielding notable improvement in terms of accuracy rates when applied to relevant datasets. It is imperative to acknowledge that the efficacy of deep learning and hybrid deep learning, comes hand in hand with the substantial computational power it demands, thereby signifying a notable cost implication. The complex nature of CNN algorithms necessitates a substantial volume of training patterns, ranging from tens of thousands to even millions in certain instances, to facilitate an optimal performance and robust generalization. Complex algorithms benefit from extensive and robust datasets, with the BCI competition IV-2a dataset serving as a key reference point for researchers. However, the EEG-based BCI field urgently requires more comprehensive and larger public datasets to propel progress and innovation. While combining datasets may seem like a viable solution to enhance dataset size, the diverse sampling rates, electrode configurations, and motor imagery tasks across different datasets present integration challenges. To address this, some researchers turn to data augmentation techniques, such as utilizing GANs, to artificially expand dataset size. This not only helps prevent overfitting but also enhances model resilience, particularly in the presence of real-world environmental noise, resulting in more robust and practical EEG-based BCI models. Besides the requirement for large datasets, these complex algorithms also require a vast amount of computational power and time to train. For this reason, some authors such

as Zhang et al. [61] have addressed this problem by using techniques such as pruning, to reduce the amount of time needed to train the network. As the field advances, it becomes imperative to not only address the computational demands of these intricate architectures but also to simultaneously enhance their efficiency and real-world applicability.

Within the realm of brain–computer interface (BCI) research, a consistent pattern emerges: the notable increases in accuracy can predominantly be attributed to the initial stages of preprocessing and the extraction of relevant features. This encompasses activities such as enhancing signals and exploring the time–frequency domain. This pivotal preparatory phase, which involves refining raw signals and extracting informative features, stands as a pivotal factor in achieving substantial progress in accuracy across various algorithmic approaches, which can be seen in Korhan et al.’s paper [64], which compares the same network architecture with different feature inputs.

The concept of “BCI illiteracy” [20] holds some significance. BCI illiteracy refers to the challenge that some individuals may face in effectively modulating their brain activity to produce distinguishable patterns during motor imagery tasks. This phenomenon can stem from various factors, such as a lack of familiarity with the task, insufficient cognitive engagement, or physiological variations that affect EEG signals. Addressing MI illiteracy is essential in designing inclusive and accessible BCIs, necessitating personalized training protocols, task adaptations, and innovative algorithms to accommodate users with varying levels of proficiency in generating discernible MI-related EEG patterns.

In the realm of motor imagery (MI), the number of commands that can be reliably extracted is inherently limited, as can be observed by the dataset’s distinct classes. This limitation arises due to the finite nature of distinct motor imagery tasks that users can effectively perform and differentiate in their mental simulations. The challenge lies in striking a balance between expanding the range of commands for diverse applications and maintaining a manageable set of tasks that users can consistently generate through mental imagery. This constraint underscores the importance of thoughtful task selection and user-centric design in the development of effective systems within the MI paradigm.

Finally, Table 10 offers a comparative analysis of the aforementioned algorithms, showcasing their respective strengths, weaknesses, and prospects for future research.

**Table 10.** Advantages and disadvantages of the reviewed methods.

Architecture	Advantages	Disadvantages
<b>Deep learning</b>	High accuracy	High computational complexity, requires a vast quantity of data
CNN	Effective feature extraction	Computationally expensive
Auto encoder	Unsupervised feature learning	Potential for vanishing/exploding gradients
Neural network	High accuracy	Hyperparameter tuning may be needed
<b>Hybrid deep learning</b>	Combination of multiple architectures	Increased complexity, requires more computational resources, requires a vast quantity of data, requires more time to train the model
CNN + LSTM	Combined spatial and temporal information for improved accuracy	Computationally expensive
CNN + AE	Benefit from both spatial feature extraction and unsupervised pre-training	Computationally expensive, potential for vanishing/exploding gradients

## 6. Future Research Directions

The consideration of computational time in real-time applications is crucial when evaluating machine learning and deep learning techniques. While deep learning indeed achieves elevated accuracy rates, the question of whether the benefits outweigh the computational burden becomes paramount. Although the inference of DL is quite fast and can be considered as real time, DL intricate architectures demand significant computational resources. That is, very powerful computers (in terms of processing power, memory, GPUs) are needed to train DL models, typically in a significant time. Balancing the pursuit of accuracy with the deployment of DL models that need less resources and less time to train remains a pivotal challenge.

The transition from a controlled lab environment with stable patients to a real-time application in a dynamic real-world setting introduces a notable shift in challenges. In the real environment, various types of noise can infiltrate the EEG signal, factors such as environmental interference, movement artifacts, and physiological variations. These noise sources can significantly degrade signal quality, making accurate interpretation and classification more difficult. Therefore, addressing the complexities of the real-world research field necessitates increased attention and consideration.

The dataset's diversity and quality are crucial aspects, encompassing a range of characteristics specific to each dataset. Nevertheless, the absence of a standard benchmark dataset presents a notable constraint. This gap could potentially be mitigated through augmentation methods, including artificial augmentation. By layering additional variations onto existing data, these techniques aim to expand the dataset's horizon, simulating diverse real-world scenarios. However, it is crucial to recognize that artificial augmentation might fall short in fully replicating the complexities of human experiences and interactions. Moreover, uncertainties persist regarding the exact conditions under which the experiments are conducted, adding an additional layer of complexity to the datasets.

## 7. Conclusions

In this paper, a thorough review of motor imagery (MI) EEG-based brain–computer interface (BCI) techniques was presented. This study delved into the realm of both deep learning and hybrid deep learning methodologies. Through an exploration of recent progress, this literature review provided a comprehensive and up-to-date insight into the advancements within this domain. Moreover, a discussion on the presented techniques was provided along with a comparison among them and future research directions.

**Author Contributions:** Conceptualization, K.K.; methodology, K.K.; formal analysis, K.K. and E.L.; investigation, E.L.; resources, E.L.; data curation, E.L.; writing—original draft preparation, E.L.; writing—review and editing, K.K.; visualization, E.L.; supervision, G.P.; 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 applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

## References

1. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller; Vaughan, T.M. Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [[CrossRef](#)] [[PubMed](#)]
2. Tariq, M.; Trivailo, P.M.; Simic, M. EEG-Based BCI Control Schemes for Lower-Limb Assistive-Robots. *Front. Hum. Neurosci.* **2018**, *12*, 312. [[CrossRef](#)] [[PubMed](#)]
3. Maksimenko, V.A.; van Heukelum, S.; Makarov, V.V.; Kelderhuis, J.; Lüttjohann, A.; Koronovskii, A.A.; Hramov, A.E.; van Luijtelaar, G. Absence Seizure Control by a Brain Computer Interface. *Sci. Rep.* **2017**, *7*, 2487. [[CrossRef](#)]

4. Bonnet, L.; Lotte, F.; Lecuyer, A. Two brains, one game: Design and evaluation of a multiuser bci video game based on motor imagery. *IEEE Trans. Comput. Intell. AI Games* **2013**, *5*, 185–198. [[CrossRef](#)]
5. Belkacem, A.N.; Lakas, A. A Cooperative EEG-based BCI Control System for Robot-Drone Interaction. In Proceedings of the 2021 International Wireless Communications and Mobile Computing, IWCMC, Harbin, China, 28 June–2 July 2021; pp. 297–302. [[CrossRef](#)]
6. Simon, M.V.; Nuwer, M.R.; Szelényi, A. Electroencephalography, electrocorticography, and cortical stimulation techniques. *Handb. Clin. Neurol.* **2022**, *186*, 11–38. [[CrossRef](#)] [[PubMed](#)]
7. Glover, G.H. Overview of Functional Magnetic Resonance Imaging. *Neurosurg. Clin. N. Am.* **2011**, *22*, 133. [[CrossRef](#)]
8. Sato, S.; Smith, P.D. Magnetoencephalography. *J. Clin. Neurophysiol.* **1985**, *2*, 173–192. [[CrossRef](#)] [[PubMed](#)]
9. Marin, T.; Moore, J. Understanding near-infrared spectroscopy. *Adv. Neonatal Care* **2011**, *11*, 382–388. [[CrossRef](#)]
10. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain Computer Interfaces, a Review. *Sensors* **2012**, *12*, 1211. [[CrossRef](#)]
11. Schupp, H.T.; Flaisch, T.; Stockburger, J.; Junghöfer, M. Emotion and attention: Event-related brain potential studies. *Prog. Brain Res.* **2006**, *156*, 31–51. [[CrossRef](#)] [[PubMed](#)]
12. Al-Nafjan, A.; Hosny, M.; Al-Ohali, Y.; Al-Wabil, A. Review and Classification of Emotion Recognition Based on EEG Brain-Computer Interface System Research: A Systematic Review. *Appl. Sci.* **2017**, *7*, 1239. [[CrossRef](#)]
13. Padfield, N.; Zabalza, J.; Zhao, H.; Masero, V.; Ren, J. EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges. *Sensors* **2019**, *19*, 1423. [[CrossRef](#)]
14. Torres, P.E.P.; Torres, E.A.; Hernández-Álvarez, M.; Yoo, S.G. EEG-Based BCI Emotion Recognition: A Survey. *Sensors* **2020**, *20*, 5083. [[CrossRef](#)] [[PubMed](#)]
15. Lan, T.; Erdogmus, D.; Adami, A.; Mathan, S.; Pavel, M. Channel selection and feature projection for cognitive load estimation using ambulatory EEG. *Comput. Intell. Neurosci.* **2007**, *2007*, 074895. [[CrossRef](#)]
16. Li, Y.; Li, X.; Ratcliffe, M.; Liu, L.; Qi, Y.; Liu, Q. A real-time EEG-based BCI system for attention recognition in ubiquitous environment. In *UAAII'11-Proceedings of the 2011 International Workshop on Ubiquitous Affective Awareness and Intelligent Interaction*; Association for Computing Machinery: New York, NY, USA, 2011; pp. 33–39. [[CrossRef](#)]
17. Fabiani, G.E.; McFarland, D.J.; Wolpaw, J.R.; Pfurtscheller, G. Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). *IEEE Trans. Neural Syst. Rehabil. Eng.* **2004**, *12*, 331–338. [[CrossRef](#)] [[PubMed](#)]
18. Guneyasu, A.; Akin, H.L. An SSVEP based BCI to control a humanoid robot by using portable EEG device. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Osaka, Japan, 3–7 July 2013; pp. 6905–6908. [[CrossRef](#)]
19. Cincotti, F.; Pichiorri, F.; Arico, P.; Aloise, F.; Leotta, F.; Fallani, F.D.V.; Millan, J.D.R.; Molinari, M.; Mattia, D. EEG-based brain-computer interface to support post-stroke motor rehabilitation of the upper limb. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, San Diego, CA, USA, 28 August–1 September 2012; pp. 4112–4115. [[CrossRef](#)]
20. Lee, M.-H.; Kwon, O.-Y.; Kim, H.-K.; Lee, Y.-E.; Williamson, J.; Fazli, S.; Lee, S.-W. EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy. *Gigascience* **2019**, *8*, giz002. [[CrossRef](#)]
21. Tran, K.A.; Kondrashova, O.; Bradley, A.; Williams, E.D.; Pearson, J.V.; Waddell, N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med.* **2021**, *13*, 152. [[CrossRef](#)]
22. Karampidis, K.; Rousouliotis, M.; Euangelos, L.; Kavallieratou, E. A comprehensive survey of fingerprint presentation attack detection. *J. Surveill. Secur. Saf.* **2021**, *2*, 117–161. [[CrossRef](#)]
23. Karampidis, K.; Kavallieratou, E.; Papadourakis, G. A Dilated Convolutional Neural Network as Feature Selector for Spatial Image Steganalysis—A Hybrid Classification Scheme. *Pattern Recognit. Image Anal.* **2020**, *30*, 342–358. [[CrossRef](#)]
24. Karampidis, K.; Linardos, E.; Kavallieratou, E. StegoPass—Utilization of Steganography to Produce a Novel Unbreakable Biometric Based Password Authentication Scheme. In *Computational Intelligence in Security for Information Systems Conference*; Springer International Publishing: Cham, Switzerland, 2022; pp. 146–155. [[CrossRef](#)]
25. Karampidis, K.; Vasillopoulos, N.; Rodríguez, C.C.; del Blanco Adán, C.R.; Kavallieratou, E.; Santos, N.G. Overview of the ImageCLEFsecurity 2019: File Forgery Detection Tasks. In Proceedings of the Conference and Labs of the Evaluation Forum (CLEF 2019) | Conference and Labs of the Evaluation Forum (CLEF 2019), Lugano, Switzerland, 9–12 September 2019.
26. Karampidis, K.; Papadourakis, G. File type identification for digital forensics. *Lect. Notes Bus. Inf. Process.* **2016**, *249*, 266–274. [[CrossRef](#)]
27. Jiang, H.; Diao, Z.; Shi, T.; Zhou, Y.; Wang, F.; Hu, W.; Zhu, X.; Luo, S.; Tong, G.; Yao, Y.-D. A review of deep learning-based multiple-lesion recognition from medical images: Classification, detection and segmentation. *Comput. Biol. Med.* **2023**, *157*, 106726. [[CrossRef](#)]
28. Lauriola, I.; Lavelli, A.; Aioli, F. An introduction to Deep Learning in Natural Language Processing: Models, techniques, and tools. *Neurocomputing* **2022**, *470*, 443–456. [[CrossRef](#)]
29. Khademi, Z.; Ebrahimi, F.; Kordy, H.M. A review of critical challenges in MI-BCI: From conventional to deep learning methods. *J. Neurosci. Methods* **2023**, *383*, 109736. [[CrossRef](#)]
30. Li, G.; Lee, C.H.; Jung, J.J.; Youn, Y.C.; Camacho, D. Deep learning for EEG data analytics: A survey. *Concurr. Comput.* **2020**, *32*, e5199. [[CrossRef](#)]

31. Cao, Z. A review of artificial intelligence for EEG-based brain–computer interfaces and applications. *Brain Sci. Adv.* **2021**, *6*, 162–170. [[CrossRef](#)]
32. Rajwal, S.; Aggarwal, S. Convolutional Neural Network-Based EEG Signal Analysis: A Systematic Review. *Arch. Comput. Methods Eng.* **2023**, *1*, 3585–3615. [[CrossRef](#)]
33. Hosseini, M.P.; Hosseini, A.; Ahi, K. A Review on Machine Learning for EEG Signal Processing in Bioengineering. *IEEE Rev. Biomed. Eng.* **2021**, *14*, 204–218. [[CrossRef](#)] [[PubMed](#)]
34. Värbu, K.; Muhammad, N.; Muhammad, Y. Past, Present, and Future of EEG-Based BCI Applications. *Sensors* **2022**, *22*, 3331. [[CrossRef](#)] [[PubMed](#)]
35. Yadav, D.; Yadav, S.; Veer, K. A comprehensive assessment of Brain Computer Interfaces: Recent trends and challenges. *J. Neurosci. Methods* **2020**, *346*, 108918. [[CrossRef](#)] [[PubMed](#)]
36. 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](#)]
37. 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](#)]
38. Habashi, A.G.; Azab, A.M.; Eldawlatly, S.; Aly, G.M. Generative adversarial networks in EEG analysis: An overview. *J. Neuroeng. Rehabil.* **2023**, *20*, 40. [[CrossRef](#)] [[PubMed](#)]
39. BCI Competition II. Available online: <https://www.bbc.de/competition/ii/> (accessed on 23 June 2023).
40. BCI Competition III. Available online: <https://www.bbc.de/competition/iii/> (accessed on 23 June 2023).
41. Tangermann, M.; Müller, K.-R.; Aertsen, A.; Birbaumer, N.; Braun, C.; Brunner, C.; Leeb, R.; Mehring, C.; Miller, K.J.; Müller-Putz, G.R.; et al. Review of the BCI competition IV. *Front. Neurosci.* **2012**, *6*, 21084. [[CrossRef](#)] [[PubMed](#)]
42. Schalk, G.; McFarland, D.; Hinterberger, T.; Birbaumer, N.; Wolpaw, J. BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1034–1043. [[CrossRef](#)] [[PubMed](#)]
43. Schirrmester, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **2017**, *38*, 5391–5420. [[CrossRef](#)] [[PubMed](#)]
44. Ma, X.; Qiu, S.; Wei, W.; Wang, S.; He, H. Deep Channel-Correlation Network for Motor Imagery Decoding from the Same Limb. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 297–306. [[CrossRef](#)] [[PubMed](#)]
45. Supporting Data for “EEG Dataset and OpenBMI Toolbox for Three BCI Paradigms: An Investigation into BCI Illiteracy”. Available online: <http://gigadb.org/dataset/100542> (accessed on 26 September 2023).
46. Kaya, M.; Binli, M.K.; Ozbay, E.; Yanar, H.; Mishchenko, Y. A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces. *Sci. Data* **2018**, *5*, 180211. [[CrossRef](#)] [[PubMed](#)]
47. Cho, H.; Ahn, M.; Ahn, S.; Kwon, M.; Jun, S.C. EEG datasets for motor imagery brain-computer interface. *Gigascience* **2017**, *6*, gix034. [[CrossRef](#)] [[PubMed](#)]
48. Brodu, N.; Lotte, F.; Lécuyer, A. Exploring two novel features for EEG-based brain–computer interfaces: Multifractal cumulants and predictive complexity. *Neurocomputing* **2012**, *79*, 87–94. [[CrossRef](#)]
49. Scherer, R.; Faller, J.; Friedrich, E.V.C.; Opisso, E.; Costa, U.; Kübler, A.; Müller-Putz, G.R. Individually Adapted Imagery Improves Brain-Computer Interface Performance in End-Users with Disability. *PLoS ONE* **2015**, *10*, e0123727. [[CrossRef](#)] [[PubMed](#)]
50. Ofner, P.; Schwarz, A.; Pereira, J.; Müller-Putz, G.R. Upper limb movements can be decoded from the time-domain of low-frequency EEG. *PLoS ONE* **2017**, *12*, e0182578. [[CrossRef](#)]
51. Saxena, A. An Introduction to Convolutional Neural Networks. *Int. J. Res. Appl. Sci. Eng. Technol.* **2015**, *10*, 943–947. [[CrossRef](#)]
52. Lun, X.; Yu, Z.; Chen, T.; Wang, F.; Hou, Y. A Simplified CNN Classification Method for MI-EEG via the Electrode Pairs Signals. *Front. Hum. Neurosci.* **2020**, *14*, 559321. [[CrossRef](#)] [[PubMed](#)]
53. Dose, H.; Möller, J.S.; Puthusserypady, S.; Iversen, H.K. A deep learning MI-EEG classification model for BCIS. In Proceedings of the European Signal Processing Conference, Rome, Italy, 3–7 September 2018; Volume 2018, pp. 1676–1679. [[CrossRef](#)]
54. Miao, M.; Hu, W.; Yin, H.; Zhang, K. Spatial-Frequency Feature Learning and Classification of Motor Imagery EEG Based on Deep Convolution Neural Network. *Comput. Math. Methods Med.* **2020**, *2020*, 1981728. [[CrossRef](#)] [[PubMed](#)]
55. Zhao, R.; Wang, Y.; Cheng, X.; Zhu, W.; Meng, X.; Niu, H.; Cheng, J.; Liu, T. A mutli-scale spatial-temporal convolutional neural network with contrastive learning for motor imagery EEG classification. *Med. Nov. Technol. Devices* **2023**, *17*, 100215. [[CrossRef](#)]
56. Liu, X.; Xiong, S.; Wang, X.; Liang, T.; Wang, H.; Liu, X. A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification. *Biomed. Signal Process Control* **2023**, *81*, 104456. [[CrossRef](#)]
57. Han, Y.; Wang, B.; Luo, J.; Li, L.; Li, X. A classification method for EEG motor imagery signals based on parallel convolutional neural network. *Biomed. Signal Process Control* **2022**, *71*, 103190. [[CrossRef](#)]
58. Ma, W.; Gong, Y.; Zhou, G.; Liu, Y.; Zhang, L.; He, B. A channel-mixing convolutional neural network for motor imagery EEG decoding and feature visualization. *Biomed. Signal Process Control* **2021**, *70*, 103021. [[CrossRef](#)]
59. Ak, A.; Topuz, V.; Midi, I. Motor imagery EEG signal classification using image processing technique over GoogLeNet deep learning algorithm for controlling the robot manipulator. *Biomed. Signal Process Control* **2022**, *72*, 103295. [[CrossRef](#)]
60. Musallam, Y.K.; AlFassam, N.I.; Muhammad, G.; Amin, S.U.; Alsulaiman, M.; Abdul, W.; Altaheri, H.; Bencherif, M.A.; Algabri, M. Electroencephalography-based motor imagery classification using temporal convolutional network fusion. *Biomed. Signal Process Control* **2021**, *69*, 102826. [[CrossRef](#)]

61. Zhang, J.; Li, K. A Pruned Deep Learning Approach for Classification of Motor Imagery Electroencephalography Signals. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Glasgow, Scotland, UK, 11–15 July 2022; Volume 2022, pp. 4072–4075. [[CrossRef](#)]
62. Vishnupriya, R.; Robinson, N.; Reddy, R.; Guan, C. Performance Evaluation of Compressed Deep CNN for Motor Imagery Classification using EEG. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Mexico, Russia, 1–5 November 2021; pp. 795–799. [[CrossRef](#)]
63. Shajil, N.; Mohan, S.; Srinivasan, P.; Arivudaiyanambi, J.; Murrugesan, A.A. Multiclass Classification of Spatially Filtered Motor Imagery EEG Signals Using Convolutional Neural Network for BCI Based Applications. *J. Med. Biol. Eng.* **2020**, *40*, 663–672. [[CrossRef](#)]
64. Korhan, N.; Dokur, Z.; Olmez, T. Motor imagery based EEG classification by using common spatial patterns and convolutional neural networks. In Proceedings of the 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019, Istanbul, Turkey, 24–26 April 2019. [[CrossRef](#)]
65. Alazrai, R.; Abuhijleh, M.; Alwanni, H.; Daoud, M.I. A Deep Learning Framework for Decoding Motor Imagery Tasks of the Same Hand Using EEG Signals. *IEEE Access* **2019**, *7*, 109612–109627. [[CrossRef](#)]
66. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process Syst.* **2012**, *25*, 84–90. [[CrossRef](#)]
67. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; Volume 2016, pp. 770–778. [[CrossRef](#)]
68. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A.; Liu, W.; et al. Going Deeper with Convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9. [[CrossRef](#)]
69. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6848–6856. [[CrossRef](#)]
70. Wei, X.; Ortega, P.; Faisal, A.A. Inter-subject deep transfer learning for motor imagery EEG decoding. In Proceedings of the International IEEE/EMBS Conference on Neural Engineering, NER, Virtual Event, 4–6 May 2021; pp. 21–24. [[CrossRef](#)]
71. Zhang, K.; Robinson, N.; Lee, S.W.; Guan, C. Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network. *Neural Netw.* **2021**, *136*, 1–10. [[CrossRef](#)] [[PubMed](#)]
72. Zhang, R.; Zong, Q.; Dou, L.; Zhao, X.; Tang, Y.; Li, Z. Hybrid deep neural network using transfer learning for EEG motor imagery decoding. *Biomed. Signal Process Control* **2021**, *63*, 102144. [[CrossRef](#)]
73. Limpiti, T.; Seetanathum, K.; Sricom, N.; Puttarak, N. Transfer Learning for Classifying Motor Imagery EEG: A Comparative Study. In Proceedings of the BMEiCON 2021-13th Biomedical Engineering International Conference, Ayutthaya, Thailand, 19–21 November 2021. [[CrossRef](#)]
74. Wei, M.; Yang, R.; Huang, M. Motor imagery EEG signal classification based on deep transfer learning. *Proc. IEEE Symp. Comput. Based Med. Syst.* **2021**, *2021*, 85–90. [[CrossRef](#)]
75. Roy, A.M. Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for EEG MI multiclassification in brain–computer interface. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105347. [[CrossRef](#)]
76. Chen, C.Y.; Wang, W.J.; Chen, C.C. Multiclass Classification of EEG Motor Imagery Signals Based on Transfer Learning. In Proceedings of the 2022 8th International Conference on Applied System Innovation, ICASI, Nantou, Taiwan, 22–23 April 2022; pp. 140–143. [[CrossRef](#)]
77. Solorzano-Espindola, C.E.; Zamora, E.; Sossa, H. Multi-subject classification of Motor Imagery EEG signals using transfer learning in neural networks. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Mexico, Russia, 1–5 November 2021; pp. 1006–1009. [[CrossRef](#)]
78. Li, D.; Wang, J.; Xu, J.; Fang, X.; Ji, Y. Cross-Channel Specific-Mutual Feature Transfer Learning for Motor Imagery EEG Signals Decoding. In *IEEE Transactions on Neural Networks and Learning Systems*; IEEE: New York, NY, USA, 2023. [[CrossRef](#)]
79. Suhaimi, N.S.; Yusoff, M.Z.; Saad, M.N.M. Artificial Neural Network Analysis on Motor Imagery Electroencephalogram. In Proceedings of the 2022 IEEE 5th International Symposium in Robotics and Manufacturing Automation (ROMA), Malacca, Malaysia, 6–8 August 2022; IEEE: New York, NY, USA, 2022. [[CrossRef](#)]
80. Cheng, D.; Liu, Y.; Zhang, L. Exploring Motor Imagery EEG Patterns for Stroke Patients with Deep Neural Networks. In Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings, Calgary, AB, Canada, 15–20 April 2018; pp. 2561–2565. [[CrossRef](#)]
81. Yohanandan, S.A.C.; Kiral-Kornek, I.; Tang, J.; Mshford, B.S.; Asif, U.; Harrer, S. A Robust Low-Cost EEG Motor Imagery-Based Brain-Computer Interface. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Honolulu, HI, USA, 18–21 July 2018; pp. 5089–5092. [[CrossRef](#)]
82. Kumar, S.; Sharma, A.; Mamun, K.; Tsunoda, T. A Deep Learning Approach for Motor Imagery EEG Signal Classification. In Proceedings of the Proceedings-Asia-Pacific World Congress on Computer Science and Engineering 2016 and Asia-Pacific World Congress on Engineering 2016, APWC on CSE/APWCE 2016, Nadi, Fiji, 5–6 December 2016; pp. 34–39. [[CrossRef](#)]

83. Pinaya, W.H.L.; Vieira, S.; Garcia-Dias, R.; Mechelli, A. Autoencoders. In *Machine Learning: Methods and Applications to Brain Disorders*; Academic Press: Cambridge, MA, USA, 2020; pp. 193–208. [[CrossRef](#)]
84. Khan, G.H.; Khan, N.A.; Altaf, M.A.B.; Abid, M.U.R. Classifying Single Channel Epileptic EEG data based on Sparse Representation using Shallow Autoencoder. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Mexico, Russia, 1–5 November 2021; pp. 643–646. [[CrossRef](#)]
85. Autthasan, P.; Chaisaen, R.; Sudhawiyangkul, T.; Rangpong, P.; Kiatthaveephong, S.; Dilokthanakul, N.; Bhakdisongkham, G.; Phan, H.; Guan, C.; Wilaiprasitporn, T. MIN2Net: End-to-End Multi-Task Learning for Subject-Independent Motor Imagery EEG Classification. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 2105–2118. [[CrossRef](#)]
86. Patrick, M.K.; Adekoya, A.F.; Mighty, A.A.; BEdward, Y. Capsule Networks—A survey. *J. King Saud. Univ.-Comput. Inf. Sci.* **2022**, *34*, 1295–1310. [[CrossRef](#)]
87. Ha, K.W.; Jeong, J.W. Decoding Two-Class Motor Imagery EEG with Capsule Networks. In Proceedings of the 2019 IEEE International Conference on Big Data and Smart Computing, BigComp 2019-Proceedings, Kyoto, Japan, 27 February–2 March 2019. [[CrossRef](#)]
88. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
89. Leon-Urbano, C.; Ugarte, W. End-to-end electroencephalogram (EEG) motor imagery classification with Long Short-Term. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence, SSCI, Canberra, ACT, Australia, 1–4 December 2020; pp. 2814–2820. [[CrossRef](#)]
90. Saputra, M.F.; Setiawan, N.A.; Ardiyanto, I. Deep Learning Methods for EEG Signals Classification of Motor Imagery in BCI. *IJITEE (Int. J. Inf. Technol. Electr. Eng.)* **2019**, *3*, 80–84. [[CrossRef](#)]
91. Hwang, J.; Park, S.; Chi, J. Improving Multi-Class Motor Imagery EEG Classification Using Overlapping Sliding Window and Deep Learning Model. *Electronics* **2023**, *12*, 1186. [[CrossRef](#)]
92. Ma, X.; Qiu, S.; Du, C.; Xing, J.; He, H. Improving EEG-Based Motor Imagery Classification via Spatial and Temporal Recurrent Neural Networks. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Honolulu, HI, USA, 18–21 July 2018; pp. 1903–1906. [[CrossRef](#)]
93. Yan, W.Q. Boltzmann Machines. In *Computational Methods for Deep Learning*; Texts in Computer Science; Springer: Cham, Switzerland, 2021. [[CrossRef](#)]
94. Xu, J.; Zheng, H.; Wang, J.; Li, D.; Fang, X. Recognition of EEG Signal Motor Imagery Intention Based on Deep Multi-View Feature Learning. *Sensors* **2020**, *20*, 3496. [[CrossRef](#)]
95. Finn, C.; Abbeel, P.; Levine, S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. In Proceedings of the 34th International Conference on Machine Learning, ICML, Sydney, Australia, 6–11 August 2017; Volume 3, pp. 1856–1868. Available online: <https://arxiv.org/abs/1703.03400v3> (accessed on 24 July 2023).
96. Li, D.; Ortega, P.; Wei, X.; Faisal, A. Model-agnostic meta-learning for EEG motor imagery decoding in brain-computer-interfacing. In Proceedings of the International IEEE/EMBS Conference on Neural Engineering, NER, Virtual, 4–6 May 2021; pp. 527–530. [[CrossRef](#)]
97. Le-Khac, P.H.; Healy, G.; Smeaton, A.F. Contrastive Representation Learning: A Framework and Review. *IEEE Access* **2020**, *8*, 193907–193934. [[CrossRef](#)]
98. Han, J.; Gu, X.; Lo, B. Semi-Supervised Contrastive Learning for Generalizable Motor Imagery EEG Classification. In Proceedings of the 2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks, BSN, Athens, Greece, 27–30 July 2021. [[CrossRef](#)]
99. Hua, Y.; Guo, J.; Zhao, H. Deep Belief Networks and deep learning. In Proceedings of the 2015 International Conference on Intelligent Computing and Internet of Things, ICIT, Harbin, China, 17–18 January 2015; pp. 1–4. [[CrossRef](#)]
100. Li, M.-A.; Zhang, M.; Sun, Y.-J. A novel motor imagery EEG recognition method based on deep learning. In *2016 International Forum on Management, Education and Information Technology Application*; Atlantis Press: Amsterdam, The Netherlands, 2016; pp. 728–733. [[CrossRef](#)]
101. Amin, S.U.; Altaheri, H.; Muhammad, G.; Abdul, W.; Alsulaiman, M. Attention-Inception and Long-Short-Term Memory-Based Electroencephalography Classification for Motor Imagery Tasks in Rehabilitation. *IEEE Trans. Ind. Inf.* **2022**, *18*, 5412–5421. [[CrossRef](#)]
102. Khademi, Z.; Ebrahimi, F.; Kordy, H.M. A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals. *Comput. Biol. Med.* **2022**, *143*, 105288. [[CrossRef](#)]
103. Ecthioui, A.; Mlaouah, A.; Zouch, W.; Ghorbel, M.; Mhiri, C.; Hamam, H. A Novel Convolutional Neural Network Classification Approach of Motor-Imagery EEG Recording Based on Deep Learning. *Appl. Sci.* **2021**, *11*, 9948. [[CrossRef](#)]
104. Li, C.; Yang, H.; Wu, X.; Zhang, Y. Improving EEG-Based Motor Imagery Classification Using Hybrid Neural Network. In Proceedings of the 2021 IEEE 9th International Conference on Information, Communication and Networks, ICICN, Xi'an, China, 25–28 November 2021; pp. 486–489. [[CrossRef](#)]
105. Li, J.; Shi, Z.; Li, Y. Research on EEG-Based Motor Imagery Tasks Recognition Using Deep Learning Approach. *Lect. Notes Electr. Eng.* **2022**, *950*, 416–425. [[CrossRef](#)]
106. Fadel, W.; Kollod, C.; Wahdow, M.; Ibrahim, Y.; Ulbert, I. Multi-Class Classification of Motor Imagery EEG Signals Using Image-Based Deep Recurrent Convolutional Neural Network. In Proceedings of the 8th International Winter Conference on Brain-Computer Interface, BCI, Gangwon, Republic of Korea, 26–28 February 2020. [[CrossRef](#)]

107. Tabar, Y.R.; Halici, U. A novel deep learning approach for classification of EEG motor imagery signals. *J. Neural Eng.* **2016**, *14*, 016003. [[CrossRef](#)]
108. 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](#)]
109. Hwaidi, J.F.; Chen, T.M. Classification of Motor Imagery EEG Signals Based on Deep Autoencoder and Convolutional Neural Network Approach. *IEEE Access* **2022**, *10*, 48071–48081. [[CrossRef](#)]
110. Gomes, J.C.; Rodrigues, M.C.A.; Santos, W.P.D. ASTERI: Image-based representation of EEG signals for motor imagery classification. *Res. Biomed. Eng.* **2022**, *38*, 661–681. [[CrossRef](#)]
111. Ma, Y.; Song, Y.; Gao, F. A novel hybrid CNN-Transformer model for EEG Motor Imagery classification. In Proceedings of the International Joint Conference on Neural Networks, Padua, Italy, 18–23 July 2022. [[CrossRef](#)]
112. Gao, S.; Yang, J.; Shen, T.; Jiang, W. A Parallel Feature Fusion Network Combining GRU and CNN for Motor Imagery EEG Decoding. *Brain Sci.* **2022**, *12*, 1233. [[CrossRef](#)] [[PubMed](#)]
113. Liu, J.; Ye, F.; Xiong, H. Multi-class motor imagery EEG classification method with high accuracy and low individual differences based on hybrid neural network. *J. Neural Eng.* **2021**, *18*, 0460f1. [[CrossRef](#)]
114. Almagor, O.; Avin, O.; Rosipal, R.; Shriki, O. Using Autoencoders to Denoise Cross-Session Non-Stationarity in EEG-Based Motor-Imagery Brain-Computer Interfaces. In Proceedings of the 2022 IEEE 16th International Scientific Conference on Informatics, Informatics 2022-Proceedings, Poprad, Slovakia, 23–25 November 2022; pp. 24–28. [[CrossRef](#)]
115. Stephe, S.; Jayasankar, T.; Kumar, K.V. Motor Imagery EEG Recognition using Deep Generative Adversarial Network with EMD for BCI Applications. *Tech. Gaz.* **2022**, *29*, 92–100. [[CrossRef](#)]
116. Jiang, R.; Sun, L.; Wang, X.; Xu, Y. Application of Transformer with Auto-Encoder in Motor Imagery EEG Signals. In Proceedings of the 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, 1–3 November 2022; Volume 2022, pp. 631–637. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.