A Survey on Deep Learning-based Non-Invasive Brain Signals: Recent Advances and New Frontiers

Xiang Zhang[1,5], Lina Yao[1], Xianzhi Wang[2], Jessica Monaghan[3], David McAlpine[3], Yu Zhang[4]

[1] University of New South Wales, Australia
[2] University of Technology Sydney, Australia
[3] Macquarie university, Australia
[4] Lehigh University, USA
[5] Harvard University, USA

E-mail: `xiang_zhang@hms.harvard.edu`,
`lina.yao@unsw.edu.au`, `xianzhi.wang@uts.edu.au`,
`{jessica.monaghan,david.mcalpine}@mq.edu.au`, `yuzi20@lehigh.edu`

Abstract. Brain signals refer to the biometric information collected from the human brain. The research on brain signals aims to discover the underlying neurological or physical status of the individuals by signal decoding. The emerging deep learning techniques have improved the study of brain signals significantly in recent years. In this work, we first present a taxonomy of non-invasive brain signals and the basics of deep learning algorithms. Then, we provide the frontiers of applying deep learning for non-invasive brain signals analysis, by summarizing a large number of recent publications. Moreover, upon the deep learning-powered brain signal studies, we report the potential real-world applications which benefit not only disabled people but also normal individuals. Finally, we discuss the opening challenges and future directions.

1. Introduction

Brain signals measure the instinct biometric information from the human brain, which reflects the user's passive or active mental state. Through precise brain signal decoding, we can recognize the underlying psychological and physical status of the user and further improve his/her life quality. Based on the signal collection, brain signals contain invasive signals and non-invasive signals. The former are acquired by electrodes deployed under the scalp while the latter are collected upon human scalp without electrodes being inserted. In this survey, we mainly consider non-invasive brain signals[1].

1.1. General Workflow

Figure 1 shows the general paradigm of brain signal decoding, which receives brain signals and produces the user's latent informatics. The workflow includes several key components: brain signal collection, signal preprocessing, feature extraction, classification, and data analysis. The brain signals are collected from humans and sent to the preprocessing component for denoising and enhancement. Then, the discriminating features are extracted from the processed signals and sent to the classifier for further analysis.

The collection methods differ from signal to signal. For example, EEG signals measure the voltage fluctuation resulting from ionic current within the neurons of the brain. Collecting EEG signals requires placing a series of electrodes on the scalp of the human head to record the electrical activity of the brain. Since the ionic current generated within the brain is measured at the scalp, obstacles (*e.g.*, skull) greatly decrease the signal quality— the fidelity of the collected EEG signals, measured as Signal-to-Noise Ratio (SNR), is only approximately 5% of that of original brain signals [1]. The collection methods of more non-invasive signals can be found in Appendix A.

Therefore, brain signals are usually preprocessed before feature extraction to increase the SNR. The preprocessing component contains multiple steps such as signal cleaning (smoothing the noisy signals or resolving the inconsistencies), signal normalization (normalizing each channel of the signals along time-axis), signal enhancement (removing direct current), and signal reduction (presenting a reduced representation of the signal).
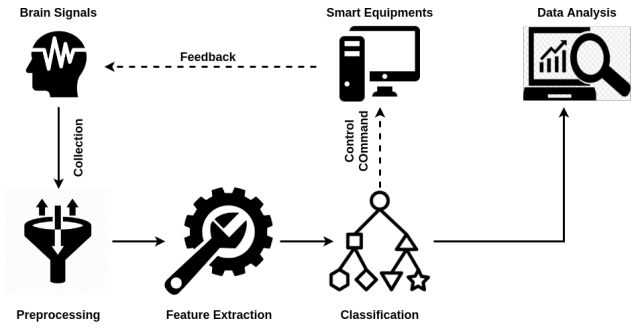

Figure 1: Generally workflow of brain signal analysis. It is named as a Brain-Computer Interface if the classified signal are used to control smart equipment (dashed lines).

Feature extraction refers to the process of extracting discriminating features from the input signals through domain knowledge. Traditional features are extracted from time-domain (*e.g.*, variance, mean value, kurtosis), frequency-domain (*e.g.*, fast Fourier transform), and time-frequency domains (*e.g.*, discrete wavelet transform). They will enrich distinguishable information regarding user intention. Feature extraction is highly dependent on the domain knowledge. For example, neuroscience knowledge is required to extract distinctive features from motor imagery EEG signals. Manual feature extraction is also time-consuming and difficult. Recently, deep learning provides a better option to automatically extract distinguishable features.

The classification component refers to the machine learning algorithms that classify the extracted features into logical control signals recognizable by external devices. Deep learning algorithms are shown to be more powerful than traditional classifiers [2, 3, 4].

The classification results reflect the user's psychological or physical status and can inspire further information analysis. This is widely used in real-world applications such as neurological disorder diagnosis, emotion measuring, and driving fatigue detection. Appropriate treatment, therapy, and precaution could be conducted based on the analysis results.

In specific, the system is called a Brain-Computer Interface (BCI) while the decoded brain signals are converted into digital commands to control the smart equipment and react with the user (dashed lines in Figure 1). BCI[2] systems interpret the human brain patterns into

---

[1] Without specification, the brain signals mentioned in this work refer to non-invasive signals.

[2] Apart from BCI, there are a number of similar terms to define the system that machines are directly controlled by human brain signals,

messages or commands to communicate with the outer world [5]. BCI is generally a closed-loop system with an external device (*e.g.*, wheelchair and robotic arm), which can directly serve the user. In contrast, brain signal analysis doesn't require a specific device as long as the analysis results can benefit society and individuals.

In this survey, we summarize the state-of-the-art studies which adopt deep learning models: 1) for feature extraction only; 2) for classification only; 3) for both feature extraction and classification. The details will be introduced in Section 4. Brain signal underpins many novel applications that are important to people's daily life. For example, the brain signal-based user identification system, with high fake-resistance, allows normal people to enjoy enhanced entertainment and security [6]; for people with psychological/physical deceases or disabilities, brain signals enable them to control smart device such as wheelchairs, home appliances, and robots. We present a wide range of deep learning-based brain signal applications in Section 5.

1.2. Why Deep Learning?

Although traditional brain signal system has made tremendous progress [7, 8], it still faces significant challenges. First, brain signals are easily corrupted by various biological (*e.g.*, eye blinks, muscle artifacts, fatigue, and the concentration level) and environmental artifacts (*e.g.*, noises) [7]. Therefore, it is crucial to distill informative data from corrupted brain signals and build a robust system that works in different situations. Second, it faces the low SNR of non-stationary electrophysiological brain signals [9]. The low SNR cannot be easily addressed by traditional preprocessing or feature extraction methods due to the time complexity of those method and the risk of information loss [10]. Third, feature extraction highly depends on human expertise in the specific domain. For example, it requires the basic biological knowledge to investigate sleep state through Electroencephalogram (EEG) signals. Human experience may help on certain aspects but fall insufficient in more general circumstances. An automatic feature extraction method is highly desirable. Moreover, most existing machine learning research focuses on static data and therefore, cannot classify rapidly changing brain signals accurately. For instance, the state-of-the-art classification accuracy for multi-class motor imagery EEG is generally below 80% [11]. It requires novel learning methods to deal with dynamical data streams in brain signal systems.

Until now, deep learning has been applied extensively in brain signal applications and shown success in addressing the above challenges [12, 13]. Deep learning has two advantages. First, it works directly on raw brain signals, thus avoiding the time-consuming preprocessing and feature extraction. Second, deep neural networks can capture both representative high-level features and latent dependencies through deep structures.

1.3. Why this Survey is Necessary?

We conduct this survey for three reasons. First, there lacks a comprehensive survey on the non-invasive brain signals. Table 1 shows a summary of the existing survey on brain signals. As our best knowledge, the limited existing surveys [14, 24, 7, 11, 5, 8, 15] only focus on partial EEG signals. For example, Lotte et al. [11] and Wang et al. [18] focus on general EEG without analyzing EEG subtypes; Cecotti et al. [28] focus on Event-Related Potentials (ERP); Haseer et al. [29] focus on functional near-infrared spectroscopy (fNIRS); Mason et al. [15] brief the neurological phenomenons like event-related desynchronization (ERD), P300, SSVEP, Visual Evoked Potentials (VEP), Auditory Evoked Potentials (AEP) but have not organized them systematically; Abdulkader et al. [7] present a topology of brain signals but have not mentioned spontaneous EEG and Rapid Serial Visual Presentation (RSVP); Lotte et al. [5] have not considered ERD and RSVP; VEP should be a subtype of ERP in [8]. Ahn et al. [21] review the performance variation in MI-EEG based BCI systems. Roy et al. [17] list some deep learning-based EEG studies but present little technical inspirations and have less analysis on deep learning algorithms, they also failed to investigate other non-invasive brain signals beyond EEG. In particular, compared to [17], this work provides a better introduction of deep learning including the basic concepts, algorithms, and popular models (Section 3 and Appendix B). Moreover, this paper discusses the high-level guidelines in brain signal analysis in terms of the brain signal paradigms, the suitable deep learning frameworks and the promising real-world applications (Section 6).

Second, few research has investigated the association between deep learning ([30, 31]) and brain signals ([32, 7, 11, 5, 8, 15]). To the best of our knowledge, this paper is in the first batch of comprehensive survey on recent advances on deep learning-based brain signals. We also point out frontiers and promising directions in this area.

Lastly, the existing surveys focus on specific areas or applications and lack an overview of broad scenarios. For example, Litjens et al. [16] summarize several deep neural network concepts aiming at medical image analysis; Soekadar et al. [20] review the BCI systems and machine learning methods for stroke-related motor paralysis based on Sensori-Motor Rhythms (SMR); Vieira et al. [33] investigate the application of brain signals on the neurological disorder and psychiatric.

1.4. Our Contributions

This survey can mainly benefit: 1) the researchers with computer science background who are inter-

---

like Brain-Machine Interface (BMI), Brain Interface (BI), Direct Brain Interface (DBI), Adaptive Brain Interface (ABI), and so on.

Table 1: The existing survey on brain signals in the last decade. The column 'Comprehensiveness' indicates whether the survey covers all subcategories of non-invasive brain signals or not. MI EEG refers to Motor Imagery EEG signals.

No.	Reference	Comprehensiveness	Signal	Deep Learning	Publication Time	Area
2	[14]	No	fMRI	Yes	2018	Mental Disease Diagnosis
3	[11]	Partial	EEG (MI EEG, P300)	No	2007	Classification
4	[5]	Partial	EEG (MI EEG, P300)	Partial	2018	Classification
5	[15]	Partial	EEG (ERD, P300, SSVEP, VEP, AEP)	No	2007	
6	[16]	No	MRI, CT	Partial	2017	Medical Image Analysis
7	[17]	No	EEG	Yes	2019	
8	[8]	No	EEG	No	2007	Signal Processing
9	[18]	Partial	EEG	No	2016	BCI Applications
10	[7]	Yes		No	2015	
11	[19]	No	EEG	Partial	2018	
12	[20]	No	EEG, fMRI	No	2015	Neurorehabilitation of Stroke
13	[21]	No	MI EEG	No	2015	
14	[22]	No	fMRI	No	2014	
15	[23]	No	ERP (P300)	No	2017	Applications of ERP"
16	[24]	No	fMRI	Yes	2018	Applications of fMRI
17	[25]	No	ERP	No	2017	Classification
18	[26]	Partial	EEG	No	2019	Brain Biometrics
19	[27]	Partial	EEG	No	2018	BCI Paradigms
20	Current Study	Yes	EEG and the subcategories, fNIRS, fMRI, MEG	Yes		

ested in the brain signal research; 2) the biomedical/medical/neuroscience experts who want to adopt deep learning techniques to solve problems in basic science.

To our best knowledge, this survey is the first comprehensive survey of the recent advances and frontiers of deep learning-based brain signal analysis. To this end, we have summarized over 200 contributions, most of which were published in the last five years. We make several key contributions in this survey:

- We review brain signals and deep learning techniques to help readers gain a comprehensive understanding of this area of research.

- We discuss the popular deep learning techniques and state-of-the-art models for brain signals, providing practical guidelines for choosing the suitable deep learning models given a specific subtype of signal.

- We review the applications of deep learning-based brain signal analysis and highlight some promising topics for future research.

The rest of this survey is structured as followed. Section 2 briefly introduces an taxonomy of brain signals in order to help the reader build a big picture in this field. Section 3 overviews the commonly used deep learning models to present the basic knowledge for researchers (*e.g.*, neurological and biomedical scholars) who are not familiar with deep learning. Section 4 presents the state-of-the-art deep learning techniques for brain signals and

Section 5 discusses the applications related to brain signals. Section 6 provides a detailed analysis and gives guidelines for choosing appropriate deep learning models based on the specific brain signal. Section 7 points out the opening challenges and future directions. Finally, Section 8 gives the concluding remarks.

2. Brain Imaging Techniques

In this section, we present a brief introduction of typical non-invasive brain imaging techniques. *More fundamental details about non-invasive brain signal (e.g., concepts, characteristics, advantages, and drawbacks) are provided in Appendix A.*

Figure 2 shows a taxonomy of non-invasive brain signals based on the signal collection method. Non-invasive signals divides into Electroencephalogram (EEG), Functional near-infrared spectroscopy (fNIRS), Functional magnetic resonance imaging (fMRI), and Magnetoencephalography (MEG) [34]. Table 2 summarizes the characteristics of various brain signals. In this survey, we mainly focus on EEG signals and its subcategories because they dominate the non-invasive signals. EEG monitors the voltage fluctuations generated by an electrical current within human neurons. The electrodes attached on scalp can measure various types of EEG signals, including spontaneous EEG [35] and evoked potentials (EP) [36]. Depending on the scenario, spontaneous EEG further diverges into sleep EEG,
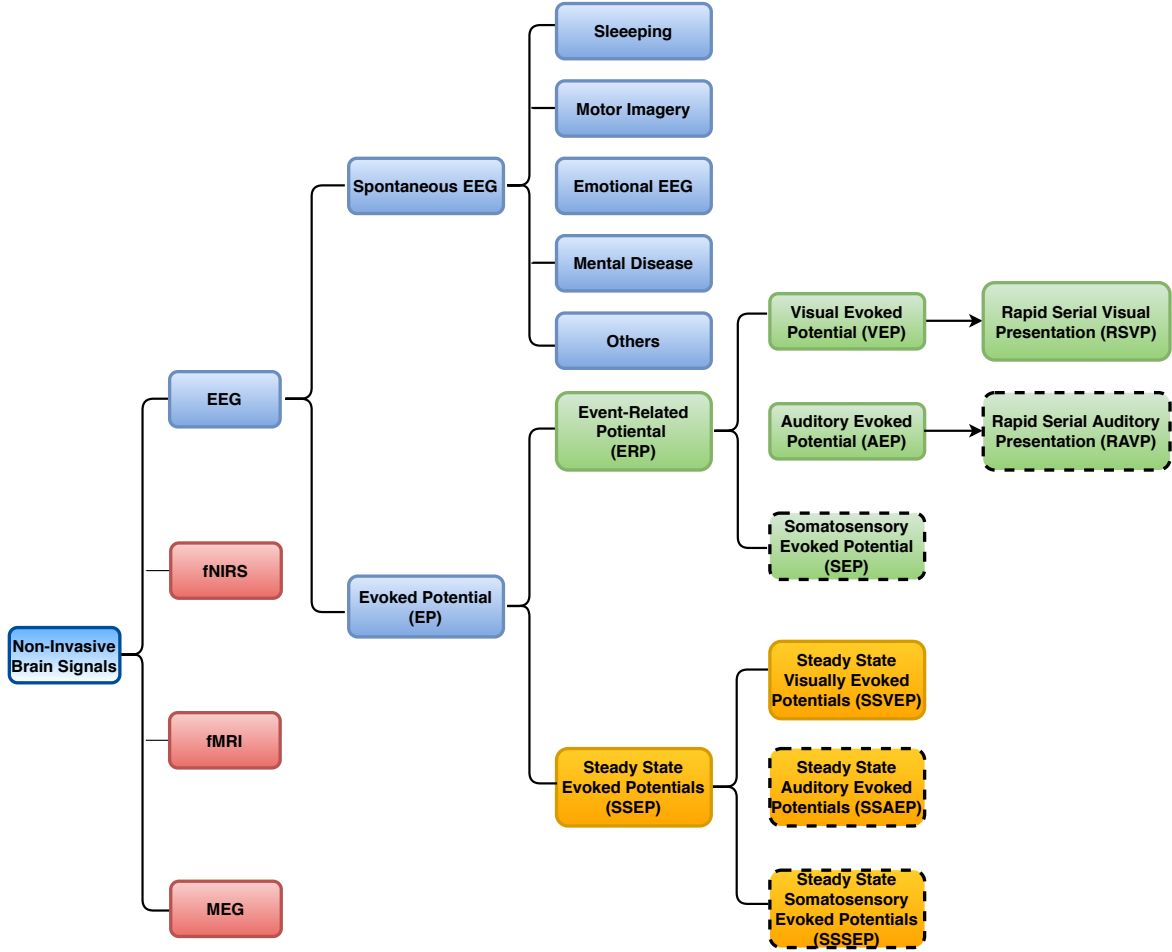
Figure 2: The taxonomy of non-invasive brain signals. The dashed quadrilaterals (RAVP, SEP, SSAEP, and SSSEP) are not included in this survey because there is no existing work focusing on them involving deep learning algorithms. P300, which is a positive potential recorded approximately 300 ms after the onset of presented stimuli, is not listed in this signal tree because it is included by ERP (which refers to all the potentials after the presented stimuli). In this classification, other brain imaging technique beyond EEG (*e.g.*, MEG and fNIRS) could also include visual/auditory tasks theoretically, but we omitted them since there is no existing work adopting deep learning on these tasks.

motor imagery EEG, emotional EEG, mental disease EEG, and others. Similarly, EP divides into event-related potentials (ERP) [28] and steady-state evoked potentials (SSEP) [37] according to the frequency of external stimuli. Each potential contains visual-, auditory-, and somatosensory-potentials based on the external stimuli types.

Regarding the other non-invasive techniques, fNIRS produces functional neuroimages by employing near-infrared (NIR) light to measure the aggregation degree of oxygenated hemoglobin (Hb) and deoxygenated-hemoglobin (deoxy-Hb), both of which have higher absorbers of light than other head components such as skull and scalp [38]; fMRI monitors brain activities by detecting the blood flow changes in brain areas [14]; MEG reflects brain activities via magnetic changes [39].

3. Overview on Deep Learning Models

In this section, we formally introduce the deep learning models including concepts, architectures, and techniques that are commonly used in the field of brain signal researches. Deep learning is a class of machine learning techniques that uses many layers of information-processing stages in hierarchical architectures for pattern classification and feature/representation learning [31]. *More detailed information about the deep learning techniques which are common-used in brain signal analysis can be find in Appendix B.*

Deep learning algorithms contain several subcategories based on the aim of the techniques (Figure 3):

- Discriminative deep learning models, which classify the input data into a pre-known label based on the adaptively learned discriminative features. Discrim-

Table 2: Summary of non-invasive brain signals' characteristics.

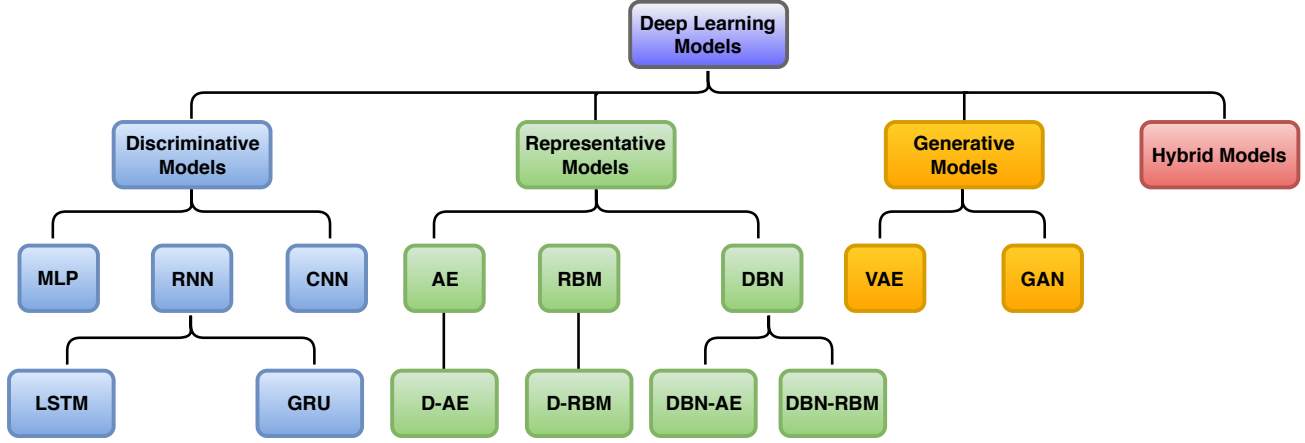| Signals	EEG	fNIRS	fMRI	MEG
Spatial resolution	Low	Intermediate	High	Intermediate
Temporal resolution	High	Low	Low	High
Signal-to-Noise Ratio	Low	Low	Intermediate	Low
Portability	High	High	Low	Low
Cost	Low	Low	High	High
Characteristic	Electrical	Metabolic	Metabolic	Magnetic

Figure 3: Deep learning models. They can be divided into discriminative, representative, generative and hybrid models based on the algorithm functions. Discriminative models (Appendix B.1) mainly include Multi-Layer Perceptron (MLP), Recurrent Neural Networks(RNN), and Convolutional Neural Networks (CNN). The two mainstreams of RNN are Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). Representative models (Appendix B.2) can be divided into Authoencoder (AE), Restricted Boltzmann Machine (RBM), and Deep Belief Networks (DBN). D-AE denotes Deep-Autoencoder which refers to the Autoencoder with multiple hidden layers. Likewise, D-RBM denotes Deep-Restricted Boltzmann Machine with multiple hidden layers. Deep Belief Network can be composed of AE or RBM, therefore, we divided DBN into DBN-AE and DBN-RBM. Generative models (Appendix B.3) that are commonly used in non-invasive brain signal analysis include Variational Autoencoder (VAE) and Generative Adversarial Networks (GAN).

Table 3: Summary of deep learning model types

Deep Learning	Input	Output	Function	Training method
Discriminative	Input data	Label	Feature extraction, Classification	Supervised
Representative	Input data	Representation	Feature extraction	Unsupervised
Generative	Input data	New Sample	Generation, Reconstruction	Unsupervised
Hybrid	Input data	–	–	–

inative algorithms are able to learn distinctive features by non-linear transformation, and classification through probabilistic prediction[3]. Thus these algorithms can play the role of both feature extraction and classification (corresponding to Figure 1). Discriminative architectures mainly include Multi-Layer

[3]The classification function is achieved by the combination of a softmax layer and one-hot label encoding. The one-hot label encoding refers to encoding the label by the one-hot method, which is a group of bits among which the only valid combinations of values are those with a single high (1) bit and all the others low (0) bits. For instance, a set of labels 0, 1, 2, 3 can be encoded as (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1).

Perceptron (MLP) [40], Recurrent Neural Networks (RNN) [41], Convolutional Neural Networks (CNN) [42], along with their variations.

- Representative deep learning models, which learn the pure and representative features from the input data. These algorithms only have the function of feature extraction (Figure 1) but cannot make classification. Commonly used deep learning algorithms for representation are Autoencoder (AE) [43], Restricted Boltzmann Machine (RBM) [44], Deep Belief Networks (DBN) [45], along with their variations.

- Generative deep learning models, which learn the

joint probability distribution of the input data and the target label. In the brain signal scope, generative algorithms are mostly used to generate a batch of brain signals samples to enhance the training set. Generative models commonly used in brain signal analysis include Variational Autoencoder (VAE)[4] [46], Generative Adversarial Networks (GANs) [47], etc.

- Hybrid deep learning models, which combine more than two deep learning models. For example, the typical hybrid deep learning model employs a representation algorithm for feature extraction and discriminative algorithms for classification.

The summary of the characteristics of each deep learning subcategories are listed in Table 3. Almost all the classification functions in neural networks are implemented by a softmax layer, which will not be regarded as an algorithmic component in this survey. For instance, a model combining a DBN and a softmax layer will still be regarded as a representative model instead of a hybrid model.

4. State-of-The-Art DL Techniques for Brain Signals

In this section, we thoroughly summarize the advanced studies on deep learning-based brain signals (Table 4). The hybrid models are divided into three parts: the combination of RNN and CNN, the combination of representative and discriminative models (denoted as 'Repre + Discri'), and others hybrid models.

4.1. EEG

Due to the advantages of high portability and low price, EEG signals have attracted much attention. Most of the latest publications on non-invasive brain signals are related to EEG. In this section, we summarize two aspects of EEG signals: spontaneous EEG and evoked potentials. As implied by the name, the former are spontaneous and the latter requires outside stimuli.

4.1.1. Spontaneous EEG We present the deep learning models for spontaneous EEG according to the application scenarios as follows.

(1) **Sleep EEG.** Sleep EEG is mainly used for recognizing the sleep stage and diagnosing sleep disorders or cultivating the healthy habit [48, 49]. According to Rechtschaffen and Kales (R&K) rules, the sleep stage includes wakefulness, non-REM (rapid eye movement) 1, non-REM 2, non-REM 3, non-REM 4, and REM. The American Academy of Sleep Medicine (AASM) recommends segmentation of sleep in five stages: wakefulness, non-REM 1, non-REM 2, slow wave sleep (SWS), and REM. The non-REM 3 and non-REM 4 are combined into

SWS since there is no clear distinction between them [49]. Generally, in the sleep stage analysis, the EEG signals are preprocessed by a filter which has various passband in different papers, but all notched at 50 Hz. The EEG signals are usually segmented into 30s windows.

(i) *Discriminative models.* CNN are frequently used for sleep stage classification on single-channel EEG [25, 50]. For example, Viamala et al. [51] manually extracted the time-frequency features and achieved a classification accuracy of 86%. Others used RNN [52] and LSTM [53] based on various features from the frequency domain, correlation, and graph theoretical features.

(ii) *Representative models.* Tan et al. [54] adopted a DBN-RBM algorithm to detect sleep spindle based on Power Spectral Density (PSD) features extracted from sleep EEG signals and achieved an F-1 of 92.78% on a local dataset. Zhang et al. [49] further combined DBN-RBM with three RBMs for sleep feature extraction.

(iii) *Hybrid models.* Manzano et al. [55] presented a multi-view algorithm in order to predict sleep stage by combining CNN and MLP. The CNN was employed to receive the raw time-domain EEG oscillations while the MLP received the spectrum singles processed by the Short-Time Fourier Transform (STFT) among 0.5-32 Hz. Fraiwan et al. [56] combined DBN with MLP for neonatal sleep state identification. Supratak et al. [57] proposed a model by combing a multi-view CNN and LSTM for automatic sleep stage scoring, in which the former was adopted to discover time-invariant dependencies while the latter (a bidirectional LSTM) was adopted the temporal features during the sleep. Dong et al. [58] proposed a hybrid deep learning model aiming at temporal sleep stage classification and took advantage of MLP for detecting hierarchical features along with LSTM for sequential information learning.

(2) **MI EEG.** Deep learning models have shown the superior on the classification of Motor-Imagery (MI) EEG and real-motor EEG [59, 60].

(i) *Discriminative models.* Such models mostly use CNN to recognize MI EEG [61]. Some are based on manually extracted features [62, 63]. For instance, Lee et al. [64] and Zhang et al. [65] employed CNN and 2-D CNN, respectively, for classification; Zhang et al. [65] learned affective information from EEG signals to built a modified LSTM control smart home appliances. Others also used CNN for feature extraction [66]. For example, Wang et al. [67] first used CNN to capture latent connections from MI-EEG signals and then applied weak classifiers to choose important features for the final classification; Hartmann et al. [59] investigated how CNN represented spectral features through the sequence of the MI EEG samples. MLP has also been applied for MI EEG recognition [68], which showed higher sensitivity to EEG phase features at earlier stages and higher sensitivity to EEG amplitude features at later stages.

---

[4]VAE is a variation of AE. However, they are working on different aspects. Therefore, we separately introduce AE and VAE.

Table 4: A summary of non-invasive brain signal studies based on deep learning models

brain signals				Deep Learning Models										
				Discriminative Models			Representative Models				Generative Models		Hybrid Models		
				MLP	RNN	CNN	AE (D-AE)	RBM (D-RBM)	DBN-AE	DBN-RBM	VAE	GAN	LSTM+CNN	Repre + Discri	Others
Non-invasive Signals	EEG	Spontaneous EEG	Sleep EEG	[69, 52]	[53, 52]	[51],[48], [25, 50], [70, 52]				[49, 54]			[57] [52]	[56]	[55], [58]
			MI EEG	[71],[68]	[6], [61, 65]	[64], [72], [60], [63],[73], [59, 62] [66]	[74, 75] [76]		[77]	[78, 79], [80]	[81]		[82, 10]	[4],[83]	[84, 85], [67, 86] [2]
			Emotional EEG	[87]	[88]	[89],[90], [91], [92, 93]	[94], [95]	[96, 97] [98]	[99]	[98],[99], [100],[101], [102, 103]			[104]	[105, 106] [107]	[108]
			Mental Disease EEG	[109]	[110],[111]	[112],[113], [114],[115], [116],[117], [118, 119] [120]	[121],[122], [123], [124]		[125]	[126, 127]			[128]	[129, 120], [130, 131]	
			Data Augmentation								[132, 81], [133] [134]				
			Others	[135, 136] [137]	[138]	[139], [140], [141], [142], [138],[143],[144], [145, 146] [147, 147] [148, 149]	[150],[151]		[152]	[153],[152], [154, 155]			[156, 147]	[157],[158, 159]	[160]
		EP	ERP VEP	[161, 162],	[163, 134]	[163],[73] [164, 147], [147, 165] [166]	[167]			[168, 169], [165]			[170],[171], [172]	[97, 173] [96]	
			ERP RSVP	[174, 175],		[176],[177], [178],[179], [180, 181], [182, 175] [183, 184] [185, 12]			[186]					[181, 175]	[12]
			ERP AEP			[187, 165], [166, 188]				[165]					
			SSEP SSVEP	[189]	[190]	[191], [192], [190, 193] [194]			[195]	[195]			[196]	[197]	
		fNIRS		[38],[198], [199],[71], [200]		[198]								[201]	
		fMRI		[202, 203]		[204],[63], [205],[206], [117],[207], [208, 194]	[209]		[210],	[211],[210],[212, 213]	[214, 215], [216, 203]			[217]	
		MEG				[218],[204]	[219]							[220]	

(ii) *Representative models.* DBN is widely used as a basis for MI EEG classification for its high representative ability [80, 79]. For example, Ren et al. [78] applied a convolutional DBN based on RBM components, showing better feature representation than hand-crafted features. Li et al. [77] processed EEG signals with discrete wavelet transformation and then applied a DBN-AE based on denoising AE. Other models include the combination of AE model (for feature extraction) and a KNN classifier [75], the combination of Genetic Algorithm (for hyper-parameter tuning) and MLP (for classification) [84], the combination AE and XGBoost for multi-person scenarios [76], and the combination of LSTM and reinforcement learning for multi-modality signal classification [85, 2].

(iii) *Hybrid models.* Several studies proposed hybrid models for the recognition of MI EEG [81]. For example, Tabar et al. [4] extracted high-level representations from the time, frequency domain and location information of EEG signals using CNN and then used a DBN-AE with seven AEs as the classifier; Tan et al. [82] used a denoising AE for dimensional reduction, a multi-view CNN combined with RNN for discovering latent temporal and spatial information, and finally achieved an average accuracy of 72.22% on a public dataset.

(3) **Emotional EEG.** The emotion of an individual can be evaluated in three aspects: valence, arousal, and dominance. The combination of the three aspects form emotions such as fear, sadness, and anger, which can be revealed by EEG signals.

(i) *Discriminative models.* MLP are traditionally used [137, 87] while CNN and RNN are increasingly popular in EEG based emotion prediction [89, 90]. Typical CNN-based work in this category includes hierarchical CNN [89, 92] and augmenting the training set for CNN [91]. Li et al. [89] were the first to propose capturing the spatial dependencies among EEG channels via converting multi-channel EEG signals into a 2-D matrix. Besides, Talathi [110] used a discriminative deep learning model composed of GRU cells. Zhang et al. [88] proposed a spatial-temporal recurrent neural network, which employs a multi-directional RNN layer to discover long-range contextual cues and a bi-directional RNN layer to capture sequential features produced by the previous spatial RNN.

(ii) *Representative models.* DBN, especially DBN-RBM, is widely used for the unsupervised representation ability in emotion recognition [100, 106, 103]. For instance, Xu et al. [99, 101] proposed a DBN-RBM algorithm with three RBMs and an RBM-AE to predict affective state; Zhao et al. [126] and Zheng et al. [102] cobmined DBN-RBM with SVM and Hidden Markov Model (HMM), respectively, addressing the same problem; Zheng et al. [96, 97] introduced a D-RBM with five hidden RBM layers to search the important frequency patterns and informative channels in affection recognition; Jia et al. [98] eliminated channels with high errors and then used D-RBM for

affective state recognition based on representative features of the residual channels.

The emotion is affected by many subjective and environmental factors (*e.g.*, gender and fatigue). Yan et al. [95] investigated the discrepancy of emotional patterns between men and women by proposing a novel model called Bimodal Deep AutoEncoder (BDAE) which received both EEG and eye movement features and shared the information in a fusion layer which connected with an SVM classifier. The results showed that the females have higher EEG signal diversity on the fearful emotion while males on sad emotion. Moreover, for women, the inter-subject differences in fear is more significant then other emotions [95]. To overcome the mismatched distribution among the samples collected from different subjects or different experimental sessions, Chai et al. [94] proposed an unsupervised domain adaptation technology which is called subspace alignment autoencoder (SAAE) by combing an AE and a subspace alignment solution. The proposed approach obtained a mean accuracy of 77.88% in person independent scenario.

(iii) *Hybrid models.* One common-used hybrid model is a combination of RNN and MLP. For example, Alhagry et al. [108] employed an LSTM architecture for feature extraction from emotional EEG signals and the features are forwarded into an MLP for classification. Furthermore, Yin et al. [107] proposed a multi-view ensemble classifier to recognize individual emotions using multimodal physiological signals. The ensemble classifier contains several D-AEs with three hidden layers and a fusion structure. Each D-AE receives one physiological signal (*e.g.*, EEG) and then sends the outputs of D-AE to a fusion structure which is composed of another D-AE. At last, an MLP classifier makes the prediction based on the mixed features. Kawde et al. [105] implemented an affect recognition system by combining a DBN-RBM for effective feature extraction and an MLP for classification.

(4) **Mental Disease EEG.** A large number of researchers exploited EEG signals to diagnose neurological disorders, especially epileptic seizure [109].

(i) *Discriminative models.* The CNN is widely used in the automatic detection of epileptic seizure [112, 114, 116, 93]. For example, Johansen et al. [118] adopted CNN to work on the high-passed (1 Hz) EEG signals of epileptic spike and achieved an AUC of 94.7%. Acharya et al. [113] employed a CNN model with 13 layers on depression detection, which was evaluated on a local dataset with 30 subjects and achieved the accuracies of 93.5% and 96.0% based on the left- and right- hemisphere EEG signals, respectively. Morabito et al. [115] tried to exploit a CNN structure to extract suitable features of multi-channel EEG signals to classify Alzheimer's Disease from the patients with Mild Cognitive Impairment and healthy control group. The EEG signals are filtered in bandpass (0.1 ∼ 30 Hz) and achieved an accuracy of around 82% for

three-class classification. Eapid Eye Movement Behavior Disorder (RBD) may cause many mental disorder diseases like Parkinson's disease (PD). Ruffini et al. [111] described an Echo State Networks (ESNs) model, a particular class of RNN, to distinguish RBD from healthy individuals. In some research, the discriminative model is only employed for feature extraction. For example, Ansari et al. [119] used CNN to extract the latent features and fed into a Random Forest classifier for the final seizure detection of neonatal babies. Chu et al. [149] combined CNN and a traditional classifier for schizophrenia recognition.

(ii) *Representative models.* For disease detection, one commonly used method is adopting a representative model (*e.g.*, DBN) followed by a softmax layer for classification [127]. Page et al. [125] adopted DBN-AE to extract informative features from seizure EEG signals. The extracted features were fed into a traditional logistic regression classifier for seizure detection. Al et al. [131] proposed a multi-view DBN-RBM structure to analyze EEG signals from depression patients. The proposed approach contains multiple input pathways, composed of two RBMs, while each corresponded to one EEG channel. All the input pathways would merge into a shared structure which is composed of another RBMs. Some papers would like to preprocess the EEG signals through dimensionality reduction methods such as PCA [129] while others prefer to directly fed the raw signals to the representative model [122]. Lin et al. [122] proposed a sparse D-AE with three hidden layers to extract the representative features from epileptic EEG signals while Hosseini et al. [129] adopted a similar sparse D-AE with two hidden layers.

(iii) *Hybrid models.* A popular hybrid method is a combination of RNN and CNN. Shah et al. [128] investigated the performance of CNN-LSTM on seizure detection after channel selection and the sensitivities range from 33% to 37% while false alarms ranges from 38% to 50%. Golmohammadi et al. [130] proposed a hybrid architecture for automatic interpretation of EEG by integrating both the temporal and spatial information. 2D and 1D CNNs capture the spatial features while LSTM networks capture the temporal features. The authors claimed a sensitivity of 30.83% and a specificity of 96.86% on the well-known TUH EEG seizure corpus. In the detection of early-stage Creutzfeldt-Jakob Disease (SJD), Morabito et al. [123] combined D-AE and MLP together. The EEG signals of SJD were first filtered by bandpass (0.5∼70 Hz) and then fed into a D-AE with two hidden layers for feature representation. At last, the MLP classifier obtained the accuracy of 81∼ 83% in a local dataset. Convolutional autoencoder, replacing the fully-connected layers in a standard AE by convolutional and de-convolutional layers, is applied to extract the seizure features in an unsupervised manner [124].

(5) **Data augmentation.** The generative models such as GAN could be used for data augmentation in brain

signal classification [132]. Palazzo et al. [133] first demonstrated that the information contained in brainwaves are empowered to distinguish the visual object and then extracted more robust and distinguishable representations of EEG data using RNN. At last, they employed the GAN paradigm to train an image generator conditioned by the learned EEG representations, which could convert the EEG signals into images [133]. Kavasidis et al. [134] aiming at converting EEG signals into images. The EEG signals were collected when the subjects were observing images on a screen. An LSTM layer was employed to extract the latent features from the EEG signals, and the extracted features were regarded as the input of a GAN structure. The generator and the discriminator of the GAN were both composed of convolutional layers. The generator was supposed to generate an image based on the input EEG signals after the pre-training. Abdelfattach et al. [132] adopted a GAN on seizure data augmentation. The generator and discriminator are both composed of fully-connected layers. The authors demonstrated that GAN outperforms other generative models such as AE and VAE. After the augmentation, the classification accuracy increased dramatically from 48% to 82%.

(6) **Others.** Some researches have explored a wide range of exciting topics. The first one is how EEG signals are affected by audio/visual stimuli. This differs from the potentials evoked by audio/visual stimulations because the stimuli in this phenomenon always exist instead of flicking in a particular frequency. Stober et al. [188, 142] claimed that the rhythm-evoked EEG signals are informative enough to distinguish the rhythm stimuli. The authors conducted an experiment where 13 participants were stimulated by 23 rhythmic stimuli, including 12 East African and 12 Western stimuli. For the 24-category classification, the proposed CNN achieved a mean accuracy of 24.4%. After that, the authors exploited convolutional AE for representation learning and CNN for recognition and achieved an accuracy of 27% for 12-class classification [157]. Sternin et al. [148] adopted CNN to capture discriminative features from the EEG oscillations to distinguish whether the subject was listening or imaging music. Similarly, Sarkar et al. [165] designed two deep learning models to recognize the EEG signals aroused by audio or visual stimuli. For this binary classification task, the proposed CNN and DBN-RBM with three RBMs achieved the accuracy of 91.63% and 91.75%, respectively. Furthermore, the spontaneous EEG could be used to distinguish the user's mental state (logical versus emotional) [172].

Moreover, some researchers focus on the impact on EEG of cognitive load [138] or physical workload [221]. Bashivan et al. [159] first extract informative features through wavelet entropy and band-specific power, which would be fed into a DBN-RBM for further refining. At last, an MLP is employed for cognitive load level recognition. The authors, in another work [171], also denoted to find the

general features which are constant in inter-/intra- subjects scenarios under various mental load. Yin et al. [150] collected the EEG signals from different mental workload levels (*e.g.*, high and low) for binary classification. The EEG signals are filtered by a low-pass filter, transformed to the frequency domain and be calculated the power spectral density (PSD). The extracted PSD features were fed into a denoising D-AE structure for future refining. They finally got an accuracy of 95.48%. Li et al. [155] worked on the recognition of mental fatigue level, including alert, slight fatigue, and severe fatigue.

In addition, EEG based driver fatigue detection is an attractive area [158, 151, 147]. Huang et al. [140] designed a 3D CNN to predict the reaction time in drowsiness driving. This is meaningful to reduce traffic accident. Hajinoroozi et al. [153] adopted a DBN-RBM to handle the EEG signals which were processed by ICA. They achieved an accuracy of around 85% in binary classification ('drowsy' or 'alert'). The strength of this paper is that it evaluated the DBN-RBM on three levels: time samples, channel epochs, and windowed samples. The experiments illustrated that the channel epoch level outperformed the other two levels. San et al. [154] combined deep learning models with a traditional classifier to detect driver fatigue. The model contains a DBN-RBM structure followed by an SVM classifier, which achieved the detection accuracy of 73.29%. Almogbel et al. [145] investigated the drivers' mental state under different low workload levels. A proposed CNN is claimed to detect the driving workload directly based on the raw EEG signals.

The research of the detection of eye state has shown exceeding accuracy. Narejo et al. [152] explored the detection of eye state (closed or open) based on EEG signals. They tried a DBN-RBM with three RBMs and a DBN-AE with three AEs and achieved a high accuracy of 98.9%. Reddy et al. [136] tried a simpler structure, MLP, and got a slightly lower accuracy of 97.5%.

Furthermore, to make this survey more complete, we provide a brief introduction of Event-related desynchronization/synchronization (ERD/ERS). ERD/ERS refers to the phenomena that the magnitude and frequency distribution of the EEG signal power changes during a specific brain state [36]. In particular, ERD denotes the power decrease of ongoing EEG signals while ERS represents the power increase of EEG signals. This characteristic of ERD/ERS of brain signals can be used to detect the event which caused the EEG fluctuation. For example, [222] presents the ERD/ERS phenomena in motor cortex recorded during a motor-imagery task.

ERD/ERS mainly appears in sensory, cognitive and motor procedures, which is not widely used in brain research due to the drawbacks like unstable accuracy cross subjects [36]. In most of the situations, the ERD/ERS is regarded as a specific feature of EEG powers for further analysis [81, 4]. The task causes an ERD in the mu band (8-13 Hz) of EEG and an ERS in the beta band (13-30 Hz). In particular, the ERD/ERS were calculated as relative changes in power concerning baseline: $ERD/ERS = (P_e - P_b)/P_b$, where P_e denotes the signals power over one-second segment when the event occurring and P_b denotes the signal power in a one-second segment during baseline which is before the event [71]. Generally, the baseline refers to the rest state. For example, Sakhavi et al. calculated the ERD/ERS map and analyzed the different patterns among different tasks. The analysis demonstrated that the dynamic of energy should be considered because the static energy does not contains enough information [86].

There are several overlooked yet promising areas. Baltatzis et al. [141] adopted CNN to detect school bullying through the EEG when watching the specific video. They achieved 93.7% and 88.58% for binary and four-class classification. Khurana et al. [223] proposed deep dictionary learning that outperformed several deep learning methods. Volker et al. [143] evaluated the use of Deep CNN in flanker task, which achieved an averaging accuracy of 84.1% on the seen subject and 81.7 on the unseen subject. Zhang et al. [160] combined CNN and graph network to discover the latent information from the EEG signal.

Miranda-Correa et al. [104] proposed a cascaded framework by combing RNN and CNN to predict individuals' affective level and personal factors (Big-five personality traits, mood, and social context). An experiment conducted by Putten et al. [146] attempted to identify the user's gender based on their EEG signals. They employed a standard CNN algorithm and achieved the binary classification accuracy of 81% over a local dataset. The detection of emergency braking intention could help to reduce the responses time. Hernandez et al. [144] demonstrated that the driver's EEG signals could distinguish braking intention and normal driving state. They combined a CNN algorithm which achieved the accuracy of 71.8% in binary classification. Behncke et al. [139] applied deep learning, a CNN model, in the context of robot assistive devices. They attempted to use CNN to improve the accuracy of decoding robot errors from EEG while the subject was watching the robot both during an object grasping and a pouring task.

Teo et al. [135] tried to combine the brain signal and recommender system, which predicted the user's preference by EEG signals. There were sixteen participants took the experiments which collected the EEG signals when the subject was presented 60 bracelet-like objects as rotating visual stimuli (a 3D object). Then, an MLP algorithm was adopted to classify the user like or dislike the object. This exploration got the prediction accuracy of 63.99%. Some researchers have tried to explore a common framework which can be used for various brain signal paradigms. Lawhern et al. [73] introduced EEGNet based on a compact CNN and evaluated its robustness in various brain signal

contexts [73].

4.1.2. Evoked Potential Next, we introduce the latest researches on evoked potentials including ERP and SSEP.

(1) **ERP.** In most situations, the ERP signals are analyzed through P300 phenomena. Meanwhile, almost all the studies on P300 are based on the scenario of ERP. Therefore, in this section, a majority of the P300 related publications are introduced in the subsection of VEP/AEP according to the scenario.

(i) *VEP.* VEP is one of the most popular subcategories of ERP [23, 224, 163]. Ma et al. [225] worked on motion-onset VEP (mVEP) by extracting representative features through deep learning and adopted genetic algorithm combined with a multi-level sensing structure to compress the raw signals. The compressed signals were sent to a DBN-RBM algorithm to capture the more abstract high-level features. Maddula et al. [170] filtered the P300 signals with visual stimuli by a bandpass filter ($2 \sim 35$ Hz) and then fed into a proposed hybrid deep learning model for further analysis. The model includes a 2D CNN structure to capture the spatial features followed by an LSTM layer for temporal feature extraction. Liu et al. [168] combined a DBN-RBM representative model with an SVM classifier for concealed information test and achieved a high accuracy of 97.3% over a local dataset. Gao et al. [167] employed an AE model for feature extraction followed by an SVM classifier. In the experiment, each segment contains 150 points, which were divided into five time-steps, and each step had 30 points. This model achieved an accuracy of 88.1% over a local dataset. A wide range of P300 related studies is based on P300 speller [173], which allows the user to write characters. Cecotti et al. [177] tried to increase the P300 detection accuracy for more precise word-spelling. A new model was presented based on CNN, which including five low-level CNN classifiers with the different feature set, and the final high-level results are voted by the low-level classifiers. The highest accuracy reached 95.5% over the dataset II from the third BCI competition. Liu et al. [164] proposed a Batch Normalized Neural Network (BN^3) which is a variant of CNN in P300 speller. The proposed method consists of six layers, and the batch normalization was operated in each batch. Kawasaki et al. [162] employed an MLP model to detect P300 segments from non-P300 segments and achieved the accuracy of 90.8%.

(ii) *AEP.* A few works focused on the recognition of AEP. For example, Carabez et al. [187] proposed and tested 18 CNN structures to classify single-trial AEP signals. In the experiment, the volunteers were required to wear on an earphone which produces auditory stimulus designed based on the oddball paradigm. The experimental analysis demonstrated that the CNN frameworks, regardless of the number of convolutional layers, were effective to extract the temporal and spatial features and provided competitive

results. The AEP signals are filtered by $0.1 \sim 8$ Hz and downsampled from 256 Hz to 25 Hz. The experimental results showed that the downsampled data work better.

(iii) *RSVP.* Among various VEP diagrams, RSVP has attracted much attention [183]. In the analysis of RSVP, a number of discriminative deep learning models (e,g., CNN [177, 178, 182] and MLP [174]) has achieved a big success. A common preprocessing method used in RSVP signals is frequency filtering. The pass bands are generally ranged from $0.1 \sim 50$ Hz [176, 185]. Cecotti et al. [12] worked on the classification of ERP signals in RSVP scenario and proposed a modified CNN model for the detection of the specific target in RSVP. In the experiment, the images of faces and cars were regarded as target or non-target, respectively. The image presenting frequency is 2 Hz. In each session, the target probability was 10%. The proposed model offered an AUC of 86.1%. Hajinoroozi et al. [179] adopted a CNN model targeting the inter-subject and inter-task detection of RSVP. The experimental results showed that CNN worked good in cross-task but failed to get satisfying performance in the cross-subject scenario. Mao et al. [175] compared three different deep neural network algorithms in the prediction of whether the subject had seen the target or not. The MLP, CNN, and DBN models obtained the AUC of 81.7%, 79.6%, and 81.6%, respectively. The author also applied a CNN model to analyze the RSVP signals for person identification [180].

The representative deep learning models are also applied in RSVP. Vareka et al. [186] verified if deep learning performs well for single trial P300 classification. They conducted an RSVP experiment while the subjects were asked to recognize the target from non-target and distracters. Then a DBN-AE was implemented and compared with some non-deep learning algorithms. The DBN-AE was composed of five AEs while the hidden layer of the last AE only has two nodes which can be used for classification through softmax function. Finally, the proposed model achieved the accuracy of 69.2%. Manor et al. [181] applied two deep neural networks to deal with the RSVP signals after lowpass filtering ($0 \sim 51$ Hz). Discriminative CNN achieved the accuracy of 85.06%. Meanwhile, the representative convolutional D-AE achieved the accuracy of 80.68%.

(2) **SSEP.** Most of deep learning-based studies in SSEP area focus on SSVEP like [191]. SSVEP refers to brain oscillations evoked by the flickering visual stimuli, which generally produced from the parietal and occipital regions [192]. Attia et al. [196] aimed at finding an intermediate representation of SSVEP. A hybrid method combined CNN and RNN was proposed to capture the meaningful features from the time domain directly, which achieved the accuracy of 93.59%. Waytowich et al. [192] applied a compact CNN model to directly work on the raw SSVEP signals without any hand-crafted features. The reported cross subject mean accuracy was approximately

80%. Thomas et al. [190] first filter the raw SSVEP signals through a bandpass filter (5 $\sim$ 48 Hz) and then operated discrete FFT on consecutive 512 points. The processed data were classified by a CNN (69.03%) and an LSTM (66.89%) independently.

Perez et al. [197] adopted a representative model, a sparse AE, to extract the distinct features from the SSVEP from multi-frequency visual stimuli. The proposed model employed a softmax layer for the final classification and achieved the accuracy of 97.78%. Kulasingham et al. [195] classified SSVEP signals in the context of guilty knowledge test. The authors applied DBN-RBM and DBN-AE independently and achieved the accuracy of 86.9% and 86.01%, respectively. Hachem et al. [189] investigated the influence of fatigue on SSVEP through an MLP model during wheelchair navigation. The goal of this study was to seek the key parameters to switch between manual, semi-autonomous, and autonomous wheelchair command. Aznan et al. [193] explored the SSVEP classification, where the signals were collected through dry electrodes. The dry signals were more challenging for the lower SNR than standard EEG signals. This study applied a CNN discriminative model and achieved the highest accuracy of 96% over a local dataset.

4.2. fNIRS

Up to now, only a few of researchers paid attention on deep learning-based fNIRS. Naseer et al. [38] analyzed the difference between two mental tasks (mental arithmetic and rest) based on fNIRS signals. The authors manually extracted six features from the prefrontal cortex fNIRS and compared six different classifiers. The results demonstrated that the MLP with the accuracy of 96.3% outperformed all the traditional classifiers, including SVM, KNN, naive Bayes, etc. Huve et al. [198] classified the fNIRS signals, which were collected from the subjects during three mental states, including substrations, word generation, and rest. The employed MLP model achieved the accuracy of 66.48% based on the hand-crafted features (*e.g.*, the concentration of OxyHb/DeoxyHb). After that, the authors study the mobile robot control through fNIRS signals and got the binary classification accuracy of 82% (offline) and 66% (online) [199]. Chiarelli et al. [71] exploited the combination of fNIRS and EEG for left/right MI EEG classification. Sixteen features extracted from fNIRS signals (eight from OxyHb and eight from DeoxyHb) were fed into an MLP classifier with four hidden layers.

On the other hand, Hiroyasu et al. [201] attempted to detect the gender of the subject through their fNIRS signals. The authors employed a denoising D-AE with three hidden layers to extract distinctive features to be fed into an MLP classifier for gender detection. The model was evaluated over a local dataset and gained the average accuracy of 81%. In this study, the authors also pointed out that, compared with Positron Emission Tomography (PET) and fMRI, fNIRS has higher time resolution and more affordable [201].

4.3. fMRI

Recently, several deep learning methods have been applied to fMRI analysis, especially on the diagnosis of cognitive impairment [14, 33].

(1) **Discriminative models.** Among the discriminative models, CNN is a promising model to analyze fMRI [206]. For example, Havaei et al. built a segmentation approach for brain tumor based on fMRI with a novel CNN algorithm which can capture both the global features and the local features simultaneously [205]. The convolutional filters have different size. Thus, the small-size and large-size filter could exploit the local and global features, independently. Sarraf et al. [226, 207] applied deep CNN to recognize Alzheimer's Disease based on fMRI and MRI data. Morenolopez et al. [227] employed a CNN model to deal with fMRI of brain tumor patients for three-class recognition (normal, edema, or active tumor). The model was evaluated over BRATS dataset and obtained the F1 score of 88%. Hosseini et al. [117] employed CNN for feature extraction. The extracted features were classified by SVM for the detection of an epileptic seizure.

Furthermore, Li et al. proposed a data completion method based on CNN. In particular, utilizing the information from fMRI data to complete PET, then train the classifier based on both fMRI and PET [208]. In the model, the input data of the proposed CNN is the fMRI patch, and the output is a PET patch. There are two convolutional layers with ten filters mapping the fMRI to PET. The experiments illustrated that the classifier trained by the combination of fMRI and PET (92.87%) outperformed the one trained by solo fMRI (91.92%) Moreover, Koyamada et al. used a nonlinear MLP to extract common features from different subjects. The model is evaluated over a dataset from the Human Connectome Project (HCP) [202].

(2) **Representative models.** A wide range of publications demonstrated the effectiveness of representative models in recognition of fMRI data [213]. Hu et al. [217] used demonstrated that deep learning outperforms other machine learning methods in the diagnosis of neurological disorders such as Alzheimer's disease. Firstly, the fMRI images were converted to a matrix to represent the activity of 90 brain regions. Secondly, a correlation matrix is obtained by calculating the correlation between each pair of brain regions to represent the functional connectivity between different brain regions. Furthermore, a targeted AE is built to classify the correlation matrix, which is sensitive to AD. The proposed approach achieved an accuracy of 87.5%. Plis et al. [211] employed a DBN-RBM with three RBM components to extract the distinctive features from ICA processed fMRI and finally achieved an average F1 measure of above 90% over four public datasets. Suk et al. compared the

effectiveness of DBN-RBM and DBN-AE on Alzheimer's disease detection and the experimental results showed that the former obtained the accuracy of 95.4%, which is slightly lower than the latter (97.9%) [210]. Suk et al. [209] applied a D-AE model to extract latent features from the resting-state fMRI data on the diagnosis of Mild Cognitive Impairment (MCI). The latent features are fed into a SVM classifier which achieved the accuracy of 72.58%. Ortiz et al. [212] proposed a multi-view DBN-RBM to receives the information of MRI and PET simultaneously. The learned representations were sent to several simple SVM classifiers which were ensembled to form a high-level stronger classifier by voting.

(3) **Generative models.** The reconstruction of natural image (*e.g.*, fMRI) has been attracted lots of attention [215, 88, 203]. Seeliger et al. [214] proposed a deep convolutional GAN for reconstructing visual stimuli from fMRI, which aimed at training a generator to create an image similar to the visual stimuli. The generator contains four convolutional layers in order to convert the input fMRI to a natural image. Han et al. [215] focused on the generation of synthetic multi-sequence fMRI using GAN. The generated image can be used for data augmentation for better diagnostic accuracy or physician training to help better understand various diseases. The authors applied the existing Deep Convolutional GAN (DCGAN) [228] and Wasserstein GAN (WGAN) [229] and found that the former works better. Shen et al. [203] presented another image recovery approach by minimizing the distance between the real image and the image generated based on real fMRI.

4.4. MEG

Garg et al. [218] worked on the refining of MEG signals by removing the artifacts like eye-blinks and cardiac activity. The MEG singles were decomposed by ICA first and then classified by a 1-D CNN model. At last, the proposed approach achieved the sensitivity of 85% and specificity of 97% over a local dataset. Hasasneh et al. [220] also focused on artifacts detection (cardiac and ocular artifacts). The proposed approach uses CNN to capture temporal features and MLP to extract spatial information. Shu et al. [219] employed a sparse AE to learn the latent dependencies of MEG signals in the task of single word decoding. The results demonstrated that the proposed approach is advantageous for some subjects, although it did not produce an overall increase in decoding accuracy. Cichy et al. [204] applied a CNN model to recognize visual object based on MEG and fMRI signals.

5. Brain Signal-based Applications

Deep learning models have contributed to various of brain signal applications as summarized in Table 5. The papers focused on signal classification without application background are not listed in this table. Therefore, the publication amounts in this table are less than in Table 4.

5.1. Health Care

In the health care area, the deep learning-based brain signal systems mainly works on the detection and diagnosis of mental diseases such as sleep disorders, Alzheimer's Disease, epileptic seizure, and other disorders. In the first place, for the sleep disorder detection, most studies are focused on the sleep stage detection based on sleep spontaneous EEG. In this situation, the researchers do not need to recruit patients with sleep disorder because the sleep EEG signals can be easily collected from healthy individuals. In terms of the algorithm, it can be observed from Table 5 that the DBN-RBM and CNN are widely adopted for feature selection and classification. Ruffini et al. [111] walk one step further by detecting the REM Behavior Disorder (RBD), which may cause neurodegenerative diseases such as Parkinson's disease. They achieved an average accuracy of 85% in recognition of the RBD from healthy controls.

Moreover, fMRI is widely applied in the diagnosis of Alzheimer's Disease. By taking advantage of the high spatial resolution of fMRI, the diagnosis achieved the accuracy of above 90% in several studies. Another reason that contributes to competitive performance is the binary classification scenario. Apart from that, there are several publications diagnose the AD based on spontaneous EEG [115, 126].

Besides, the diagnosis of epileptic seizure attracted much attention. The seizure detection mainly based on spontaneous EEG. The popular deep learning models in this scenario contain the independent CNN and RNN, along with hybrid models combined RNN and CNN. Some models integrated the deep learning models for feature extraction and traditional classifier for detection [127, 125]. For example, Yuan et al. [121] applied a D-AE in feature extraction followed by SVM for seizure diagnosis. Ullah et al. [112] adopted the voting for post-processing, which proposed several different CNN classifiers and predicted the final result by voting.

Furthermore, there are a lot of other healthcare issues can be solved by brain signal research. The cardiac artifacts in MEG can be automatically detected by deep learning models[218, 220]. Several modified CNN structures are proposed to detect brain tumor based on fMRI from the public BRATS dataset [205, 206]. Researchers have demonstrated the effectiveness of deep learning models in the detection of a wide number of mental diseases such as depression [113], Interictal Epileptic Discharge (IED) [230], schizophrenia [211], Creutzfeldt-Jakob Disease (CJD) [123], and Mild Cognitive Impairment (MCI) [209].
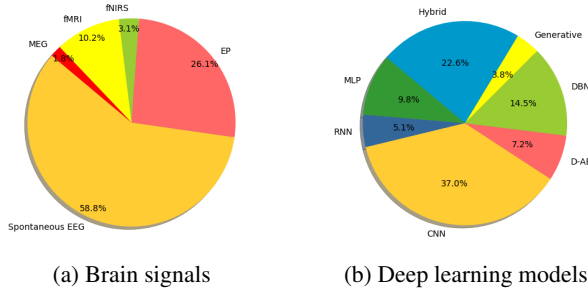
(a) Brain signals   (b) Deep learning models

Figure 4: Illustration of the publications proportion for crucial brain signals and deep learning models.

5.2. Smart Environment

The smart environment is a promising application scenario for brain signals in the future. With the development of Internet of Things (IoT), an increasing number smart environment can be connected to brain signals. For example, the assisting robot can be used in smart home [65, 2], in which the robot can be controlled by brain signals of the individuals. Moreover, Behncke et al. [139] and Huve et al. [199] investigated the robot control problem based on the visual stimulated spontaneous EEG and fNIRS signals. The brain signal controlled exoskeleton could help the disabilities who damaged the motor system in sub-limb in walking and daily activities [191]. In the future, the research on brain-controlled appliances may be beneficial to the elders or disabilities in smart home and smart hospital.

5.3. Communication

One of the biggest advantages of brain signals, compared to other human-machine interface techniques, is that brain signal enables the patient who lost most motor abilities like speaking to communicate with the outer world. The deep learning technology improved the efficiency of brain signal based communications. One typical diagram which enables individual typing without any motor system is P300 speller, which can convert the user's intent into text [162]. The powerful deep learning models empower the brain signal systems to recognize the P300 segment from the non-P300 segment while the former contains the communication information of the user [166]. In a higher level, the representative deep learning models can help to detect what character the user is focusing on and print it on the screen to chat with others [166, 170, 164]. Additionally, Zhang et al. [10] proposed a hybrid model that combined RNN, CNN, and AE to extract the informative features from MI EEG to recognize what letter the user wants to speak.

5.4. Security

Brain signals can be used in security scenarios such as identification (or recognition) and authentication (or verification). The former conducts multi-class classification

to recognize a person's identity [6]. The latter conducts binary classification to decide whether a person is authorized [61].

The majority of the existing biometric identification/authentication systems rely on individuals' intrinsic physiological features such as face, iris, retina, voice, and fingerprint [6]. They are vulnerable to various attacks based on anti-surveillance prosthetic masks, contact lenses, vocoder, and fingerprint films. EEG-based biometric person identification is a promising alternative given its highly resilient to spoofing attacks—individual's EEG signals are virtually impossible for an imposter to mimic. Koike et al. [161] have adopted deep neural networks to identify the user's ID based on the VEP signals; Mao et al. [180] applied CNN for person identification based on RSVP signals; Zhang et al. [6] proposed an attention-based LSTM model and evaluated it over both public and local datasets. EEG signals are also combined with gait information in a hybrid deep learning model for a dual-authentication system [61].

5.5. Affective Computing

Affective states of a user provide critical information for many applications such as personalized information (*e.g.*, multimedia content) retrieval or intelligent human-computer interface design [99]. Recent research illustrated that deep learning models can enhance the performance in affective computing. The most widely used circumplex model believe the emotions are distributed in two dimensions: arousal and valence. The arousal refers to the intensity of the emotional stimuli or how strong is the emotion. The valence refers to the relationship within the person who experiences the emotion. In some other models, the dominance and liking dimensions are deployed.

Some research [89, 90, 91] attempts to classify users' emotional state into two (positive/negative) or three categories (positive, neutral, and negative) based on EEG signals using deep learning algorithms such as CNN and its variants [87]. DBN-RBM is the most representative deep learning model to discover the concealed features from emotional spontaneous EEG [99, 96]. Xu et al. [99] applied DBN-RBM as feature extractors to classify affective states based on EEG.

Further, some researchers aim to recognize the positive/negative state of each specific emotional dimension. For example, Yin et al. [107] employed an ensemble classifier of AE in order to recognize the user's affection. Each AE uses three hidden layers to filter out noises and to derive stable physiological feature representations. The proposed model was evaluated over the benchmark, DEAP, and achieved the arousal of 77.19% and valence of 76.17%.

5.6. Driver Fatigue Detection

Vehicle drivers' ability to keep alert and maintain optimal performance will dramatically affect the traffic safety [145].

Table 5: Summary of deep learning-based brain signal applications. The 'local' dataset refers to private or not available dataset. The public datasets (along with download links) will be introduced in Section 5.9. In the signals, S-EEG, MD EEG, and E-EEG separately denote sleep EEG, mental disease EEG, and emotional EEG. The single 'EEG' refers to the other subcategory of spontaneous EEG. In the models, RF and LR denote to random forest and logistic regression algorithms, respectively. In the performance column, 'N/A', 'sen', 'spe', 'aro', 'val', 'dom', and 'like' denote not-found, sensitivity, specificity, arousal, valence, dominance, and liking, respectively. For each application scenario, the literature are sorted out by signal types and deep learning models.

Brain Signal Applications		Reference	Signals	Deep Learning Models	Dataset	Performance
Health Care	**Sleep Quality Evaluation**	Shahin et al. [69]	S-EEG	MLP	University Hospital in Berlin	0.9
		Biswai et al. [52]	S-EEG	RNN	Local	0.8576
		Ruffini et al. [111]	S-EEG	RNN	Local	0.85
		Vilamala et al. [51]	S-EEG	CNN	Sleep-EDF	0.86
		Tsinalis et al. [25]	S-EEG	CNN	Sleep-EDF	0.82
		Sors et al. [50]	S-EEG	CNN	SHHS	0.87
		Chambon et al. [48]	S-EEG	Multi-view CNN	MASS session 3	N/A
		Manzano et al. [55]	S-EEG	CNN + MLP	Sleep-EDF	0.732
		Fraiwan et al. [56]	S-EEG	DBN-AE + MLP	Local	0.804
		Tan et al. [54]	S-EEG	DBN-RBM	Local	0.9278 (F1)
		Zhang et al. [49]	S-EEG	DBN + voting	UCD	0.9131
		Fernandez et al. [70]	S-EEG	CNN	SHHS	0.9 (F1)
		Supratak et al. [57]	S-EEG	CNN + LSTM	MASS/ Sleep-EDF	0.862/0.82
	AD Detection	Morabito et al. [115]	MD EEG	CNN	Local	0.82
		Zhao et al. [126]	MD EEG	DBN-RBM	Local	0.92
		Suk et al. [210]	fMRI	DBN-AE; DBN-RBM	ADNI	0.979; 0.954
		Sarraf et al. [207]	fMRI	CNN	ADNI	0.9685
		Li et al. [208]	fMRI	CNN + LR	ADNI	0.9192
		Hu et al. [217]	fMRI	D-AE + MLP	ADNI	0.875
		Ortiz et al. [212]	fMRI, PET	DBN-RBM + SVM	ADNI	0.9
	Seizure Detection	Hosseini et al. [120]	EEG	CNN	Local	0.96
		Yuan et al. [109]	MD EEG	Attention-MLP	CHB-MIT	0.9661
		Tsiouris et al. [53]	MD EEG	LSTM	CHB-MIT	>0.99
		Talathi et al. [110]	MD EEG	GRU	BUD	0.996
		Acharya et al. [114]	MD EEG	CNN	UBD	0.8867
		Schirmeister et al. [116]	MD EEG	CNN	TUH	0.854
		Hosseini et al. [117]	MD EEG	CNN	Local	N/A
		Johansen et al. [118]	MD EEG	CNN	Local	0.947 (AUC)
		Ansari et al. [119]	MD EEG	CNN + RF	Local	0.77
		Ullah et al. [112]	MD EEG	CNN + voting	UBD	0.954
		Wen et al. [124]	MD EEG	AE	Local	0.92
		Lin et al.[122]	MD EEG	D-AE	UBD	0.96
		Yuan et al. [121]	MD EEG	D-AE + SVM	CHB-MIT	0.95
		Page et al. [125]	MD EEG	DBN-AE + LR	N/A	0.8 ∼ 0.9
		Turner et al. [127]	MD EEG	DBN-RBM + LR	Local	N/A
		Hosseini et al. [129]	MD EEG	D-AE + MLP	Local	0.94
		Golmohammadi et al. [130]	MD EEG	RNN+CNN	TUH	Sen: 0.3083; Spe: 0.9686
		Shah et al. [128]	MD EEG	CNN+ LSTM	TUH	Sen: 0.39; Spe: 0.9037

Table 5. Summary of deep learning-based brain signal applications (Continued). IEF and CJD refer to Interictal Epileptic Discharge and Creutzfeldt-Jakob Disease, respectively.

Brain Signal Applications		Reference	Signals	Deep Learning Models	Dataset	Performance
Health Care	**Others:** **IED** **CJD**	Antoniades et al. [231]	EEG	AE + CNN	Local	0.68
		Morabito et al. [123]	MD EEG	D-AE	Local	0.81 ∼ 0.83
	Depression	Acharya et al. [113]	MD EEG	CNN	Local	0.935 ∼ 0.9596
		Al et al. [131]	MD EEG	DBN-RBM + MLP	Local	0.695
	Brain Tumor	Morenolopez et al. [227]	fMRI	CNN	BRATS	0.88 (F1)
		Shreyas et al. [206]	fMRI	CNN	BRATS	0.83
		Havaei et al. [205]	fMRI	Muli-scale CNN	BRATS	0.88 (F1)
	Schizophrenia	Plils et al. [211]	fMRI	DBN-RBM	Combined	0.9 (F1)
		Chu et al. [149]		CNN + RF + Voting	Local	0.816, 0.967, 0.992
	Mild Cognitive Impairment (MCI)	Suk et al. [209]	fMRI	AE + SVM	ADNI2	0.7258
	Cardiac Detection	Garg [218]	MEG	CNN	Local	Sen: 0.85, Spe: 0.97
		Hasasneh et al. [220]	MEG	CNN + MLP	Local	0.944
Smart Environment	**Robot Control Smart Home Exoskeleton Control**	Behncke et al. [139]	EEG	CNN	Local	0.75
		Zhang et al. [65]	MI EEG	RNN	EEGMMI	0.9553
		Kwak et al. [191]	SSVEP	CNN	Local	0.9403
		Huve et al. [199]	fNIRS	MLP	Local	0.82
Communication		Zhang et al. [10]	MI EEG	LSTM+CNN +AE	Local	0.9452
		Kawasaki et al. [162]	VEP	MLP	Local	0.908
		Cecotti et al. [166]	VEP	CNN	The third BCI competition, Dataset II	0.945
		Liu et al. [164]	VEP	CNN	The third BCI competition, Dataset II	0.92 ∼ 0.96
		Cecotti et al. [166]	VEP	CNN + Voting	The third BCI competition, Dataset II	0.955
		Maddula et al. [170]	VEP	RCNN	Local	0.65∼0.76
Security	**Identification**	Zhang et al. [6]	MI-EEG	Attention-based RNN	EEGMMI + local	0.9882
		Koike et al. [161]	VEP	MLP	Local	0.976
		Mao et al. [180]	RSVP	CNN	Local	0.97
	Authentication	Zhang et al. [61]	MI EEG	Hybrid	EEGMMI + local	0.984
Affective Computing		Frydenlund et al. [87]	E-EEG	MLP	DEAP	N/A
		Zhang et al. [88]	E-EEG	RNN	SEED	0.895
		Li et al. [201]	E-EEG	CNN	SEED	0.882
		Liu et al. [90]	E-EEG	CNN	Local	0.82
		Li et al. [89]	E-EEG	Hierarchical CNN	SEED	0.882
		Chai et al. [94]	E-EEG	AE	SEED	0.818
		Xu et al. [99]	E-EEG	DBN-AE, DBN-RBM	DEAP	>0.86 (F1)
		Jia et al. [98]	E-EEG	DBN-RBM	DEAP	0.8 ∼ 0.85 (AUC)
		Li et al. [100]	E-EEG	DBN-RBM	DEAP	Aro:0.642, Val:0.584, Dom 0.658
		Xu et al. [101]	E-EEG	DBN-RBM	DEAP	Aro:0.6984, Val:0.6688, Lik: 0.7539

Table 5. Summary of deep learning-based brain signal applications (Continued).

Brain Signal Applications		Reference	Signals	Deep Learning Models	Dataset	Performance
Affective Computing		Zheng et al. [102]	E-EEG	DBN-RBM + HMM	Local	0.8762
		Zhang et al. [96, 97]	E-EEG	DBN-RBM + MLP	SEED	0.8608
		Gao et al. [106]	E-EEG	DBN-RBM + MLP	Local	0.684
		Yin et al. [107]	E-EEG	Multi-view D-AE + MLP	DEAP	Aro: 0.7719; Val: 0.7617
		Mioranda et al. [104]	E-EEG	RNN + CNN	AMIGOS	¡0.7
		Alhagry et al. [108]	E-EEG	LSTM + MLP	DEAP	Aro:0.8565, Val:0.8545, Lik: 0.8799
		Liu et al. [95]	EEG	AE	SEED, DEAP	0.9101, 0.8325
		Kawde et al. [105]	EEG	DBN-RBM	DEAP	Aro: 0.7033; Val: 0.7828; Dom: 0.7016
Drive Fatigue Detection		Hung et al. [140, 140]	EEG	CNN	Local	0.572 (RMSE)
		Hung et al. [140]	EEG	CNN	Local	
		Almogbel et al. [145]	EEG	CNN	Local	0.9531
		Hajinoroozi et al. [147, 147]	EEG	CNN	Local	0.8294
		Hajinoroozi et al. [153]	EEG	DBN-RBM	Local	0.85
		San et al. [154]	EEG	DBN-RBM + SVM	Local	0.7392
		Chai et al. [158]	EEG	DBN + MLP	Local	0.931
		Du et al. [151]	EEG	D-AE + SVM	Local	0.094 (RMSE)
		Hachem et al. [189]	SSVEP	MLP	Local	0.75
Mental Load Measurement		Yin et al. [150]	EEG	D-AE	Local	0.9584
		Bashivan et al. [159]	EEG	DBN-RBM	Local	0.92
		Li et al. [155]	EEG	DBN-RBM	Local	0.9886
		Bashivan et al. [171]	EEG	R-CNN	Local	0.9111
		Bashivan et al. [172]	EEG	DBN + MLP	Local	N/A
		Naseer et al. [38]	fNIRS	MLP	Local	0.963
		Hennrich et al. [200]	fNIRS	MLP	Local	0.641
Other Appli--cations	**School Bullying**	Baltatzis et al. [141]	EEG	CNN	Local	0.937
	Music Detection	Stober et al. [142]	EEG	CNN	Local	0.776
		Stober et al. [157]	EEG	AE + CNN	Open MIIR	0.27 for 12-class
		Stober et al. [188]	EEG	CNN	Local	0.244
		Sternin et al. [148]	EEG	CNN	Local	0.75
	Number Choosing	Waytowich et al. [192]	SSVEP	CNN	Local	0.8
		Cichy et al. [204]	fMRI, MEG	CNN	N/A	N/A
	Visual Object Recognition	Manor et al. [176]	RSVP	CNN	Local	0.75
		Cecotti et al. [177]	RSVP	CNN	Local	0.897 (AUC)
		Hajinoroozi et al. [179]	RSVP	CNN	Local	0.7242 (AUC)
		Shamwell et al. [185]	RSVP	CNN	Local	0.7252 (AUC)
		Perez et al. [197]	SSVEP	AE	Local	0.9778
	Guilty Knowledge Test	Kulasingham et al. [195]	SSVEP	DBN-RBM; DBN-AE	Local	0.869; 0.8601
	Concealed Information Test	Liu et al. [168]	EEG	DBN-RBM	Local	0.973
	Flanker Task	Volker et al. [143]	EEG	CNN	Local	0.841
	Eye State	Narejo et al. [152]	EEG	DBN-RBM	UCI	0.989
		Reddy et al. [136]	EEG	MLP	Local	0.975
	User Preference	Teo et al. [135]	EEG	MLP	Local	0.6399
	Emergency Braking	Hernandez et al. [144]	EEG	CNN	Local	0.718
	Gender Detection	Putten et al. [146]	EEG	CNN	Local	0.81
		Hiroyasu et al. [201]	fNIRS	D-AE + MLP	Local	0.81

EEG signals have proven useful in evaluating the human's cognitive state in different context. Generally, a driver is regarded as in an alert state if the reaction time is lower than 0.7 seconds and in fatigue state if it is higher than 2.1 seconds. Hajinoroozi et al. [153] considered the detection of driver's fatigue from EEG signals by discovering the distinct features. They explored an approach based on DBN for dimension reduction.

Detecting driver fatigue is crucial because the drowsiness of the driver may lead to disaster. Driver fatigue detection is feasible in practice. In the hardware aspect, the collection equipment of EEG singles is off-the-shelf and portable enough to be used in a car. Moreover, the price of an EEG headset is affordable for most people. In the algorithm aspect, deep learning models have enhanced the performance of fatigue detection. As we summarized, the EEG based driving drowsiness can be recognized with high accuracy ($82\% \sim 95\%$).

Future scope of drive fatigue detection is in the self-driving scenario. As we know, in the most situation of self-driving (*e.g.*, Automation level 3[5]), the human driver is expected to respond appropriately to a request to intervene, which indicates that the driver should keep alert state. Therefore, we believe the application of brain signal-based drive fatigue detection will benefit the development of the self-driving car.

5.7. Mental Load Measurement

The EEG oscillations can be used to measure the mental workload level, which can sustain decision making and strategy development in the context of human-machine interaction [150]. Additionally, the appropriate mental workload is essential for maintaining human health and preventing accidents. For example, the abnormal mental workload of the human operator may result in performance degradation which could cause catastrophic accidents [232]. Evaluation of operator Mental Workload levels via ongoing EEG is quite promising in Human-Machine collaborative task environment to alarm the temporal operator performance degradation.

Several researchers have been paid attention to this topic. The mental workload can be measured from fNIRS signals or spontaneous EEG. Naseer et al. adopted a

MLP algorithm for fNIRS-based binary mental task level classification (mental arithmetic and rest) [38]. The experiment results showed that the MLP outperformed the traditional classifiers like SVM, KNN, and achieved the highest accuracy of 96.3%. Bashivan et al. [159] presented a statistical approach, a DBN model, for the recognition of mental workload level based on single-trial EEG. Before the DBN, the authors manually extracted the wavelet entropy and band-specific power from three frequency bands (theta, alpha, and beta). At last, the experiments demonstrated the recognition of mental workload achieved an overall accuracy of 92%. Zhang et al. [156] investigate the mental load measurement across multiple mental tasks via a recurrent-convolutional framework. The model simultaneously learns EEG features from the spatial, spectral, and temporal dimensions, which results in the accuracy of 88.9% in binary classification (high/low workload levels).

5.8. Other Applications

There are plenty of interesting scenarios beyond the above where deep learning-based brain signals can apply, such as recommender system [135] and emergency braking [144]. One possible topic is the recognition of a visual object, which may be used in guilty knowledge test [195] and concealed information test [168]. The neurons of the participant will produce a pulse when he/she suddenly watch a similar object. Based on the theory, the visual target recognition is mainly used RSVP signals. Cecotti et al. [177] aimed to build a common model for target recognition, which can work for various subjects instead of a specific subject.

Besides, researchers have investigated to distinguish the subject's gender by the fNIRS [201] and spontaneous EEG [146]. Hiriyasu et al. [201] adopted deep learning to recognize the gender of the subject based on the cerebral blood flow. The experiment results suggested that the cerebral blood flow changes in different ways for male and female. Putten et al. [146] tried to discover the sex-specific information from the brain rhythms and adopted a CNN model to recognize the participant's gender. This paper illustrated that fast beta activity (20 $\sim$25 Hz) is one of the most distinctive attributes.

5.9. Benchmark Datasets

We have extensively explored the benchmark datasets usable for deep learning-based brain signals (Table 6). We provide a bunch of public datasets with download links, which cover most brain signal types. In particular, BCI competition IV (BCI-C IV) contains five datasets via the same link. For better understanding, we present the number of subjects, the number of class (how many categories), sampling rate, and the number of channels of each dataset. In the '# Channel' column, the default channel is for EEG

---

[5]https://en.wikipedia.org/wiki/Self-driving_car
[6]https://physionet.org/physiobank/database/sleep-edfx/
[7]https://massdb.herokuapp.com/en/
[8]https://physionet.org/pn3/shhpsgdb/
[9]https://physionet.org/pn6/chbmit/
[10]https://www.isip.piconepress.com/projects/tuh_eeg/html/downloads.shtml
[11]https://physionet.org/pn4/eegmmidb/
[12]http://www.bbci.de/competition/ii/
[13]http://www.eecs.qmul.ac.uk/mmv/datasets/amigos/readme.html
[14]http://bcmi.sjtu.edu.cn/ seed/download.html
[15]https://www.eecs.qmul.ac.uk/mmv/datasets/deap/
[16]https://owenlab.uwo.ca/research/the_openmiir_dataset.html
[17]http://adni.loni.usc.edu/data-samples/access-data/
[18]https://www.med.upenn.edu/sbia/brats2018/data.html

Table 6: The summary of public dataset for brain signal studies. The '# Sub', '# Cla', and S-Rate denote the number of subject, number of class, and sampling rate, respectively. FM denote finger movement while BCI-C denote the BCI Competition. The '# channel' refers to the number of brain signal channels.

Brain Signals		Name Link	# Sub	# Cla	S-Rate	# Channel
EEG	Sleep EEG	Sleep-EDF[6]: Telemetry	22	6	100	2
		Sleep-EDF: Cassette	78	6	100, 1	2
		MASS-1[7]	53	5	256	17
		MASS-2	19	6	256	19
		MASS-3	62	5	256	20
		MASS-4	40	6	256	4
		MASS-5	26	6	256	20
		SHHS[8]	5804	N/A	125, 50	2
	Seizure EEG	CHB-MIT[9]	22	2	256	18
		TUH[10]	315	2	200	19
	MI EEG	EEGMMI[11]	109	4	160	64
		BCI-C II[12], Dataset III	1	2	128	3
		BCI-C III, Dataset III a	3	4	250	60
		BCI-C III, Dataset III b	3	2	125	2
		BCI-C III, Dataset IV a	5	2	1000	118
		BCI-C III, Dataset IV b	1	2	1001	119
		BCI-C III, Dataset IV c	1	2	1002	120
		BCI-C IV, Dataset I	7	2	1000	64
		BCI-C IV, Dataset II a	9	4	250	22
		BCI-C IV, Dataset II b	9	2	250	3
	Emotional EEG	AMIGOS[13]	40	4	128	14
		SEED[14]	15	3	200	62
		DEAP[15]	32	4	512	32
	Others EEG	Open MIIR[16]	10	12	512	64
	VEP	BCI-C II, Dataset II b	1	36	240	64
		BCI-C III, Dataset II	2	26	240	64
fMRI		ADNI[17]	202	3	N/A	N/A
		BRATS[18]2013	65	4	N/A	N/A
MEG		BCI-C IV, Dataset III	2	4	400	10

signals. Some datasets contain more biometric signals (*e.g.*, ECG), but we only list the channels related to brain signals.

6. Analysis and Guidelines

In this section, we first analyze what is the most suitable deep learning models for each brain signal. Then, we summarize the popular deep learning models in brain signal research. At last, we investigate the brain signals in terms of application. We hope this survey could help our readers to select the most effective and efficient methods when dealing with brain signals. Please recall Table 4 where we summarize the brain signals and the corresponding deep learning models of the state-of-the-art papers. Figure 4 illustrated of the publications proportion for crucial brain signals and deep learning models.

6.1. Brain Signal Acquisition

Among the non-invasive signals, the studies on EEG is far more than the sum of all the other brain signal paradigms

(fNIRS, fMRI, and MEG). Furthermore, there are about 70% of the EEG papers pay attention to the spontaneous EEG (133 publications). For better understanding, we split the spontaneous EEG into several aspects: the sleep, the motor imagery, the emotional, the mental disease, the data augmentation, and others.

First, the classification of the sleep EEG mainly depends on the discriminative and the hybrid models. Among the nineteen studies about sleep stage classification, there are six employed CNN and the modified CNN models independently while two papers adopted RNN models. There are three hybrid models built on the combination of CNN and RNN.

Second, in terms of the research on MI EEG (30 publications), the independent CNN and CNN-based hybrid models are widely used. As for the representative models, DBN-RBM is often applied to capture the latent features from the MI EEG signals.

Third, there are twenty-five publications related to spontaneous emotional EEG. More than half of them employed representative models (such as D-AE, D-RBM,

especially DBN-RBM) for unsupervised feature learning. The most typical state recognition works recognize the user's emotion as positive, neutral, or negative. Some researchers take a further step to classify the valence, and the arouse rate, which is more complex and challenging.

Fourth, the research on mental disease diagnosis is promising and attracting. The majority of the related research focuses on the detection of epileptic seizure and Alzheimer's Disease. Since the detection is a binary classification problem which is rather easier than multi-class classification, many studies can achieve a high accuracy like above 90%. In this area, the standard CNN model and the D-AE are prevalent. One possible reason is that CNN and AE are the most well-known and effective deep learning models for classification and dimensionality reduction.

Fifth, several publications pay attention to the GAN based data augmentation. At last, about thirty studies are investigating other spontaneous EEG such as driving fatigue, audio/visual stimuli impact, cognitive/mental load, and eye state detection. These studies extensively apply standard CNN models and variants.

Moreover, apart from spontaneous EEG, evoked potentials also attracted much attention. On the one hand, in ERP, VEP and the subcategory RSVP has drawn lots of investigations because visual stimuli, compared to other stimuli, is easier to be conducted and more applicable in the real world (*e.g.*, P300 speller can be used for brain typing). For VEP (twenty-one publications), there are elven studies applied discriminative models, and six works adopted hybrid models. In terms of RSVP, the sole CNN dominates the algorithms. Apart from them, five papers focused on the analysis of AEP signals. On the other hand, among the steady-state related researches, only SSVEP has been studied by deep learning models. Most of them only applied discriminative models on the recognition of the target image.

Furthermore, beyond the diverse EEG diagrams, a wide range of papers paid attention to fNIRS and fMRI. The fNIRS images are rarely studied by deep learning, and the major studies just employed the simple MLP models. We believe more attention should be paid to the research on fNIRS for the high portability and low cost. As for the fMRI, twenty-three papers proposed deep learning models to the classification. The CNN model is widely used for its outstanding performance in feature learning from images. There are also several papers interested in image reconstruction based on fMRI signals. One reason why fMRI is so hot is that several public datasets are available on the Internet, although the fMRI equipment is expensive. The MEG signals are mainly used in the medical area, which is insensitive to the deep learning algorithm. Thus, we only found very few studies on MEG. The sparse AE and CNN algorithms have a positive influence on the feature refining and classification of MEG.

6.2. Selection Criteria for Deep Learning Models

Our investigation shows that discriminative models are most frequent in the summarized publications. This is reasonable at a high level because a large proportion of brain signal issues can be regarded as a classification problem. Another observation is that CNN and its variants are adopted in more than 70% of the discriminative models, for which we provide reasons as follows.

First, the design of CNN is powerful enough to extract the latent discriminative features and spatial dependencies from the EEG signals for classification. As a result, CNN structures are adopted for classification in some studies while adopted for feature extraction in some other studies.

Second, CNN has been achieved great success in some research areas (*e.g.*, computer vision), which makes it extremely famous and feasible (public codes). Thus, the brain signal researchers have more chance to understand and apply CNN on their works.

Third, some brain signal diagrams (*e.g.*, fMRI) are naturally formed as two-dimension images that are conducive to be processedg by CNN. Meanwhile, other 1-D signals (*e.g.*, EEG) could be converted into 2-D images for further analysis by CNN. Here, we provide several methods converting 1-D EEG signals (with multiple channels) to the 2-D matrix: 1) convert each time-point[19] to a 2-D image; 2) convert a segment into a 2-D matrix. In the first situation, suppose we have 32 channels, and we can collect 32 elements (each element corresponding to a channel) at each time-point. As described in [89], the collected 32 elements could be converted into a 2-D image based on the spatial position. In the second situation, suppose we have 32 channels, and the segment contains 100 time-points. The collected data can be arranged as a matrix with the shape of $[32, 100]$ where each row and column refers to a specific channel and time-point, respectively.

Fourth, there are a lot of variants of CNN which are suitable for a wide range of brain signal scenarios. For example, the single-channel EEG signals can be processed by 1-D CNN. In terms of RNN, only about 20% of discriminative model-based papers adopted RNN, which is much less than we expected since RNN has demonstrated powerful in temporal feature learning. One possible reason for this phenomena is that processing a long sequence by RNN is time-consuming and the EEG signals are generally formed as a long sequence. For example, the sleep signals are usually sliced into segments with 30 seconds, which has 3000 time-points under 100 Hz sampling rate. For a sequence with 3000 elements, through our preliminary experiments, RNN takes more than 20 folds training time than CNN. Moreover, MLP is not popular due to its inferior effectiveness (*e.g.*, non-linear ability) to the other algorithms its simple deep learning architecture.

---

[19] Time-point represents one sampling point. For example, we can have 100 time-points if the sampling rate is 100 Hz.

As for representative models, DBN, especially DBN-RBM, is the most popular model for feature extraction. DBN is widely used in brain signal for two reasons: 1) it learns the generative parameters that reveal the relationship of variables in neighboring layers efficiently; 2) it makes it straightforward to calculate the values of latent variables in each hidden layer [233]. However, most works that employed the DBN-RBM model were published before 2016. It can be inferred that the researchers prefer to use DBN for feature learning followed by a non-deep learning classifier before 2016; but recently, an increasing number of studies would like to adopt CNN or hybrid models for both feature learning and classification.

Moreover, generative models are rarely employed independently. The GAN- and VAE-based data augmentation and image reconstruction are mainly focused on fMRI and EEG signals. It is demonstrated that the trained classifier will achieve more competitive performance after data augmentation. Therefore, this is a promising research prospect in the future.

Last but not the least, there are fifty-three publications proposed hybrid models for brain signal studies. Among them, the combinations of RNN and CNN take about one-fifth proportion. Since RNN and CNN are illustrated having excellent temporal and spatial feature extraction ability, it is natural to combine them for both temporal and spatial feature learning. Another type of hybrid models is the combination of representative and discriminative models. This is easy to understand because the former is employed for feature refining, and the latter is employed for classification. There are twenty-eight publications which almost covered all the brain signals proposed this type of hybrid deep learning models. The adopted representative models are mostly AE or DBN-RBM; at the meanwhile, the adopted discriminative models are mostly CNN. Apart from that, there are twelve papers proposed other hybrid models such as two discriminative models. For example, several studies proposed the combination of CNN and MLP where a CNN structure is used for extract spatial features and an MLP is used for classification.

6.3. Application Performance

In order to have a closer observation of the recent advances on deep learning-based brain signal analysis, we analyze the brain signal acquisition methods and the deep learning algorithms in terms of application performance. In some cases, various studies adopt the same deep architecture working on the same dataset but results in different performance, which maybe caused by the different pre-processing methods and hyper-parameter settings.

To begin with, the most appealing and hot field is that using brain signal analysis on health care area. For sleep quality evaluation, the dominate brain signals are spontaneous EEG which are measured while the patient is sleeping. The single RNN or CNN models seem have a good discriminative feature learning ability and lead to a comprehensive performance. Generally, most of the deep learning algorithms can achieve the accuracy of above 85% in the context of multiple sleep stage scenario. Upon this, the combined hybrid models (*e.g.*, CNN integrates with LSTM) can only have incremental improvements.

One key method to detect Alzheimer's Disease is brain signal analysis by measuring the functions of specific brain regions. In detail, the diagnosis can be conducted by spontaneous EEG signals or fMRI images. For MD EEG, DBN is supposed to outperform CNN since the EEG signals contains more temporal instead of spatial information. As for the fMRI pictures, CNN have great advantages in the grid-arranged spatial information learning, which makes it obtain a very comprehensive classification accuracy (above 90%). As for epileptic seizure, the diagnosis are generally based on EEG signals. The single RNN classifier (*e.g.*, LSTM or GRU) seems work better than its counterparts due to the excellent temporal dependency representing ability. Here, the complex hybrid models indeed outperform the single component. For example, [130] achieves a better specification than [116] on the same dataset because of combing with RNN. Most of the epileptic seizure detection models claim a rather high classification accuracy (above 95%). One possible reason is that the binary recognition scenario is much easier than multi-class classification.

The brain signal-controlled smart environment only appear in a small number of publications. Among them, the brain signals are collected through very different methods. This is an emerging but promising field because it is easy to integrate with smart home and smart hospital to benefit the individuals whether healthy or disable. Another advantage of brain signals is bridging people's inside and outer world by communication techniques. In this area, lots of investigations are focusing on the VEP signals because the visual evoked potential is obvious and easy to be detected. One important data source is from the third BCI competition. In addition, brain signal analysis can be widely implement in security systems since the brain signals are invisible and very hard to be mimicked. The characteristic of high fake-resistance enables brain signal a raising star in the identification/authentication in confidential scenarios. The drawbacks of brain signal-based security systems are the expensive equipment and inconvenient (*e.g.*, the subject have to wear an EEG headset to monitor the brainwaves).

Affective computing has drawn much attention in recent years. The EEG signals have high temporal resolution and able to capture the quick-varying emotions. Therefore, almost all the studies are based on spontaneous EEG signals. The signals are gathered when the subject is watching video which is supposed to arouse the subject's specific emotion. Another reason for this phenomenon is that there are several open-source EEG-based affecting analysis datasets (*e.g.*, DEAP and SEED) which greatly

promote the investigation in this area. The EEG-based affective computing contains two mainstreams. One of them focuses on developing powerful discriminative classifiers (such as hierarchical CNN) which are designed to perform feature extraction and classification in the same step. The other tries to learn the latent features through deep representative models (*e.g.*, DBN-RBM) and then send the learned representations into a powerful classifier (such as HMM and MLP). It can be observed that the former models ([88, 201]) seem outperform the latter methods ([96]) with a small margin on the SEED dataset.

Drive fatigue detection can be easily integrated in the platforms such as self-driving vehicles. Nevertheless, there are only a few publications in this area due to the expensive experimental cost and the lack of accessible dataset. Moreover, there are a lot of interesting applications (*e.g.*, guilty knowledge test and gender detection) have been explored by deep learning models.

7. Open Issues

Although deep learning has lifted the performance of brain signal systems, technical and usability challenges remain. The technical challenges concern the classification ability in complex scenarios, and the usability challenges refer to limitations in large scale real-world deployment. In this section, we introduce these challenges and point out the possible solutions.

7.1. Explainable General Framework

Until now, we have introduced several types of brain signals (*e.g.*, spontaneous EEG, ERP, fMRI) and deep learning models that have been applied for each type. One promising research direction for deep learning-based brain signal research is to develop a general framework that can handle various brain signals regardless of the number of channels used for signal collection, the sample dimensions (*e.g.*, 1-D or 2-D sample), and stimulation types (*e.g.*, visual or audio stimuli), etc. The general framework would require two key capabilities: the attention mechanism and the ability to capture latent feature. The former guarantees the framework can focus on the most valuable parts of input signals, and the latter enables the framework to capture the distinctive and informative features.

The attention mechanism can be implemented based on attention scores or by various machine learning algorithms such as reinforcement learning. The attention scores can be inferred from the input data and work as a weight to help the framework to pay attention to the parts with high attention scores. Reinforcement learning has shown to be able to find the most valuable part through a policy search [85]. CNN is the most suitable structure for capturing features at various levels and ranges. In the future, CNN could be used as a fundamental feature learning tool and be integrated with suitable attention mechanisms to form a general classification framework.

One additional direction we may consider is how to interpret the feature representation derived by the deep neural network, what is the intrinsic relationship between the learned features and the task-related neural pattern, or neuropathology of mental disorders. More and more people are realizing that interpretation could be even more important than prediction performance, since we usually just treat deep learning as a black box.

7.2. Subject-Independent Classification

Until now, most brain signal classification tasks focus on person-dependent scenarios, where the training samples and testing samples are collected from the identical individual. The future direction is to realize person-independent classification so that the testing data will never appear in the training set. High-performance person-independent classification is compulsory for the wide application of brain signals in the real world.

One possible solution to achieving this goal is to build a personalized model with transfer learning. A personalized affective model can adopt a transductive parameter transfer approach to construct individual classifiers and to learn a regression function that maps the relationship between data distribution and classifier parameters [234]. Another potential solution is mining the subject-independent component from the input data. The input data can be decomposed into two parts: a subject-dependent component, which depends on the subject and a subject-independent component, which is common for all subjects. A hybrid multi-task model can work on two tasks simultaneously, one focusing on person identification and the other on class recognition. A well-trained and converged model is supposed to extract the subject-independent features in the class recognition task.

7.3. Semi-supervised and Unsupervised Classification

The performance of deep learning highly depends on the size of training data, which, however, requires expensive and time-consuming manual labeling to collect abundant class labels for a wide range of scenarios such as sleep EEG. While supervised learning requires both observations and labels for the training, unsupervised learning requires no labels, and semi-supervised learning only requires partial labels [98]. They are, therefore, more suitable for problems with little ground truth.

Zhang et al. proposed an Adversarial Variational Embedding (AVAE) framework that combines a VAE++ model (as a high-quality generative model) and semi-supervised GAN (as a posterior distribution learner) [235] for robust and effective semi-supervised learning. Jia et al. proposed a semi-supervised framework by leveraging the data distribution of unlabelled data to prompt the representation learning of labelled data [98].

Two methods may enhance the unsupervised learning: one is to employ crowd-sourcing to label the unlabeled observations; the other is to leverage unsupervised domain adaption learning to align the distribution of source brain signals and the distribution of target signals with a linear transformation.

7.4. Online Implementation

Most of the existing brain signal systems focus on offline procedure which means that the training and testing dataset are pre-collected and evaluated offline. However, in the real-world scenarios, the brain signal systems are supposed to receive live data stream and produce classification results in real time, which is still very challenging.

For EEG signals, in the online system, compared to the offline procedure, the gathered live signals are more noisy and unstable due to lots of factors such as the less-concentrating of the subject [236] and the inherent destabilization of the equipment (*e.g.*, fluctuating sampling rate). Through our empirical experiments, online brain signal systems generally perform a lower accuracy of 10% than their counterparts. One future scope of online implementation is to develop a batch of robust algorithms in order to handle the influence factors and discover the latent distinctive patterns underlying the noisy live brain signals. [237] implemented an EEG-based online system that achieves comparable performance, however, this work only investigates a very high-level target (*i.e.*, human attention). Discovering the latent invariant representations through covariance matrices of EEG signals can help to mitigate the influence of extinct perturbations [238]. Some post-processing methods (*e.g.*, voting and aggregating) [166, 149] can help to improve the decoding performance by averaging the results from multiple continues samples. However, these methods will inevitably bring higher latency. Thus, the post-processing requires a trade-off between the high-accuracy and low-latency.

For fNIRS and fMRI, the online evaluation is relatively less challenging since they have a rather low temporal resolution. The online images with less dynamic can be regarded as static images to some extent, which makes the online system approximating to the offline system. Furthermore, most fMRI and MEG signals are used to evaluate the user's neurological status (*e.g.*, detect the effects of tumor) which does not require an instantaneous response. Thus, they have less demand for a real-time monitoring system.

7.5. Hardware Portability

Poor portability of hardware has been preventing brain signals from wide application in the real world. In most scenarios, users would like to use small, comfortable, or even wearable brain signal hardware to collect brain signals and to control appliances and assistant robots.

Currently, there are three types of EEG collection equipment: the unportable, the portable headset, and ear-EEG sensors. The unportable equipment has high sampling frequency, channel numbers, and signal quality but is expensive. It is suitable for physical examination in a hospital. The portable headsets (*e.g.*, Neurosky, Emotiv EPOC) have $1 \sim 14$ channels and $128 \sim 256$ sampling rate but has inaccuracy readings and cause discomfort after long-time use. The ear-EEG sensors, which are attached to the outer eat, have gained increasing attention recently but remain mostly at the laboratory stage [239]. The ear-EEG sensors contain a series of electrodes which are placed in each ear canal and concha [240]. The EEGrids, to the best of our knowledge, is the only commercial ear-EEG. It has multi-channel sensor arrays placed around the ear using an adhesive [20] and is even more expensive. A promising future direction is to improve the usability by developing a cheaper (*e.g.*, lower than 200$) and more comfortable (*e.g.*, can last longer than 3 hours without feeling uncomfortable) wireless ear-EEG equipment.

8. Conclusion

In this paper, we thoroughly summarize the recent advances in deep learning models for non-invasive brain signal analysis. Compared with traditional machine learning methods, deep learning not only enables to learn high-level features automatically from brain signals but also have less dependency on domain knowledge. We organize brain signals and dominant deep learning models, followed by discussing state-of-the-art deep learning techniques for brain signals. Moreover, we provide guidelines to help researchers to find the suitable deep learning algorithms for each category of brain signals. Finally, we overview deep learning-based brain signal applications and point out the open challenges and future directions.

References

[1] T. Ball, M. Kern, I. Mutschler, A. Aertsen, and A. Schulze-Bonhage, "Signal quality of simultaneously recorded invasive and non-invasive eeg," *Neuroimage*, vol. 46, no. 3, pp. 708–716, 2009.

[2] X. Zhang, L. Yao, S. Zhang, S. Kanhere, M. Sheng, and Y. Liu, "Internet of things meets brain-computer interface: A unified deep learning framework for enabling human-thing cognitive interactivity," *IEEE Internet of Things Journal*, 2018.

[3] X. An, D. Kuang, X. Guo, Y. Zhao, and L. He, "A deep learning method for classification of eeg data based on motor imagery," in *International Conference on Intelligent Computing*, 2014, pp. 203–210.

[4] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of eeg motor imagery signals," *Journal of neural engineering*, vol. 14, no. 1, p. 016003, 2016.

[5] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year

---

[20]http://ceegrid.com/home/concept/

update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[6] X. Zhang, L. Yao, S. S. Kanhere, Y. Liu, T. Gu, and K. Chen, "Mindid: Person identification from brain waves through attention-based recurrent neural network," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, p. 149, 2018.

[7] S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213–230, 2015.

[8] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals," *Journal of Neural engineering*, vol. 4, no. 2, p. R32, 2007.

[9] W. Samek, K.-R. Müller, M. Kawanabe, and C. Vidaurre, "Brain-computer interfacing in discriminative and stationary subspaces," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 2873–2876.

[10] X. Zhang, L. Yao, Q. Z. Sheng, S. S. Kanhere, T. Gu, and D. Zhang, "Converting your thoughts to texts: Enabling brain typing via deep feature learning of eeg signals," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2018, pp. 1–10.

[11] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.

[12] H. Cecotti, M. P. Eckstein, and B. Giesbrecht, "Single-trial classification of event-related potentials in rapid serial visual presentation tasks using supervised spatial filtering," *IEEE transactions on neural networks and learning systems*, vol. 25, no. 11, pp. 2030–2042, 2014.

[13] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 6, pp. 2063–2079, 2018.

[14] D. Wen, Z. Wei, Y. Zhou, G. Li, X. Zhang, and W. Han, "Deep learning methods to process fmri data and their application in the diagnosis of cognitive impairment: A brief overview and our opinion," *Frontiers in neuroinformatics*, vol. 12, p. 23, 2018.

[15] S. Mason, A. Bashashati, M. Fatourechi, K. Navarro, and G. Birch, "A comprehensive survey of brain interface technology designs," *Annals of biomedical engineering*, vol. 35, no. 2, pp. 137–169, 2007.

[16] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[17] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019.

[18] X. Wang, G. Gong, N. Li, and Y. Ma, "A survey of the bci and its application prospect," in *Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems*, 2016, pp. 102–111.

[19] F. Movahedi, J. L. Coyle, and E. Sejdić, "Deep belief networks for electroencephalography: A review of recent contributions and future outlooks," *IEEE journal of biomedical and health informatics*, vol. 22, no. 3, pp. 642–652, 2018.

[20] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen, "Brain–machine interfaces in neurorehabilitation of stroke," *Neurobiology of disease*, vol. 83, pp. 172–179, 2015.

[21] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain–computer interface: a brief review," *Journal of neuroscience methods*, vol. 243, pp. 103–110, 2015.

[22] S. Ruiz, K. Buyukturkoglu, M. Rana, N. Birbaumer, and R. Sitaram, "Real-time fmri brain computer interfaces: self-regulation of single brain regions to networks," *Biological psychology*,

vol. 95, pp. 4–20, 2014.

[23] A. Haider and R. Fazel-Rezai, "Application of p300 event-related potential in brain-computer interface," in *Event-Related Potentials and Evoked Potentials*, 2017.

[24] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, and J. Wang, "Applications of deep learning to mri images: a survey," *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1–18, 2018.

[25] O. Tsinalis, P. M. Matthews, Y. Guo, and S. Zafeiriou, "Automatic sleep stage scoring with single-channel eeg using convolutional neural networks," *arXiv preprint arXiv:1610.01683*, 2016.

[26] Q. Gui, M. Ruiz-blondet, S. Laszlo, and Z. Jin, "A survey on brain biometrics," *ACM Computing Surveys*, vol. 51, no. 112, 2019.

[27] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of eeg-based brain–computer interface paradigms," *Journal of neural engineering*, vol. 16, no. 1, p. 011001, 2019.

[28] H. Cecotti and A. J. Ries, "Best practice for single-trial detection of event-related potentials: Application to brain-computer interfaces," *International Journal of Psychophysiology*, vol. 111, pp. 156–169, 2017.

[29] N. Naseer and K.-S. Hong, "fnirs-based brain-computer interfaces: a review," *Frontiers in human neuroscience*, vol. 9, p. 3, 2015.

[30] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[31] L. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," *APSIPA Transactions on Signal and Information Processing*, vol. 3, 2014.

[32] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "Emg and eog artifacts in brain computer interface systems: A survey," *Clinical neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.

[33] S. Vieira, W. H. Pinaya, and A. Mechelli, "Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications," *Neuroscience & Biobehavioral Reviews*, vol. 74, pp. 58–75, 2017.

[34] P. Aricò, G. Borghini, G. Di Flumeri, N. Sciaraffa, and F. Babiloni, "Passive bci beyond the lab: current trends and future directions," *Physiological measurement*, vol. 39, no. 8, p. 08TR02, 2018.

[35] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.

[36] D. Huang, K. Qian, D.-Y. Fei, W. Jia, X. Chen, and O. Bai, "Electroencephalography (eeg)-based brain–computer interface (bci): A 2-d virtual wheelchair control based on event-related desynchronization/synchronization and state control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 3, pp. 379–388, 2012.

[37] D. Regan, "Steady-state evoked potentials," *JOSA*, vol. 67, no. 11, pp. 1475–1489, 1977.

[38] N. Naseer, N. K. Qureshi, F. M. Noori, and K.-S. Hong, "Analysis of different classification techniques for two-class functional near-infrared spectroscopy-based brain-computer interface," *Computational intelligence and neuroscience*, vol. 2016, 2016.

[39] S. Singh, S. Jain, T. Ahuja, Y. Sharma, and N. Pathak, "Study for reduction of pollution level in diesel engines, petrol engines and generator sets by bio signal ring," *International Journal of Advance Research and Innovation*, vol. 6, no. 3, pp. 175–181, 2018.

[40] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," *IEEE Transactions on neural networks*, vol. 3, no. 5, pp. 683–697, 1992.

[41] T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, and S. Khudanpur, "Recurrent neural network based language model," in *Eleventh annual conference of the international speech communication association*, 2010.

[42] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for lvcsr," in *2013 IEEE international conference on acoustics, speech and signal processing*, 2013, pp. 8614–8618.

[43] M. A. Kramer, "Nonlinear principal component analysis using

autoassociative neural networks," *AIChE journal*, vol. 37, no. 2, pp. 233–243, 1991.

[44] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, no. 5786, pp. 504–507, 2006.

[45] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009.

[46] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[47] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[48] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, and A. Gramfort, "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018.

[49] J. Zhang, Y. Wu, J. Bai, and F. Chen, "Automatic sleep stage classification based on sparse deep belief net and combination of multiple classifiers," *Transactions of the Institute of Measurement and Control*, vol. 38, no. 4, pp. 435–451, 2016.

[50] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, and J.-F. Payen, "A convolutional neural network for sleep stage scoring from raw single-channel eeg," *Biomedical Signal Processing and Control*, vol. 42, pp. 107–114, 2018.

[51] A. Vilamala, K. H. Madsen, and L. K. Hansen, "Neural networks for interpretable analysis of eeg sleep stage scoring," in *International Workshop on Machine Learning for Signal Processing 2017*, 2017.

[52] S. Biswal, J. Kulas, H. Sun, B. Goparaju, M. B. Westover, M. T. Bianchi, and J. Sun, "Sleepnet: automated sleep staging system via deep learning," *arXiv preprint arXiv:1707.08262*, 2017.

[53] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in biology and medicine*, vol. 99, pp. 24–37, 2018.

[54] D. Tan, R. Zhao, J. Sun, and W. Qin, "Sleep spindle detection using deep learning: a validation study based on crowdsourcing," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 2828–2831.

[55] M. Manzano, A. Guillén, I. Rojas, and L. J. Herrera, "Combination of eeg data time and frequency representations in deep networks for sleep stage classification," in *International Conference on Intelligent Computing*, 2017, pp. 219–229.

[56] L. Fraiwan and K. Lweesy, "Neonatal sleep state identification using deep learning autoencoders," in *Signal Processing & its Applications (CSPA), 2017 IEEE 13th International Colloquium on*, 2017, pp. 228–231.

[57] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998–2008, 2017.

[58] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 2, pp. 324–333, 2018.

[59] K. G. Hartmann, R. T. Schirrmeister, and T. Ball, "Hierarchical internal representation of spectral features in deep convolutional networks trained for eeg decoding," in *Brain-Computer Interface (BCI), 2018 6th International Conference on*, 2018, pp. 1–6.

[60] E. Nurse, B. S. Mashford, A. J. Yepes, I. Kiral-Kornek, S. Harrer, and D. R. Freestone, "Decoding eeg and lfp signals using deep learning: heading truenorth," in *Proceedings of the ACM International Conference on Computing Frontiers*, 2016, pp. 259–266.

[61] X. Zhang, L. Yao, K. Chen, X. Wang, Q. Sheng, and T. Gu, "Deepkey: An eeg and gait based dual-authentication system," *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2017.

[62] H. Yang, S. Sakhavi, K. K. Ang, and C. Guan, "On the use of convolutional neural networks and augmented csp features for multi-class motor imagery of eeg signals classification," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 2620–2623.

[63] L. Jingwei, C. Yin, and Z. Weidong, "Deep learning eeg response representation for brain computer interface," in *Control Conference (CCC), 2015 34th Chinese*, 2015, pp. 3518–3523.

[64] H. K. Lee and Y.-S. Choi, "A convolution neural networks scheme for classification of motor imagery eeg based on wavelet time-frequecy image," in *International Conference on Information Networking (ICOIN) 2018*, 2018, pp. 906–909.

[65] X. Zhang, L. Yao, C. Huang, Q. Z. Sheng, and X. Wang, "Intent recognition in smart living through deep recurrent neural networks," in *International Conference on Neural Information Processing*, 2017, pp. 748–758.

[66] Z. Tang, C. Li, and S. Sun, "Single-trial eeg classification of motor imagery using deep convolutional neural networks," *Optik-International Journal for Light and Electron Optics*, vol. 130, pp. 11–18, 2017.

[67] Q. Wang, Y. Hu, and H. Chen, "Multi-channel eeg classification based on fast convolutional feature extraction," in *International Symposium on Neural Networks*, 2017, pp. 533–540.

[68] I. Sturm, S. Lapuschkin, W. Samek, and K.-R. Müller, "Interpretable deep neural networks for single-trial eeg classification," *Journal of neuroscience methods*, vol. 274, pp. 141–145, 2016.

[69] M. Shahin, B. Ahmed, S. T.-B. Hamida, F. L. Mulaffer, M. Glos, and T. Penzel, "Deep learning and insomnia: Assisting clinicians with their diagnosis," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1546–1553, 2017.

[70] I. Fernández-Varela, D. Athanasakis, S. Parsons, E. Hernández-Pereira, and V. Moret-Bonillo, "Sleep staging with deep learning: a convolutional model," in *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2018)*.

[71] A. M. Chiarelli, P. Croce, A. Merla, and F. Zappasodi, "Deep learning for hybrid eeg-fnirs brain–computer interface: application to motor imagery classification," *Journal of neural engineering*, vol. 15, no. 3, p. 036028, 2018.

[72] T. Uktveris and V. Jusas, "Application of convolutional neural networks to four-class motor imagery classification problem," *Information Technology And Control*, vol. 46, no. 2, pp. 260–273, 2017.

[73] V. Lawhern, A. Solon, N. Waytowich, S. M. Gordon, C. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, 2018.

[74] J. Li, Z. Struzik, L. Zhang, and A. Cichocki, "Feature learning from incomplete eeg with denoising autoencoder," *Neurocomputing*, vol. 165, pp. 23–31, 2015.

[75] S. Redkar, "Using deep learning for human computer interface via electroencephalography," *IAES International Journal of Robotics and Automation*, vol. 4, no. 4, 2015.

[76] X. Zhang, L. Yao, D. Zhang, X. Wang, Q. Z. Sheng, and T. Gu, "Multi-person brain activity recognition via comprehensive eeg signal analysis," in *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2017, pp. 28–37.

[77] J. Li and A. Cichocki, "Deep learning of multifractal attributes from motor imagery induced eeg," in *International Conference on Neural Information Processing*, 2014, pp. 503–510.

[78] Y. Ren and Y. Wu, "Convolutional deep belief networks for feature extraction of eeg signal," in *International Joint Conference on Neural Networks (IJCNN)*, 2014, pp. 2850–2853.

[79] S. Kumar, A. Sharma, K. Mamun, and T. Tsunoda, "A deep learning approach for motor imagery eeg signal classification,"

in *Computer Science and Engineering (APWC on CSE), 2016 3rd Asia-Pacific World Congress on*, 2016, pp. 34–39.

[80] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted boltzmann machines," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 25, no. 6, pp. 566–576, 2017.

[81] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "Eeg classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, no. 3, p. 551, 2019.

[82] C. Tan, F. Sun, W. Zhang, J. Chen, and C. Liu, "Multimodal classification with deep convolutional-recurrent neural networks for electroencephalography," in *International Conference on Neural Information Processing*, 2017, pp. 767–776.

[83] L. Duan, M. Bao, J. Miao, Y. Xu, and J. Chen, "Classification based on multilayer extreme learning machine for motor imagery task from eeg signals," *Procedia Computer Science*, vol. 88, pp. 176–184, 2016.

[84] E. S. Nurse, P. J. Karoly, D. B. Grayden, and D. R. Freestone, "A generalizable brain-computer interface (bci) using machine learning for feature discovery," *PloS one*, vol. 10, no. 6, p. e0131328, 2015.

[85] X. Zhang, L. Yao, C. Huang, S. Wang, M. Tan, G. Long, and C. Wang, "Multi-modality sensor data classification with selective attention," *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2018.

[86] S. Sakhavi, C. Guan, and S. Yan, "Parallel convolutional-linear neural network for motor imagery classification," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, 2015, pp. 2736–2740.

[87] A. Frydenlund and F. Rudzicz, "Emotional affect estimation using video and eeg data in deep neural networks," in *Canadian Conference on Artificial Intelligence*, 2015, pp. 273–280.

[88] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," *IEEE transactions on cybernetics*, no. 99, pp. 1–9, 2018.

[89] J. Li, Z. Zhang, and H. He, "Implementation of eeg emotion recognition system based on hierarchical convolutional neural networks," in *International Conference on Brain Inspired Cognitive Systems*, 2016, pp. 22–33.

[90] W. Liu, H. Jiang, and Y. Lu, "Analyze eeg signals with convolutional neural network based on power spectrum feature selection," *Proceedings of Science*, 2017.

[91] F. Wang, S.-h. Zhong, J. Peng, J. Jiang, and Y. Liu, "Data augmentation for eeg-based emotion recognition with deep convolutional neural networks," in *International Conference on Multimedia Modeling*, 2018, pp. 82–93.

[92] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for eeg-based emotion recognition," *Cognitive Computation*, pp. 1–13, 2017.

[93] K. Wang, Y. Zhao, Q. Xiong, M. Fan, G. Sun, L. Ma, and T. Liu, "Research on healthy anomaly detection model based on deep learning from multiple time-series physiological signals," *Scientific Programming*, vol. 2016, 2016.

[94] X. Chai, Q. Wang, Y. Zhao, X. Liu, O. Bai, and Y. Li, "Unsupervised domain adaptation techniques based on auto-encoder for non-stationary eeg-based emotion recognition," *Computers in biology and medicine*, vol. 79, pp. 205–214, 2016.

[95] W. Liu, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using multimodal deep learning," in *International Conference on Neural Information Processing*, 2016, pp. 521–529.

[96] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, 2015.

[97] W.-L. Zheng, H.-T. Guo, and B.-L. Lu, "Revealing critical channels and frequency bands for emotion recognition from eeg with deep belief network," in *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*, 2015, pp. 154–157.

[98] X. Jia, K. Li, X. Li, and A. Zhang, "A novel semi-supervised deep learning framework for affective state recognition on eeg signals," in *IEEE International Conference on Bioinformatics and Bioengineering (BIBE)*, 2014, pp. 30–37.

[99] H. Xu and K. N. Plataniotis, "Affective states classification using eeg and semi-supervised deep learning approaches," in *Multimedia Signal Processing (MMSP), 2016 IEEE 18th International Workshop on*, 2016, pp. 1–6.

[100] X. Li, P. Zhang, D. Song, G. Yu, Y. Hou, and B. Hu, "Eeg based emotion identification using unsupervised deep feature learning," 2015.

[101] H. Xu and K. N. Plataniotis, "Eeg-based affect states classification using deep belief networks," in *Digital Media Industry & Academic Forum (DMIAF)*, 2016, pp. 148–153.

[102] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "Eeg-based emotion classification using deep belief networks," in *Multimedia and Expo (ICME), 2014 IEEE International Conference on*, 2014, pp. 1–6.

[103] K. Li, X. Li, Y. Zhang, and A. Zhang, "Affective state recognition from eeg with deep belief networks," in *2013 IEEE International Conference on Bioinformatics and Biomedicine*, 2013, pp. 305–310.

[104] J. A. Mioranda-Correa and I. Patras, "A multi-task cascaded network for prediction of affect, personality, mood and social context using eeg signals," in *Automatic Face & Gesture Recognition (FG 2018), 2018 13th IEEE International Conference on*, 2018, pp. 373–380.

[105] P. Kawde and G. K. Verma, "Deep belief network based affect recognition from physiological signals," in *Electrical, Computer and Electronics (UPCON), 2017 4th IEEE Uttar Pradesh Section International Conference on*, 2017, pp. 587–592.

[106] Y. Gao, H. J. Lee, and R. M. Mehmood, "Deep learninig of eeg signals for emotion recognition," in *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on*, 2015, pp. 1–5.

[107] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Computer methods and programs in biomedicine*, vol. 140, pp. 93–110, 2017.

[108] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on eeg using lstm recurrent neural network," *Emotion*, vol. 8, no. 10, 2017.

[109] Y. Yuan, G. Xun, F. Ma, Q. Suo, H. Xue, K. Jia, and A. Zhang, "A novel channel-aware attention framework for multi-channel eeg seizure detection via multi-view deep learning," in *International Conference on Biomedical & Health Informatics (BHI)*, 2018, pp. 206–209.

[110] S. S. Talathi, "Deep recurrent neural networks for seizure detection and early seizure detection systems," *arXiv preprint arXiv:1706.03283*, 2017.

[111] G. Ruffini, D. Ibañez, M. Castellano, S. Dunne, and A. Soria-Frisch, "Eeg-driven rnn classification for prognosis of neurode-generation in at-risk patients," in *International Conference on Artificial Neural Networks*, 2016, pp. 306–313.

[112] I. Ullah, M. Hussain, H. Aboalsamh *et al.*, "An automated system for epilepsy detection using eeg brain signals based on deep learning approach," *Expert Systems with Applications*, vol. 107, pp. 61–71, 2018.

[113] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, and D. P. Subha, "Automated eeg-based screening of depression using deep convolutional neural network," *Computer methods and programs in biomedicine*, vol. 161, pp. 103–113, 2018.

[114] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals," *Computers in biology and medicine*, vol. 100, pp. 270–278, 2018.

[115] F. C. Morabito, M. Campolo, C. Ieracitano, J. M. Ebadi, L. Bonanno, A. Bramanti, S. Desalvo, N. Mammone, and P. Bramanti, "Deep convolutional neural networks for classification of mild cognitive impaired and alzheimer's disease patients from scalp eeg recordings," in *Research and Technologies for Society and Industry Leveraging a better*

tomorrow (RTSI), 2016 IEEE 2nd International Forum on*, 2016, pp. 1–6.

[116] R. Schirrmeister, L. Gemein, K. Eggensperger, F. Hutter, and T. Ball, "Deep learning with convolutional neural networks for decoding and visualization of eeg pathology," in *Signal Processing in Medicine and Biology Symposium (SPMB), 2017 IEEE*, 2017, pp. 1–7.

[117] M.-P. Hosseini, T. X. Tran, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Deep learning with edge computing for localization of epileptogenicity using multimodal rs-fmri and eeg big data," in *Autonomic Computing (ICAC), 2017 IEEE International Conference on*, 2017, pp. 83–92.

[118] A. R. Johansen, J. Jin, T. Maszczyk, J. Dauwels, S. S. Cash, and M. B. Westover, "Epileptiform spike detection via convolutional neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 754–758.

[119] A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. De Vos, and S. Van Huffel, "Neonatal seizure detection using deep convolutional neural networks," *International journal of neural systems*, p. 1850011, 2018.

[120] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for eeg big data and seizure prediction bci via internet of things," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 392–404, 2017.

[121] Y. Yuan, G. Xun, K. Jia, and A. Zhang, "A novel wavelet-based model for eeg epileptic seizure detection using multi-context learning," in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 694–699.

[122] Q. Lin, S.-q. Ye, X.-m. Huang, S.-y. Li, M.-z. Zhang, Y. Xue, and W.-S. Chen, "Classification of epileptic eeg signals with stacked sparse autoencoder based on deep learning," in *International Conference on Intelligent Computing*, 2016, pp. 802–810.

[123] F. C. Morabito, M. Campolo, N. Mammone, M. Versaci, S. Franceschetti, F. Tagliavini, V. Sofia, D. Fatuzzo, A. Gambardella, A. Labate *et al.*, "Deep learning representation from electroencephalography of early-stage creutzfeldt-jakob disease and features for differentiation from rapidly progressive dementia," *International journal of neural systems*, vol. 27, no. 02, p. 1650039, 2017.

[124] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of eeg signals," *IEEE Access*, vol. 6, pp. 25 399–25 410, 2018.

[125] A. Page, J. Turner, T. Mohsenin, and T. Oates, "Comparing raw data and feature extraction for seizure detection with deep learning methods." in *FLAIRS Conference*, 2014.

[126] Y. Zhao and L. He, "Deep learning in the eeg diagnosis of alzheimer's disease," in *Asian Conference on Computer Vision*, 2014, pp. 340–353.

[127] J. Turner, A. Page, T. Mohsenin, and T. Oates, "Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection," in *2014 AAAI Spring Symposium Series*, 2014.

[128] V. Shah, M. Golmohammadi, S. Ziyabari, E. Von Weltin, I. Obeid, and J. Picone, "Optimizing channel selection for seizure detection," in *Signal Processing in Medicine and Biology Symposium (SPMB), 2017 IEEE*, 2017, pp. 1–5.

[129] M.-P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, and D. Pompili, "Cloud-based deep learning of big eeg data for epileptic seizure prediction," *arXiv preprint arXiv:1702.05192*, 2017.

[130] M. Golmohammadi, S. Ziyabari, V. Shah, S. L. de Diego, I. Obeid, and J. Picone, "Deep architectures for automated seizure detection in scalp eegs," *arXiv preprint arXiv:1712.09776*, 2017.

[131] A. M. Al-kaysi, A. Al-Ani, and T. W. Boonstra, "A multichannel deep belief network for the classification of eeg data," in *International Conference on Neural Information Processing*, 2015, pp. 38–45.

[132] S. M. Abdelfattah, G. M. Abdelrahman, and M. Wang, "Augmenting the size of eeg datasets using generative adversarial networks," in *2018 International Joint Conference on Neural Net-*

works (IJCNN)*, 2018, pp. 1–6.

[133] S. Palazzo, C. Spampinato, I. Kavasidis, D. Giordano, and M. Shah, "Generative adversarial networks conditioned by brain signals," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 3410–3418.

[134] P. S. S. C. G. D. S. M. Kavasidis, I., "Brain2image: Converting brain signals into images," in *Proceedings of the 25th ACM international conference on Multimedia*, 2017, pp. 1809–1817.

[135] J. Teo, C. L. Hou, and J. Mountstephens, "Deep learning for eeg-based preference classification," in *AIP Conference Proceedings*, vol. 1891, no. 1, 2017, p. 020141.

[136] T. K. Reddy and L. Behera, "Online eye state recognition from eeg data using deep architectures," in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*, 2016, pp. 000 712–000 717.

[137] A. J. Yepes, J. Tang, and B. S. Mashford, "Improving classification accuracy of feedforward neural networks for spiking neuromorphic chips," *arXiv preprint arXiv:1705.07755*, 2017.

[138] J. Shang, W. Zhang, J. Xiong, and Q. Liu, "Cognitive load recognition using multi-channel complex network method," in *International Symposium on Neural Networks*, 2017, pp. 466–474.

[139] J. Behncke, R. T. Schirrmeister, W. Burgard, and T. Ball, "The signature of robot action success in eeg signals of a human observer: Decoding and visualization using deep convolutional neural networks," in *International Conference on Brain-Computer Interface (BCI) 2018*, 2018, pp. 1–6.

[140] Y.-C. Hung, Y.-K. Wang, M. Prasad, and C.-T. Lin, "Brain dynamic states analysis based on 3d convolutional neural network," in *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on*, 2017, pp. 222–227.

[141] V. Baltatzis, K.-M. Bintsi, G. K. Apostolidis, and L. J. Hadjileontiadis, "Bullying incidences identification within an immersive environment using hd eeg-based analysis: A swarm decomposition and deep learning approach," *Scientific reports*, vol. 7, no. 1, p. 17292, 2017.

[142] S. Stober, D. J. Cameron, and J. A. Grahn, "Classifying eeg recordings of rhythm perception." in *ISMIR*, 2014, pp. 649–654.

[143] M. Völker, R. T. Schirrmeister, L. D. Fiederer, W. Burgard, and T. Ball, "Deep transfer learning for error decoding from non-invasive eeg," in *Brain-Computer Interface (BCI), 2018 6th International Conference on*, 2018, pp. 1–6.

[144] L. G. Hernández, O. M. Mozos, J. M. Ferrández, and J. M. Antelis, "Eeg-based detection of braking intention under different car driving conditions," *Frontiers in neuroinformatics*, vol. 12, 2018.

[145] M. A. Almogbel, A. H. Dang, and W. Kameyama, "Eeg-signals based cognitive workload detection of vehicle driver using deep learning," in *Advanced Communication Technology (ICACT), 2018 20th International Conference on*, 2018, pp. 256–259.

[146] M. J. Putten, S. Olbrich, and M. Arns, "Predicting sex from brain rhythms with deep learning," *Scientific reports*, vol. 8, no. 1, p. 3069, 2018.

[147] M. Hajinoroozi, Z. Mao, and Y. Huang, "Prediction of driver's drowsy and alert states from eeg signals with deep learning," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015 IEEE 6th International Workshop on*, 2015, pp. 493–496.

[148] A. Sternin, S. Stober, J. Grahn, and A. Owen, "Tempo estimation from the eeg signal during perception and imagination of music," in *International Symposium on Computer Music Multidisciplinary Research*, 2015.

[149] L. Chu, R. Qiu, H. Liu, Z. Ling, T. Zhang, and J. Wang, "Individual recognition in schizophrenia using deep learning methods with random forest and voting classifiers: Insights from resting state eeg streams," *arXiv preprint arXiv:1707.03467*, 2017.

[150] Z. Yin and J. Zhang, "Cross-session classification of mental workload levels using eeg and an adaptive deep learning model," *Biomedical Signal Processing and Control*, vol. 33, pp. 30–47, 2017.

[151] L.-H. Du, W. Liu, W.-L. Zheng, and B.-L. Lu, "Detecting driving fatigue with multimodal deep learning," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, 2017, pp. 74–77.

[152] S. Narejo, E. Pasero, and F. Kulsoom, "Eeg based eye state classification using deep belief network and stacked autoencoder," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 6, pp. 3131–3141, 2016.

[153] M. Hajinoroozi, T.-P. Jung, C.-T. Lin, and Y. Huang, "Feature extraction with deep belief networks for driver's cognitive states prediction from eeg data," in *Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on*, 2015, pp. 812–815.

[154] P. P. San, S. H. Ling, R. Chai, Y. Tran, A. Craig, and H. Nguyen, "Eeg-based driver fatigue detection using hybrid deep generic model," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, 2016, pp. 800–803.

[155] P. Li, W. Jiang, and F. Su, "Single-channel eeg-based mental fatigue detection based on deep belief network," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, 2016, pp. 367–370.

[156] P. Zhang, X. Wang, W. Zhang, and J. Chen, "Learning spatial–spectral–temporal eeg features with recurrent 3d convolutional neural networks for cross-task mental workload assessment," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 27, no. 1, pp. 31–42, 2018.

[157] S. Stober, A. Sternin, A. M. Owen, and J. A. Grahn, "Deep feature learning for eeg recordings," *arXiv preprint arXiv:1511.04306*, 2015.

[158] R. Chai, S. H. Ling, P. P. San, G. R. Naik, T. N. Nguyen, Y. Tran, A. Craig, and H. T. Nguyen, "Improving eeg-based driver fatigue classification using sparse-deep belief networks," *Frontiers in neuroscience*, vol. 11, p. 103, 2017.

[159] P. Bashivan, M. Yeasin, and G. M. Bidelman, "Single trial prediction of normal and excessive cognitive load through eeg feature fusion," in *Signal Processing in Medicine and Biology Symposium (SPMB), 2015 IEEE*, 2015, pp. 1–5.

[160] X. Zhang, L. Yao, C. Huang, S. S. Kanhere, and D. Zhang, "Brain2object: Printing your mind from brain signals with spatial correlation embedding," *arXiv preprint arXiv:1810.02223*, 2018.

[161] T. Koike-Akino, R. Mahajan, T. K. Marks, Y. Wang, S. Watanabe, O. Tuzel, and P. Orlik, "High-accuracy user identification using eeg biometrics," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 854–858.

[162] K. Kawasaki, T. Yoshikawa, and T. Furuhashi, "Visualizing extracted feature by deep learning in p300 discrimination task," in *Soft Computing and Pattern Recognition (SoCPaR), 2015 7th International Conference of*, 2015, pp. 149–154.

[163] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly, and M. Shah, "Deep learning human mind for automated visual classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6809–6817.

[164] M. Liu, W. Wu, Z. Gu, Z. Yu, F. Qi, and Y. Li, "Deep learning based on batch normalization for p300 signal detection," *Neurocomputing*, vol. 275, pp. 288–297, 2018.

[165] S. Sarkar, K. Reddy, A. Dorgan, C. Fidopiastis, and M. Giering, "Wearable eeg-based activity recognition in phm-related service environment via deep learning," *Int. J. Progn. Health Manag*, vol. 7, pp. 1–10, 2016.

[166] H. Cecotti and A. Graser, "Convolutional neural networks for p300 detection with application to brain-computer interfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 3, pp. 433–445, 2011.

[167] W. Gao, J.-a. Guan, J. Gao, and D. Zhou, "Multi-ganglion ann based feature learning with application to p300-bci signal classification," *Biomedical Signal Processing and Control*, vol. 18, pp. 127–137, 2015.

[168] Q. Liu, X.-G. Zhao, Z.-G. Hou, and H.-G. Liu, "Deep belief networks for eeg-based concealed information test," in *International Symposium on Neural Networks*, 2017, pp. 498–506.

[169] T. Ma, H. Li, H. Yang, X. Lv, P. Li, T. Liu, D. Yao, and P. Xu, "The extraction of motion-onset vep bci features based on deep learning and compressed sensing," *Journal of neuroscience methods*, vol. 275, pp. 80–92, 2017.

[170] R. Maddula, J. Stivers, M. Mousavi, S. Ravindran, and V. de Sa, "Deep recurrent convolutional neural networks for classifying p300 bci signals," in *Proceedings of the 7th Graz Brain-Computer Interface Conference, Graz, Austria*, 2017, pp. 18–22.

[171] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from eeg with deep recurrent-convolutional neural networks," *ICLR*, 2016.

[172] P. Bashivan, I. Rish, and S. Heisig, "Mental state recognition via wearable eeg," *arXiv preprint arXiv:1602.00985*, 2016.

[173] A. Shanbhag, A. P. Kholkar, S. Sawant, A. Vicente, S. Martires, and S. Patil, "P300 analysis using deep neural network," in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017, pp. 3142–3147.

[174] Z. Mao, V. Lawhern, L. M. Merino, K. Ball, L. Deng, B. J. Lance, K. Robbins, and Y. Huang, "Classification of non-time-locked rapid serial visual presentation events for brain-computer interaction using deep learning," in *Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, 2014, pp. 520–524.

[175] Z. Mao, "Deep learning for rapid serial visual presentation event from electroencephalography signal," Ph.D. dissertation, The University of Texas at San Antonio, 2016.

[176] R. Manor and A. B. Geva, "Convolutional neural network for multi-category rapid serial visual presentation bci," *Frontiers in computational neuroscience*, vol. 9, p. 146, 2015.

[177] H. Cecotti, "Convolutional neural networks for event-related potential detection: impact of the architecture," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 2031–2034.

[178] A. J. Solon, S. M. Gordon, B. Lance, and V. Lawhern, "Deep learning approaches for p300 classification in image triage: Applications to the nails task," in *Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, NTCIR-13, Tokyo, Japan*, 2017, pp. 5–8.

[179] M. Hajinoroozi, Z. Mao, Y.-P. Lin, and Y. Huang, "Deep transfer learning for cross-subject and cross-experiment prediction of image rapid serial visual presentation events from eeg data," in *International Conference on Augmented Cognition*, 2017, pp. 45–55.

[180] Z. Mao, W. X. Yao, and Y. Huang, "Eeg-based biometric identification with deep learning," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, 2017, pp. 609–612.

[181] R. Manor, L. Mishali, and A. B. Geva, "Multimodal neural network for rapid serial visual presentation brain computer interface," *Frontiers in computational neuroscience*, vol. 10, p. 130, 2016.

[182] Z. Lin, Y. Zeng, L. Tong, H. Zhang, C. Zhang, and B. Yan, "Method for enhancing single-trial p300 detection by introducing the complexity degree of image information in rapid serial visual presentation tasks," *PloS one*, vol. 12, no. 12, p. e0184713, 2017.

[183] S. M. Gordon, M. Jaswa, A. J. Solon, and V. J. Lawhern, "Real world bci: cross-domain learning and practical applications," in *Proceedings of the 2017 ACM Workshop on An Application-oriented Approach to BCI out of the laboratory*, 2017, pp. 25–28.

[184] J. Yoon, J. Lee, and M. Whang, "Spatial and time domain feature of erp speller system extracted via convolutional neural network," *Computational intelligence and neuroscience*, vol. 2018, 2018.

[185] J. Shamwell, H. Lee, H. Kwon, A. R. Marathe, V. Lawhern, and W. Nothwang, "Single-trial eeg rsvp classification using convolutional neural networks," in *Micro-and Nanotechnology*

Sensors, Systems, and Applications VIII, vol. 9836, 2016, p. 983622.

[186] L. Vařeka and P. Mautner, "Stacked autoencoders for the p300 component detection," *Frontiers in neuroscience*, vol. 11, p. 302, 2017.

[187] E. Carabez, M. Sugi, I. Nambu, and Y. Wada, "Identifying single trial event-related potentials in an earphone-based auditory brain-computer interface," *Applied Sciences*, vol. 7, no. 11, p. 1197, 2017.

[188] S. Stober, D. J. Cameron, and J. A. Grahn, "Using convolutional neural networks to recognize rhythm stimuli from electroencephalography recordings," in *Advances in neural information processing systems*, 2014, pp. 1449–1457.

[189] A. HACHEM, M. M. B. Khelifa, A. M. Alimi, P. Gorce, S. V. ARASU, S. BAULKANI, S. K. BISOY, P. K. PATTNAIK, S. RAVINDRAN, N. PALANISAMY *et al.*, "Effect of fatigue on ssvep during virtual wheelchair navigation," *Journal of Theoretical and Applied Information Technology*, vol. 65, no. 1, 2014.

[190] J. Thomas, T. Maszczyk, N. Sinha, T. Kluge, and J. Dauwels, "Deep learning-based classification for brain-computer interfaces," in *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on*, 2017, pp. 234–239.

[191] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PloS one*, vol. 12, no. 2, p. e0172578, 2017.

[192] N. R. Waytowich, V. Lawhern, J. O. Garcia, J. Cummings, J. Faller, P. Sajda, and J. M. Vettel, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *arXiv preprint arXiv:1803.04566*, 2018.

[193] N. K. N. Aznan, S. Bonner, J. D. Connolly, N. A. Moubayed, and T. P. Breckon, "On the classification of ssvep-based dry-eeg signals via convolutional neural networks," *arXiv preprint arXiv:1805.04157*, 2018.

[194] T. Tu, J. Koss, and P. Sajda, "Relating deep neural network representations to eeg-fmri spatiotemporal dynamics in a perceptual decision-making task," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1985–1991.

[195] J. Kulasingham, V. Vibujithan, and A. De Silva, "Deep belief networks and stacked autoencoders for the p300 guilty knowledge test," in *Biomedical Engineering and Sciences (IECBES), 2016 IEEE EMBS Conference on*, 2016, pp. 127–132.

[196] M. Attia, I. Hettiarachchi, M. Hossny, and S. Nahavandi, "A time domain classification of steady-state visual evoked potentials using deep recurrent-convolutional neural networks," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*, 2018, pp. 766–769.

[197] J. Pérez-Benítez, J. Pérez-Benítez, and J. Espina-Hernández, "Development of a brain computer interface interface using multifrequency visual stimulation and deep neural networks," in *Electronics, Communications and Computers (CONIELECOMP), 2018 International Conference on*, 2018, pp. 18–24.

[198] G. Huve, K. Takahashi, and M. Hashimoto, "Brain activity recognition with a wearable fnirs using neural networks," in *Mechatronics and Automation (ICMA), 2017 IEEE International Conference on*, 2017, pp. 1573–1578.

[199] ——, "Brain-computer interface using deep neural network and its application to mobile robot control," in *Advanced Motion Control (AMC), 2018 IEEE 15th International Workshop on*, 2018, pp. 169–174.

[200] J. Hennrich, C. Herff, D. Heger, and T. Schultz, "Investigating deep learning for fnirs based bci." in *EMBC*, 2015, pp. 2844–2847.

[201] T. Hiroyasu, K. Hanawa, and U. Yamamoto, "Gender classification of subjects from cerebral blood flow changes using deep learning," in *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on*, 2014, pp. 229–233.

[202] S. Koyamada, Y. Shikauchi, K. Nakae, M. Koyama, and S. Ishii,

"Deep learning of fmri big data: a novel approach to subject-transfer decoding," *arXiv preprint arXiv:1502.00093*, 2015.

[203] G. Shen, T. Horikawa, K. Majima, and Y. Kamitani, "Deep image reconstruction from human brain activity," *PLoS computational biology*, vol. 15, no. 1, p. e1006633, 2019.

[204] R. M. Cichy, A. Khosla, D. Pantazis, A. Torralba, and A. Oliva, "Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence," *Scientific reports*, vol. 6, p. 27755, 2016.

[205] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Medical image analysis*, vol. 35, pp. 18–31, 2017.

[206] V. Shreyas and V. Pankajakshan, "A deep learning architecture for brain tumor segmentation in mri images," in *Multimedia Signal Processing (MMSP), 2017 IEEE 19th International Workshop on*, 2017, pp. 1–6.

[207] S. Sarraf and G. Tofighi, "Deep learning-based pipeline to recognize alzheimer's disease using fmri data," in *Future Technologies Conference (FTC)*, 2016, pp. 816–820.

[208] R. Li, W. Zhang, H.-I. Suk, L. Wang, J. Li, D. Shen, and S. Ji, "Deep learning based imaging data completion for improved brain disease diagnosis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2014, pp. 305–312.

[209] H.-I. Suk, C.-Y. Wee, S.-W. Lee, and D. Shen, "State-space model with deep learning for functional dynamics estimation in resting-state fmri," *NeuroImage*, vol. 129, pp. 292–307, 2016.

[210] H.-I. Suk, D. Shen, A. D. N. Initiative *et al.*, "Deep learning in diagnosis of brain disorders," in *Recent Progress in Brain and Cognitive Engineering*, 2015, pp. 203–213.

[211] S. M. Plis, D. R. Hjelm, R. Salakhutdinov, E. A. Allen, H. J. Bockholt, J. D. Long, H. J. Johnson, J. S. Paulsen, J. A. Turner, and V. D. Calhoun, "Deep learning for neuroimaging: a validation study," *Frontiers in neuroscience*, vol. 8, p. 229, 2014.

[212] A. Ortiz, J. Munilla, J. M. Gorriz, and J. Ramirez, "Ensembles of deep learning architectures for the early diagnosis of the alzheimer's disease," *International journal of neural systems*, vol. 26, no. 07, p. 1650025, 2016.

[213] N. F. M. Suhaimi, Z. Z. Htike, and N. K. A. M. Rashid, "Studies on classification of fmri data using deep learning approach," 2015.

[214] K. Seeliger, U. Güçlü, L. Ambrogioni, Y. Güçlütürk, and M. Van Gerven, "Generative adversarial networks for reconstructing natural images from brain activity," *NeuroImage*, vol. 181, pp. 775–785, 2018.

[215] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri, and H. Nakayama, "Gan-based synthetic brain mr image generation," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 734–738.

[216] P. Zhang, F. Wang, W. Xu, and Y. Li, "Multi-channel generative adversarial network for parallel magnetic resonance image reconstruction in k-space," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 180–188.

[217] C. Hu, R. Ju, Y. Shen, P. Zhou, and Q. Li, "Clinical decision support for alzheimer's disease based on deep learning and brain network," in *Communications (ICC), 2016 IEEE International Conference on*, 2016, pp. 1–6.

[218] P. Garg, E. Davenport, G. Murugesan, B. Wagner, C. Whitlow, J. Maldjian, and A. Montillo, "Automatic 1d convolutional neural network-based detection of artifacts in meg acquired without electrooculography or electrocardiography," in *Pattern Recognition in Neuroimaging (PRNI), 2017 International Workshop on*, 2017, pp. 1–4.

[219] M. Shu and A. Fyshe, "Sparse autoencoders for word decoding from magnetoencephalography," in *Proceedings of the third NIPS Workshop on Machine Learning and Interpretation in NeuroImaging (MLINI)*, 2013.

[220] A. Hasasneh, N. Kampel, P. Sripad, N. J. Shah, and J. Dammers, "Deep learning approach for automatic classification of ocular and cardiac artifacts in meg data," *Journal of Engineering*, vol. 2018, 2018.

[221] Y. Gordienko, S. Stirenko, Y. Kochura, O. Alienin, M. Novotarskiy, and N. Gordienko, "Deep learning for fatigue estimation on the basis of multimodal human-machine interactions," *arXiv preprint arXiv:1801.06048*, 2017.

[222] G. Pfurtscheller and F. L. Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.

[223] P. Khurana, A. Majumdar, and R. Ward, "Class-wise deep dictionaries for eeg classification," in *International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 3556–3563.

[224] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain–computer interface based on the fusion of p300 and ssvep scores," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 693–701, 2015.

[225] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Briefings in bioinformatics*, vol. 18, no. 5, pp. 851–869, 2017.

[226] S. Sarraf, G. Tofighi *et al.*, "Deepad: Alzheimer's disease classification via deep convolutional neural networks using mri and fmri," *bioRxiv*, p. 070441, 2016.

[227] L. Marc Moreno, "Deep learning for brain tumor segmentation," *Master diss. University of Colorado Colorado Springs.*, 2017.

[228] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *International Conference on Learning Representations (ICLR)*, 2016.

[229] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International Conference on Machine Learning (ICML)*, 2017, pp. 214–223.

[230] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial eeg data," in *Machine Learning for Signal Processing (MLSP), 2016 IEEE 26th International Workshop on*, 2016, pp. 1–6.

[231] A. Antoniades, L. Spyrou, D. Martin-Lopez, A. Valentin, G. Alarcon, S. Sanei, and C. C. Took, "Deep neural architectures for mapping scalp to intracranial eeg," *International journal of neural systems*, p. 1850009, 2018.

[232] R. Parasuraman and Y. Jiang, "Individual differences in cognition, affect, and performance: Behavioral, neuroimaging, and molecular genetic approaches," *Neuroimage*, vol. 59, no. 1, pp. 70–82, 2012.

[233] L. Deng, "Three classes of deep learning architectures and their applications: a tutorial survey," *APSIPA transactions on signal and information processing*, 2012.

[234] W.-L. Zheng and B.-L. Lu, "Personalizing eeg-based affective models with transfer learning," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016, pp. 2732–2738.

[235] X. Zhang, L. Yao, and F. Yuan, "Adversarial variational embedding for robust semi-supervised learning," in *SIGKDD 2019*, 2019.

[236] S. Koyama, S. M. Chase, A. S. Whitford, M. Velliste, A. B. Schwartz, and R. E. Kass, "Comparison of brain–computer interface decoding algorithms in open-loop and closed-loop control," *Journal of computational neuroscience*, vol. 29, no. 1-2, pp. 73–87, 2010.

[237] S. Aliakbaryhosseinabadi, E. N. Kamavuako, N. Jiang, D. Farina, and N. Mrachacz-Kersting, "Online adaptive synchronous bci system with attention variations," in *Brain-Computer Interface Research*, 2019, pp. 31–41.

[238] E. K. Kalunga, S. Chevallier, Q. Barthélemy, K. Djouani, E. Monacelli, and Y. Hamam, "Online ssvep-based bci using riemannian geometry," *Neurocomputing*, vol. 191, pp. 55–68, 2016.

[239] M. Pacharra, S. Debener, and E. Wascher, "Concealed around-the-ear eeg captures cognitive processing in a visual simon task," *Frontiers in human neuroscience*, vol. 11, p. 290, 2017.

[240] K. B. Mikkelsen, S. L. Kappel, D. P. Mandic, and P. Kidmose, "Eeg recorded from the ear: Characterizing the ear-eeg method," *Frontiers in neuroscience*, vol. 9, p. 438, 2015.

[241] S. B. Rutkove, "Introduction to volume conduction," in *The clinical neurophysiology primer*, 2007, pp. 43–53.

[242] B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq, and F. Vidal, "Spatial and temporal resolutions of eeg: Is it really black and white? a scalp current density view," *International Journal of Psychophysiology*, vol. 97, no. 3, pp. 210–220, 2015.

[243] J. Malmivuo, R. Plonsey *et al.*, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*, 1995.

[244] M. Tortella-Feliu, A. Morillas-Romero, M. Balle, J. Llabrés, X. Bornas, and P. Putman, "Spontaneous eeg activity and spontaneous emotion regulation," *International Journal of Psychophysiology*, vol. 94, no. 3, pp. 365–372, 2014.

[245] A. Salek-Haddadi, K. Friston, L. Lemieux, and D. Fish, "Studying spontaneous eeg activity with fmri," *Brain research reviews*, vol. 43, no. 1, pp. 110–133, 2003.

[246] A. Ikeda and Y. Washizawa, "Spontaneous eeg classification using complex valued neural network," in *International Conference on Neural Information Processing*, 2019, pp. 495–503.

[247] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: a review," *Journal of vision*, vol. 15, no. 6, pp. 4–4, 2015.

[248] S. Lees, N. Dayan, H. Cecotti, P. Mccullagh, L. Maguire, F. Lotte, and D. Coyle, "A review of rapid serial visual presentation-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 2, p. 021001, 2018.

[249] K. H. Chiappa, *Evoked potentials in clinical medicine*, 1997.

[250] V. Mayya, B. Mainsah, and G. Reeves, "Information-theoretic analysis of refractory effects in the p300 speller," *arXiv preprint arXiv:1701.03313. Non-exclusive-distrib License. https://arxiv.org/licenses/nonexclusive-distrib/1.0/license.html*, 2017.

[251] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, "How many people are able to control a p300-based brain–computer interface (bci)?" *Neuroscience letters*, vol. 462, no. 1, pp. 94–98, 2009.

[252] A. Belitski, J. Farquhar, and P. Desain, "P300 audio-visual speller," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025022, 2011.

[253] M. Welvaert and Y. Rosseel, "On the definition of signal-to-noise ratio and contrast-to-noise ratio for fmri data," *PloS one*, vol. 8, no. 11, 2013.

[254] R. M. Cichy, A. Khosla, D. Pantazis, and A. Oliva, "Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks," *Neuroimage*, vol. 153, pp. 346–358, 2017.

[255] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[256] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[257] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[258] P. O. Glauner, "Comparison of training methods for deep neural networks," *arXiv preprint arXiv:1504.06825*, 2015.

[259] G. St-Yves and T. Naselaris, "Generative adversarial networks conditioned on brain activity reconstruct seen images," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018, pp. 1054–1061.

Appendices

A. Non-invasive Brain Signals

Here, we present a detailed introduction of brain signals as shown in Figure 2. Non-invasive brain imaging technique can be collected using electrical, magnetic or metabolic methods, which mainly include Electroencephalogram (EEG), Functional near-infrared spectroscopy (fNIRS), Functional magnetic resonance imaging (fMRI), and Magnetoencephalography (MEG).

A.1. Electroencephalography (EEG)

Electroencephalography (EEG) is the most commonly used non-invasive technique for measuring brain activities. EEG monitors the voltage fluctuations generated by an electrical current within human neurons. Electrodes placed on the scalp measure the amplitude of EEG signals. EEG signals have a low spatial resolution due to the effect of volume conduction which refers to the complex effects of measuring electrical potentials a distance from the source generators [241, 242]. EEG electrode locations generally follow the international 10-20 system [243]. The specific placement of electrodes is presented in Figure 5 [10]. The EEG signals are collected while the subject is undertaking imagination task. Each line represents the signal stream collected from a single EEG electrode (also called 'channel') over time.

The temporal resolution of EEG signals is much better than the spatial resolution. The ionic current changes rapidly, which offers a temporal resolution higher than 1000 Hz. The SNR of EEG is generally very poor due to both objective and subjective factors. Objective factors include environmental noises, the obstruction of the skull and other tissues between cortex and scalp, and different stimulations. Subjective factors contain the subject's mental stage, fatigue status, the variance among different subjects, and so on.

EEG recording equipment can be installed in a cap-like headset. The EEG headset can be mounted on the user's head to gather signals. Compared to other equipment used to measure brain signals, EEG headsets are portable and more accessible for most applications.

The EEG signals collected from any typical EEG hardware have several non-overlapping frequency bands (Delta, Theta, Alpha, Beta, and Gamma) based on the strong intra-band correlation with a distinct behavioral state [10]. Each EEG pattern contains signals associated with particular brain information. Table 7 shows EEG frequency patterns and the corresponding characteristics. Here, the degree of awareness denotes the perception of individuals when presented with external stimuli.

Compared to other signals (e.g., fMRI, fNIRS, MEG), EEG has several important advantages: 1) the hardware has higher portability with much lower price; 2) the temporal resolution is very high (milliseconds level). Among other non-invasive techniques, only MEG has the same level of temporal resolution; 3) EEG is relatively tolerant of subject movement and artifacts, which can be minimized by existing signal processing methods; 4) the subject doesn't need to be exposed to high-intensity (>1 Tesla) magnetic fields, therefore, EEG can serve subjects that have metal implants in their body (such as metal-containing pacemakers).

As the most commonly used signals, there are a large number of sub-classes of EEG signals. In this section, we present a methodical introduction of EEG sub-class signals. As shown in Figure 2, we divided EEG signals into spontaneous EEG and evoked potentials. Evoked potentials can be split into event-related potentials and steady-state evoked potentials based on the frequency of the external stimuli [7]. Each potential contains visual-, auditory-, and somatosensory- potentials based on the external stimuli types. The dashed quadrilaterals in Figure 2, such as Intracortical, SEP, SSAEP, SSSEP, and RSAP, are not included in this survey because there are very few existing studies working on them with deep learning algorithms. We list these signals for systematic completeness.
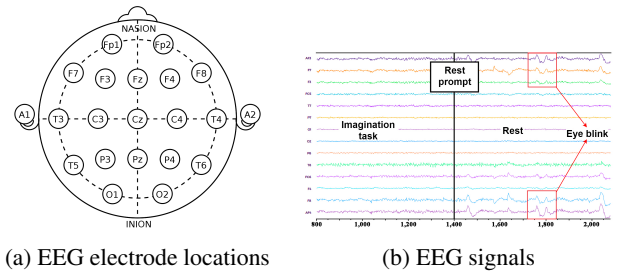

(a) EEG electrode locations   (b) EEG signals

Figure 5: EEG electrode locations on scalp (10-20 system) and the gathered EEG signals [10]. The electrodes' names are marked by their position: Fp (pre-frontal), F (frontal), T (temporal), P (parietal), O (occipital), and C (central).

Table 7: EEG patterns and corresponding characters. Awareness Degree denotes the degree of being aware of an external world. The awareness degree mentioned here is mainly defined in physiology instead of psychology.

Patterns	Frequency (Hz)	Amplitude	Brain State	Awareness Degree	Produced Location
Delta	0.5-4	Higher	Deep sleep pattern	Lower	Frontally and posteriorly
Theta	4-8	High	Light sleep pattern	Low	Entorhinal cortex, hippocampus
Alpha	8-12	Medium	Closing the eyes, relax state	Intermediate	Posterior regions of head
Beta	12-30	Low	Active thinking, focus, high alert, anxious	High	Most evident frontally, motor areas
Gamma	30-100	Lower	During cross-modal sensory processing	Higher	Somatosensory, auditory cortices

A.1.1. Spontaneous EEG Typically, when we talk about the term 'EEG,' we refer to *spontaneous* EEG which measures the brain signals under a specific state without external stimulation [244, 245, 246]. In particular, spontaneous EEG includes the EEG signals while the individual is sleeping, undertaking a mental task (e.g., counting), suffering brain disorders, undertaking motor imagery tasks, in a certain emotion, etc.

The EEG signals recorded while a user stares at a color/shape/image belong to this category. While the subject is gazing at a specific image, the visual stimuli are steady without any change. This scenario differs from the visual stimuli in evoked potential, where the visual stimuli are changing at a specific frequency. Thus, we regard the image stimulation as a particular state and regard it as spontaneous EEG. Spontaneous EEG-based systems are challenging to train, due to the lower SNR and the larger variation across subjects [35].

According to the gathering scenarios, the spontaneous EEG contains several subordinates: sleeping, motor imagery, emotional, mental disease and others.

A.1.2. Evoked Potential (EP) Evoked Potentials (EP) or evoked responses refers to the EEG signals which are evoked by an external stimulus instead of spontaneously. An EP is time-locked to the external stimulus while the aforementioned spontaneous EEG is non-time-locked. In contrast to spontaneous EEG, EP generally has higher amplitude and lower frequency. As a result, the EP signals are more robust across subjects. According to the stimulation method, there exist two categories of EP: the Event-Related Potential (ERP) and the Steady State Evoked Potential (SSEP). ERP records the EEG signals in response to an isolated discrete stimulus event (or event change). To achieve this isolation, stimuli in an ERP experiment are typically separated from each other by a long inter-stimulus interval, allowing for the estimation of a stimulus-independent baseline reference [247]. The stimuli frequency of ERP is generally lower than 2 Hz. In contrast, SSEP is generated in response to a periodic stimulus at a fixed rate. The stimuli frequency of SSEP generally ranges within 3.5-75 Hz.

Event-related potential (ERP). There are three kinds of evoked potentials in extensive research and clinical use: Visual Evoked Potentials (VEP); Auditory Evoked Potentials (AEP); and Somatosensory Evoked Potentials (SEP) [28]. The VEP signals are mainly on the occipital lobe, and the highest signal amplitudes are collected at the Calcarine sulcus.

1) Visual Evoked Potentials (VEP). Visual Evoked Potentials are a specific category of ERP which is caused by visual stimulus (e.g., an alternating checkerboard pattern on a computer screen). VEP signals are hidden within the normal spontaneous EEG. To separate VEP signals from the background EEG readings, repetitive stimulation and time-locked signal-averaging techniques are generally employed.

Rapid Serial Visual Presentation (RSVP) [248] can be regarded as one kind of VEP. An RSVP diagram is commonly used to examine the temporal characteristics of attention. The subject is required to stare at a screen where a series of items (e.g., images) are presented one-by-one. There is a specific item (called the target) separates from the rest of the other items (called distracters). The subject knows which is the target before the RSVP experiment. For instance, the distracters can be a color change or letters among numbers. RSVP contains a static mode (the items appear on the screen and then disappear without moving) and a moving mode (the items appear on the screen, move to another place, and finally disappear). Nowadays, brain signal research mainly focuses on the static mode RSVP. Usually, the frequency of RSVP is 10Hz which means that each item will stay on the screen for 0.1 seconds.

2) Auditory Evoked Potentials (AEP). Auditory Evoked Potentials are a specific subclass of ERP in which responses to auditory (sound) stimuli are recorded. AEP is mainly recorded from the scalp but originates at the brainstem or cortex. The most common AEP measured is the auditory brainstem response (ABR) which is generally employed to test the hearing ability of newborns and infants. In the brain signal area, AEP is mainly used in clinical tests for its accuracy and reliability in detecting unilateral loss [249]. Similar to RSVP, Rapid Serial Auditory Presentation (RSAP) refers to the experiments with rapid serial presentation of sound stimuli. The task for the subject is to recognize the target audio among the distracters.

3) Somatosensory Evoked Potentials (SEP).[21] Somatosensory Evoked Potentials are another commonly used

---

[21]Generally, Somatosensory Evoked Potentials is abbreviated as SSEP or SEP. In this paper, we choose SEP as the abbreviation in case of the conflict with Steady-State Evoked Potentials (SSEP).

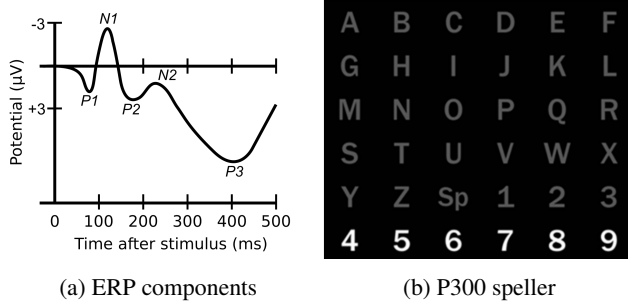(a) ERP components (b) P300 speller

Figure 6: P300 waves and visual-based P300 speller [250].

subcategory of ERP which is elicited by electrical stimulation of the peripheral nerves. SEP signals conclude a series of amplitude deflection that can be elicited by virtually any sensory stimuli.

P300. P300 (also called P3) is an important component in ERP [251]. Here we introduce P300 signal separately since it is widely-used in brain signal analysis. Figure 6a shows the ERP signal fluctuation in the 500 ms after the stimuli onset. The waveform mainly concludes five components, P1, N1, P2, N2, and P3. The capital character P/N represents positive/negative electrical potentials. The following number refers to the occurrence time of the specific potential. Thus, P300 denotes the positive potential of ERP waveform at approximately 300 ms after the presented stimuli. Compared to other components, P300 has the highest amplitude and is easiest to detect. Thus, a large number of brain signal studies focus on P300 analysis. P300 is more of an informative feature instead of a type of brain signal (e.g., VEP). Therefore, we do no list P300 in Figure 2. P300 can be analyzed in most of ERP signals such as VEP, AEP, SEP.

In practice, P300 can be elicited by rare, task-relevant events in an 'oddball' paradigm (e.g., P300 speaker). In the oddball paradigm, the subject receives a series of stimuli where low-probability target items are mixed with high-probability non-target items. Visual and auditory stimuli are the most commonly used in the oddball paradigm. Figure 6b shows an example of visual-based P300 speller which enables the subject the spell letters/numbers directly through brain signals [250]. The 26 letters of the alphabet and the Arabic numbers are displayed on a computer screen which serves as the keyboard. The subject focuses attention successively on the characters they wish to spell. The computer detects the chosen character online in real time. This detection is achieved by repeatedly flashing rows and columns of the matrix. When the elements containing the selected characters are flashing, a P300 fluctuation is elicited. In the 6×6 matrix screen, the rows and columns flash in mixed random order. The flash duration and interval among adjacent flashes are generally set as 100 ms [252]. The columns and rows flash separately. First, the columns

flash six times with each column flashing one time. Second, the rows will flash for six times. After that, this paradigm repeats for several times (e.g., N times). The P300 signals of the total $12N$ flash will be analyzed to output a single outcome (i.e., one letter/number).

Steady State Evoked Potentials (SSEP). Steady State Evoked Potential is another subcategory of evoked potentials, which are periodic cortical responses evoked by certain repetitive stimuli with a constant frequency. It has been demonstrated that the brain oscillations generally maintain a steady level over time while the potentials are evoked by steady state stimuli (e.g., a flickering light with fixed frequency). Technically, SSEP is defined as a form of response to repetitive sensory stimulation in which the constituent frequency components of the response remain constant over time in both amplitude and phase [37]. Depending on the type of stimuli, SSEP divides into three subcategories: Steady-State Visually Evoked Potentials (SSVEP), Steady-State Auditory Evoked Potentials (SSAEP), and Steady-State Somatosensory Evoked Potentials (SSSEP). In the brain signal area, most studies are focused on visual evoked steady potentials, and only rarely do papers focus on auditory and somatosensory stimuli. Therefore, in this survey, we mainly introduce SSVEP rather than SSAEP and SSSEP.

Commonly Used Visual-Related Potentials. Visual evoked potentials are the most common used potentials. Therefore, it is essential to distinguish the three different visual evoked potential paradigms: VEP, RSVP, SSVEP. Here, we theoretically introduce the characteristics of each paradigm and then give three demonstration videos to provide a better understanding. First, the frequencies are different: the frequency of VEP is less than 2Hz while the frequency of RSVP is around 10Hz, and the frequency of SSVEP ranges from 3.5 to 75Hz. Second, they have various presentation protocols. In the VEP paradigm, different visual patterns will be presented on the screen to check the user's brain signals changes. For instance, in this video[22], the image pattern is full of the screen and changes dramatically. In RSVP diagram, several items will be presented on a screen one-by-one. All the items are shown in the same place and share the same frequency. For example, the video[23] shows an RSVP scenario which is called speed reading. In SSVEP paradigm, several items will be presented on a screen at the same time while the items are shown at *variant positions* with different frequencies. For example, in this demonstration video[24], there are four circles distributed on the up, down, left, and right sides of a screen and the frequency of each item differs from each other.

---

[22]https://www.youtube.com/watch?v=iUW_l5YAEEM
[23]https://www.youtube.com/watch?v=5yddeRrd0hA&t=36s
[24]https://www.youtube.com/watch?v=t96rl1SFHlI

A.2. Functional Near-infrared Spectroscopy (fNIRS)

Functional near-infrared spectroscopy (fNIRS) is a non-invasive functional neuro-imaging technology using near-infrared (NIR) light [38]. In specific, fNIRS employs NIR light to measure the aggregation degree of oxygenated hemoglobin (Hb) and deoxygenated-hemoglobin (deoxy-Hb) because Hb and deoxy-Hb have higher absorbence of light than other head components such as the skull and scalp. fNIRS relies on blood-oxygen-level-dependent (BOLD) response or hemodynamic response to form a functional neuro-image. The BOLD response can detect the oxygenated or deoxygenated blood level in the brain blood. The relative levels reflect the blood flow and neural activation, where increased blood flow implies a higher metabolic demand caused by active neurons. For example, when the user is concentrating on a mental task, the prefrontal cortex neurons will be activated, and the BOLD response in the prefrontal cortex area will be stronger [200].

Single or multiple emitter-detector pairs measure the Hb and deoxy-Hb: the emitter transmits NIR light through the blood vessels to the detector. Most existing studies use fNIRS technologies to measure the status of prefrontal and motor cortex. The former response to mental tasks and music/image imagery while the latter is a response to motor-related tasks (e.g., motor imagery). The monitored Hb and deoxy-Hb change slowly since the blood speed varies in a relatively slow ratio compared to electrical signals. Temporal resolution refers to the smallest time of neural activity reliably separated by the signal. The fNIRS has lower temporal resolution compared with electrical or magnetic signals. The spatial resolution depends on the number of emitter-detector pairs. In current studies, three emitters and eight detectors would suffice for adequately acquiring the prefrontal cortex signals; and six emitters and six detectors would suffice for covering the motor cortex area [29]. fNIRS has a drawback in that it cannot be used to measure cortical activity occurring deeper than 4cm in the brain, due to the limitations in light emitter power and spatial resolution.

A.3. Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) monitors brain activities by detecting changes associated with blood flow in brain areas [14]. Similar to fNIRS, fMRI relies on the BOLD response. The main differences between fNIRS and fMRI are as follows [24]. First, as the name implies, fMRI measures BOLD response through magnetic instead of optical methods. Hemoglobin differs in how it responds to magnetic fields, depending on whether it has a bound oxygen molecule. The magnetic fields are more sensitive to and are more easily distorted by deoxy-Hb than Hb molecules. Second, the magnetic fields have higher penetration than NIR light, which gives fMRI greater ability to capture information from deep parts of the brain than fNIRS. Third, fMRI has a higher spatial resolution than fNIRS since the latter's spatial resolution is limited by the emitter-detector pairs. However, the temporal resolutions of fMRI and fNIRS are at an equal level because they both constrained by the blood flow speed.

fMRI has several flaws compared to fNIRS: 1) fMRI requires an expensive scanner to generate magnetic fields; 2) the scanner is heavy and has poor portability. In order to measure the signal of interest, CNR (Contrast-to-Noise Ratio) has been investigated to measure the image quality of fMRI because researchers are more interested in the contrast between images rather than the raw images. So for fMRI data, using the CNR of the time series instead of (t)SNR is more preferred because CNR compares a measure of the activation fluctuations to the noise [253].

A.4. Magnetoencephalography (MEG)

Magnetoencephalography (MEG) is a functional neuroimaging technique for mapping brain activity by recording magnetic fields produced by electrical currents occurring naturally in the brain, using very sensitive magnetometers [254]. The ionic currents of active neurons will create weak magnetic fields. The generated magnetic fields can be measured by magnetometers like SQUIDs (superconducting quantum interference devices). However, producing a detectable magnetic field requires massive (e.g., 50,000) active neurons with similar orientation. The source of the magnetic field measured by MEG is the pyramidal cells which are perpendicular to the cortex surface.

MEG has a relatively low spatial resolution since the signal quality highly depends on the measurement factors (e.g., brain area, neuron orientations, neuron depth). However, MEG can provide very high temporal resolution ($\geq$1000Hz) since MEG directly monitors the brain activity from the neuron level, which is in the same level of intracortical signals. MEG equipment is expensive and not portable which limits its real-world deployment.

B. Basic Deep Learning in Brain Signal Analysis

In this part, we will give relative detail introduction of various deep learning models for the reason that a part of the potential readers who are from non-computer area (e.g., biomedical) are not familiar to deep learning.

For simplification, we first define an operation $\mathcal{T}(\cdot)$ as

$$\mathcal{T}(\boldsymbol{x}) = \boldsymbol{w} * \boldsymbol{x} + \boldsymbol{b} \quad (1)$$

$$\mathcal{T}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{w} * \boldsymbol{x} + \boldsymbol{b} + \boldsymbol{w}' * \boldsymbol{x}' + \boldsymbol{b}' \quad (2)$$

where $\boldsymbol{x}$ and $\boldsymbol{x}'$ denote two variables while $\boldsymbol{w}$, $\boldsymbol{w}'$, $\boldsymbol{b}$, and $\boldsymbol{b}'$ denote the corresponding weights and basis.

B.1. Discriminative Deep Learning Models

Since the main task of brain signal analysis is brain signal recognition, the discriminative deep learning models are the most popular and powerful algorithms. Suppose we have a dataset of brain signal samples $\{\mathbb{X}, \mathbb{Y}\}$ where $\mathbb{X}$ denotes the set of brain signal observations and $\mathbb{Y}$ denotes the set of sample ground truth (i.e., labels). Suppose an specific sample-label pair $\{\boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{y} \in \mathbb{R}^M\}$ where N and M denote the dimension of observations and the number of sample categories, respectively. The aim of discriminative deep learning models is to learn a function with the mapping: $\boldsymbol{x} \rightarrow \boldsymbol{y}$. In short, the discriminative models receive the input data and output the corresponding category or label. All the discriminative models introduced in this section are supervised learning techniques which require the information of both the observations and the ground truth.

B.1.1. Multi-Layer Perceptron (MLP) The most basic neural network is fully-connected neural networks (Figure 7a) which only contains one hidden layer. The input layer receives the raw data or extracted features of brain signals while the output layer shows the classification results. The term 'fully-connected' denotes each node in a specific layer is connected with all the nodes in the previous and next layer. This network is too 'shallow' and generally not regarded as 'deep' neural networks.

Multilayer Perceptron is the simplest and the most basic deep learning model. The key difference between MLP and the fully-connected neural network is that MLP has more than one hidden layers. All the nodes are fully-connected with the nodes of the adjacent layers but without connection with the other nodes of the same layer. MLP includes multiple hidden layers. As shown in Figure 7b, we take a structure with two hidden layers as an example to describe the data flow in MLP.

The input layer receives the observation $\boldsymbol{x}$ and feeds forward to the first hidden layer,

$$\boldsymbol{x}^{h1} = \sigma(\mathcal{T}(\boldsymbol{x})) \quad (3)$$

where $\boldsymbol{x}^{h1}$ denotes the data flow in the first hidden layer and σ represents the non-linear activation function. There are several commonly used activation functions such as sigmoid/Logistic, Tanh, ReLU, we choose sigmoid activation function as an example in this section. Then, the data flow to the second hidden layer and the output layer,

$$\boldsymbol{x}^{h2} = \sigma(\mathcal{T}(\boldsymbol{x}^{h1})) \quad (4)$$

$$\boldsymbol{y}' = \sigma(\mathcal{T}(\boldsymbol{x}^{h2})) \quad (5)$$

where $\boldsymbol{y}'$ denotes the predict results in one-hot format. The error (i.e., loss) could be calculated based on the distance between $\boldsymbol{y}'$ and the ground truth $\boldsymbol{y}$. For instance, the Euclidean-distance based error can be calculated by

$$error = \|\boldsymbol{y}' - \boldsymbol{y}\|_2 \quad (6)$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Afterward, the error will be back-propagated and optimized by a suitable optimizer. The optimizer will adjust all the weights and basis in the model until the error converges. The most widely used loss functions includes cross-entropy, negative log likelihood, mean square estimation, etc. The most widely used optimizers include Adaptive moment estimation (Adam), Stochastic Gradient Descent (SGD), Adagrad (Adaptive subgradient method), etc.

Several terms may be easily confused with each other: Artificial Neural Network (ANN), Deep Neural Network (DNN), and MLP. These terms have no strict difference and often mixed in literature and commonly used as synonyms. Generally, ANN and DNN can be used to describe deep learning models overall, including not only fully-connected networks but also other networks (e.g., recurrent, convolutional networks), but MLP can only refer to fully-connected network. Additionally, ANN contains all the models of neural networks, can be either shallow (one hidden layer) or deep (multiple hidden layers) while DNN doesn't cover shallow neural network [30, 31].
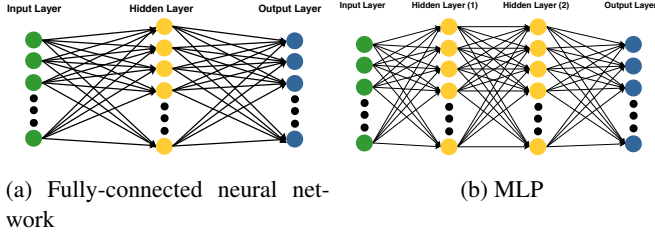
(a) Fully-connected neural network

(b) MLP

Figure 7: Illustration of standard neural network and multilayer perceptron. (a) The basic structure of the fully-connected neural network. The input layer receives the raw data or extracted features of brain signals while the output layer shows the classification results. The term 'fully-connected' denotes each node in a specific layer is connected with all the nodes in the previous and next layer. (b) MLP could have multiple hidden layers, the more, the deeper. This is an example of MLP with two hidden layers, which is the simplest MLP model.

B.1.2. Recurrent Neural Networks (RNN) Recurrent Neural Network is a specific subclass of discriminative deep learning model which are designed to capture temporal dependencies among input data [41]. Figure 8a describes the activity of a specific RNN node in the time domain. At each time ranges from $[1, t+1]$, the node receives an input I (the subscript represents the specific time) and a hidden state c from the previous time (except the first time). For instance, at time t it receives not only the input I_t but also the hidden state of the previous node c_{t-1}. The hidden state can be regarded as the 'memory' of the nodes which can help the RNN 'remember' the historical input.

Next, we will report two typical RNN architectures which have attracted much attention and achieved great success: long short-term memory and gated recurrent units. They both follow the basic principles of RNN, and we will pay our attention to the complicated internal structures in each node. Since the structure is much more complicated than general neural nodes, we call it a 'cell.' Cells in RNN are equivalent to nodes in MLP.

Long Short-Term Memory (LSTM). Figure 9a shows the structure of a single LSTM cell at time t [255]. The LSTM cell has three inputs (I_t, O_{t-1}, and c_{t-1}) and two outputs (c_t and O_t). The operation is as follows:

$$I_t, O_{t-1}, c_{t-1} \rightarrow c_t, O_t \tag{7}$$

I_t denotes the input value at time t, O_{t-1} denotes the output at the previous time (i.e., time $t-1$), and c_{t-1} denotes the hidden state at the previous time. c_t and O_t separately denote the hidden state and the output at time t. Therefore, we can observe that the output O_t at time t not only related to the input I_t but also related to the information at the previous time. In this way, LSTM is empowered to remember the important information in the time domain. Moreover, the essential idea of LSTM is to control the memory of specific information. For this aim,

LSTM cell adopts four gates: the input gate, forget gate, output gate, and input modulation gate. Each gate is a weight to control how much information can flow through this gate. For example, if the weight of the forget gate is zero, the LSTM cell would remember all the information passed from the previous time $t-1$; if the weight is one, the LSTM cell would remember nothing. The corresponding activation function determines the weight. The detailed data flow as follows:

$$f = \sigma(\mathcal{T}(I_t, O_{t-1})) \tag{8}$$

$$i = \sigma(\mathcal{T}(I_t, O_{t-1})) \tag{9}$$

$$o = \sigma(\mathcal{T}(I_t, O_{t-1})) \tag{10}$$

$$m = tanh(\mathcal{T}(I_t, O_{t-1})) \tag{11}$$

$$c_t = f * c_{t-1} + i * m \tag{12}$$

$$h_t = o * tanh(c_t) \tag{13}$$

where i, f, o and m represent the input gate, forget gate, output gate and input modulation gate, respectively.

Gated Recurrent Units (GRU). Another widely used RNN architecture is GRU [256]. Similar to LSTM, GRU attempts to exploit the information from the past. GRU does not require hidden states, however, it receives temporal information only from the output of time $t-1$. Thus, as shown in Figure 9b, GRU has two inputs (I_t and O_{t-1}) and one output (O_t). The mapping can be described as:

$$I_t, O_{t-1} \rightarrow O_t \tag{14}$$

GRU contains two gates: reset gate r and update gate z. The former decides how to combine the input with previous memory. The latter decides how much of previous memory to keep around, which is similar to the forget gate of LSTM. The data flow as follows:

$$z = \sigma(\mathcal{T}(I_t, O_{t-1})) \tag{15}$$

$$r = \sigma(\mathcal{T}(I_t, O_{t-1})) \tag{16}$$

$$\bar{O}_t = tanh(\mathcal{T}(I_t, r * O_{t-1})) \tag{17}$$

$$O_t = (1 - z) * O_{t-1} + z * \bar{O}_t \tag{18}$$

It can be observed that there's an intermediate variable $\bar{O}_t$ which is similar to the hidden state of LSTM. However, $\bar{O}_t$ only works on this time point and unable to pass to the next time point.

We here give a brief comparison between LSTM and GRU since they are very similar. First, LSTM and GRU have comparable performance as studied by literature. For any specific task, it is recommended to try both of them to determine which provides better performance. Second, GRU is lightweight since it only has two gates and without the hidden state. Therefore, GRU is faster to train and requires few data for generalization. Third, in contrast, LSTM generally works better if the training dataset is big
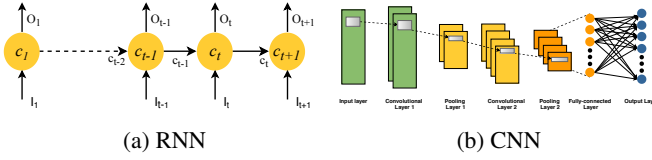
(a) RNN (b) CNN

Figure 8: Illustration of RNN and CNN models. (a) The recurrent procedure of the RNN model. This procedure describes the recurrent procedure of a specific node in time range $[1, t + 1]$. The node at time t receives two inputs variables (I_t denotes the input at time t and c_{t-1} denotes the hidden state at time $t - 1$) and exports two variables (the output O_t and the hidden state c_t at time t). (b) The paradigm of CNN model which includes two convolutional layers, two pooling layers, and one fully-connected layer.
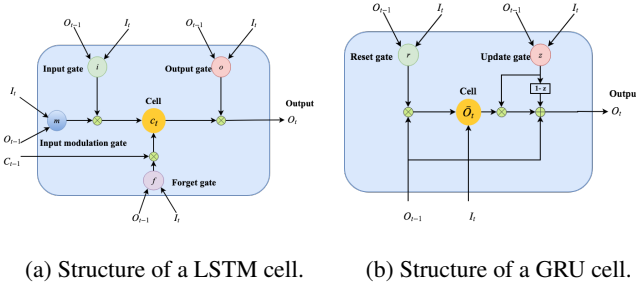

(a) Structure of a LSTM cell. (b) Structure of a GRU cell.

Figure 9: Illustration of detailed LSTM and GRU cell structures. (a) LSTM cell receives three inputs (I_t denotes the input at time t, O_{t-1} denotes the output of previous time, and c_{t-1} denotes the hidden state of the previous time) and exports two outputs (the output of this time O_t and the hidden state of this time c_t). LSTM cell contains four gates in order to control the data flow, which are the input gate, output gate, forget gate, and input modulation gate. (b) GRU cell receives two inputs (the input of this time I_t and the output of the previous time O_{t-1}) and exports its output O_t. GRU cell only contains two gates which are the reset gate and the update gate. Unlike the hidden state c_t in LSTM cell, there is no transmittable hidden state in GRU cell except one intermediate variable $\bar{O}_t$.

enough. The reason is that LSTM has better non-linearity than GRU since LSTM has two more control gates (input modulation gate and forget gate). As a result, LSTM, compared with GRU, is more powerful to discover the latent distinct information from large-level training dataset.

B.1.3. Convolutional Neural Networks (CNN)

Convolutional Neural Networks is one of the most popular deep learning models specialized in spatial information exploration [42]. This section will briefly introduce the working mechanism of CNN. CNN is widely used to discover the latent spatial information in applications such as image recognition, ubiquitous, and object searching due to their salient features such as regularized structure, good spatial locality, and translation invariance. In the area of brain signal, specifically, CNN is supposed to capture the distinctive dependencies among the patterns associated with different brain signals.

We present a standard CNN architecture as shown in Figure 8b. The CNN contains one input layer, two convolutional layers with each followed by a pooling layer, one fully-connected layer, and one output layer. The square patch in each layer shows the processing progress of a specific batch of input values. The key to the CNN is to reduce the input data into a form which is easier to recognize, with as little information loss as possible. CNN has three stacked layers: the convolutional Layer, pooling Layer, and fully-connected Layer.

The convolutional layer is the core block of CNN, which contains a set of filters to convolve the input data followed by a nonlinear transformation to extract the geographical features. In the deep learning implementation, there are several key hyper-parameters should be set in the convolutional layer, like the number of filters, the size of each filter, etc. The pooling layer generally follows the convolutional layer. The pooling layer aims to reduce the spatial size of the features progressively. In this way, it can help to decrease the number of parameters (e.g., weights and basis) and the computing burden. There are three kinds of pooling operation: max, min, average. Take max pooling for example. The pooling operation outputs the maximum value of the pooling area as a result. The hyper-parameters in the pooling layer includes the pooling operation, the size of the pooling area, the strides, etc. In the fully-connected layer, as in the basic neural network, the nodes have full connections to all activations in the previous layer.

The CNN is the most popular deep learning model in brain signal research, which can be used to exploit the latent spatial dependencies among the input brain signals like fMRI image, spontaneous EEG, and so on. More details will be reported in Section 4.

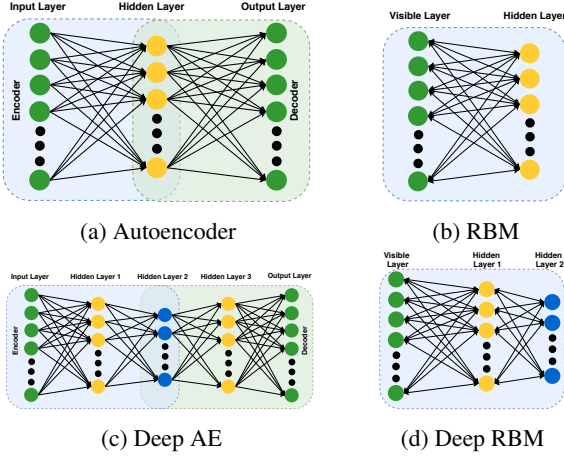(a) Autoencoder (b) RBM

(c) Deep AE (d) Deep RBM

Figure 10: Illustration of several standard representative deep learning models. (a) A basic autoencoder contains one hidden layer. The process from the input layer to the hidden layer is an encoder while the process from the hidden layer to the output layer is a decoder. (b) Restricted Boltzmann Machine, the encoder and the decoder share the same transformation weights. The input layer and the output layer are merged into the visible layer. (c) Deep AE with hidden layers. Generally, the number of hidden layers is odd, and the middle layer is the learned representative features. (d) Deep RBM has one visible layer and multiple hidden layers, the last layer is the encoded representation.

B.2. Representative Deep Learning Models

The term of representative deep learning refers to use deep neural network for representation learning. It aims to learn representations of input data that makes it easier to perform a downstream task (e.g., classification, generation, and clustering) [257].

The essential blocks of representative deep learning models are autoencoders, and restricted Boltzmann machines[25]. Deep Belief Networks are composed of AE or RBM. The representative models including AE, RBM[26], and DBN, are unsupervised learning methods. Thus, they can learn the representative features from only the input observations x without the ground truth y. In short, representative models receive the input data and output a dense representation of the data. There are various definitions in different studies for several models (such as DBN, Deep RBM, and Deep AE), in this survey, we choose the most understandable definitions and will present them in detail in this section.

B.2.1. Autoencoder (AE) As shown in Figure 10a, A autoencoder is a neural network that has three layers: the

[25]AE and RBM are generally regarded as kind of deep learning although they only have three and two layers, respectively.
[26]We regard AE, and RBMas representative methods as most researches in brain researches adopt them for feature representation.

input layer, the hidden layer, and the output layer [43]. It

differs from the standard neural network, in that the AE is trained to reconstruct its inputs, which forces the hidden layer to try to learn good representations of the inputs.

The structure of AE contains two blocks. The first block is called the encoder, which embeds the observation to a latent representation (also called 'code'),

$$x^h = \sigma(\mathcal{T}(x)) \tag{19}$$

where x^h represents the hidden layer. The second block is called the decoder, which decodes the representation into the original space,

$$y' = \sigma(\mathcal{T}(x^h)) \tag{20}$$

where y' represents the output.

AE forces y' to be equal to the input x and calculates the error based on the distance between them. Thus, AE can compute the loss function only by x without the ground truth y

$$error = \|y' - x\|_2 \tag{21}$$

Compared to Equation 6, this equation does not involve the variable y because it takes the input x as the ground truth. This is why AE is able to perform unsupervised learning.

Naturally, one variant of AE is Deep-AE (D-AE) which has more than one hidden layer. We present the structure of D-AE with three hidden layers in Figure 10c. From the figure, we can observe that there is one more hidden layer in both the encoder and the decoder. The symmetrical structure ensures the smoothness of encoding and decoding procedure. Thus, D-AE generally has an odd number of hidden layers (e.g., $2n + 1$) where the first n layers belong to the encoder, the $(n + 1)$-th layer works as the code which belongs to both encoder and decoder, and the last n layers belong to the decoder. The data flow of D-AE (Figure 10c) can be represented as

$$x^{h1} = \sigma(\mathcal{T}(x)) \tag{22}$$

$$x^{h2} = \sigma(\mathcal{T}(x^{h2})) \tag{23}$$

where x^{h2} denotes the median hidden layer (the code). Then decode the hidden layer, we can get

$$x^{h3} = \sigma(\mathcal{T}(x^{h2})) \tag{24}$$

$$y' = \sigma(\mathcal{T}(x^{h3})) \tag{25}$$

It is almost the same as AE except that D-AE has more hidden layers. Apart from D-AE, AE has many other variants like denoising autoencoder, sparse autoencoder, contractive AE, etc. Here we only introduce the D-AE because it is easily confused with the AE-based deep belief network. The key difference between them will be provided in Section B.2.3.

The core idea of AE and its variants is simple, which is that condensing the input data x into a code x^h (generally
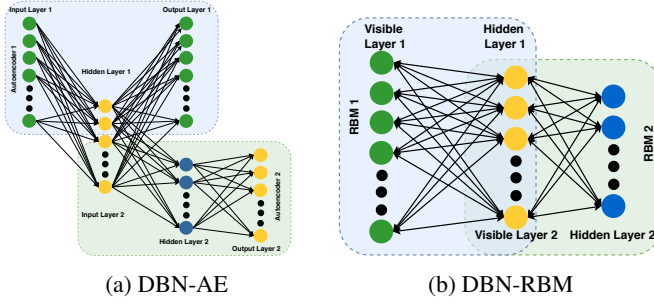
(a) DBN-AE   (b) DBN-RBM

Figure 11: Illustration of deep belief networks. (a) DBN composed of autoencoders. DBN-AE contains multiple AE components (in this case, two AE), with the hidden layer of the previous AE working as the input layer of the next AE. The hidden layer of the last AE is the learned representation. (b) DBN composed of RBM. In this illustration, there are two RBM components with the hidden layer of the first RBM working as the visible layer of the second RBM. The last hidden layer is the encoded representation. While DBN-RBM and D-RBM (Figure 10d) have similar architecture, the former is trained greedily while the latter is trained jointly .

the code layer has lower dimension) and then reconstructing the data based on the code. If the reconstructed y' can approximate to the input data x, it can be demonstrated that the condensed code x^h carries enough information about x, thus, we can regard x^h as a representation of the input data for future operation (e.g., classification).

B.2.2. Restricted Boltzmann Machine (RBM) Restricted Boltzmann Machine is a stochastic artificial neural network that can learn a probability distribution over its set of inputs [44]. It contains two layers including one visible layer (input layer) and one hidden layer, as shown in Figure 10b. From the figure, we can see that the connection lines between the two layers are bidirectional. RBM is a variant of Boltzmann Machine with stronger restriction of being without intra-layer connections. In a general Boltzmann machine, the nodes in the same hidden layer will connect. Similar to AE, the procedure of RBM also includes two steps. The first step condenses the input data from the original space to the hidden layer in a latent space. After that, the hidden layer is used to reconstruct the input data in an identical way. Compared to AE, RBM has a stronger constraint which is that the encoder weights and the decoder weights should be equal. We have

$$x^h = \sigma(\mathcal{T}(x)) \qquad (26)$$
$$x' = \sigma(\mathcal{T}(x^h)) \qquad (27)$$

In the above two equations, the weights of $\mathcal{T}(\cdot)$ are the same. Then, the error for training can be calculated by

$$error = \|x' - x\|_2 \qquad (28)$$

We can observe from the Figure 10d that the Deep-RBM (D-RBM) is an RBM with multiple hidden layers. The input data from the visible layer firstly flow to the first hidden layer and then the second hidden layer. Then, the code will flow backward into the visible layer for reconstruction.

B.2.3. Deep Belief Networks (DBN) A Deep Belief Network (DBN) is a stack of simple networks, such as AEs or RBMs [258]. Thus, we divided DBN into DBN-AE (also called stacked AE) which is composed of AE and DBN-RBM (also called stacked RBM) which is composed of RBM.

As shown in Figure 11a, the DBN-AE contains two AE structures while the hidden layer of the first AE works as the input layer of the second AE. This diagram has two stages. In the first stage, the input data feed into the first AE follows the rules introduced in Section B.2.1. The reconstruction error is calculated and back propagated to adjust the corresponding weights and basis. This iteration continues until the AE converges. We get the mapping,

$$x^1 \to x^{h1} \qquad (29)$$

Then, we move on to the second stage where the learned representative code in the hidden layer x^{h1} will be used as the input layer of the second AE, which is

$$x^2 = x^{h1} \qquad (30)$$

and then, after the second AE converges, we have

$$x^2 \to x^{h2} \qquad (31)$$

where x^{h2} denotes the hidden layer of the second AE, meanwhile, it is the final outcome of the DBN-AE.

The core idea of AE is that of learning a representative code with lower dimensionality but containing most information of the input data. The idea behind DBN-AE is to learn a more representative and purer code.

Similarly, the DBN-RBM is composed of several single RBM structures. Figure 11b shows a DBN with two RBMs where the hidden layer of the first RBM is used as the visible layer of the second RBM.

Compare the DBN-RBM (Figure 11b) and D-RBM (Figure 10d). They almost have the same architecture. Moreover, DBN-AE (Figure 11a) and D-AE (Figure 10c) have similar architecture. The most important difference between the DBN and the deep AE/RBM is that the former is trained greedily while the latter is trained jointly. In particular, for the DBN, the first AE/RBM is trained first, after it converges, the second AE/RBM is trained[44]. For the deep AE/RBM, jointly training means that the whole structure is trained together, no matter how layers it has.
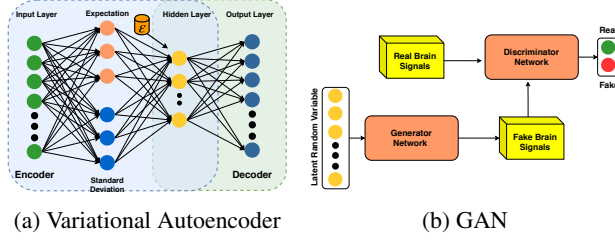
(a) Variational Autoencoder   (b) GAN

Figure 12: Illustration of generative deep learning models. (a) VAE contains two hidden layers. The first hidden layer is composed of two components: the expectation and the standard deviation, which are learned separately from the input layer. The second hidden layer represents the encoded information. ϵ denotes the standard normal distribution. (b) GAN mainly contain two crucial components: the generator and the discriminator network. The former receives a latent random variable to generate a fake brain signal while the latter receives both the real and the generated brain signals and attempts to determine if its generated or not. In the are of brain signals, GAN reconstructs or augments data instead of classification.

B.3. Generative Deep Learning Models

Generative deep learning models are mainly used to generate training samples or data augmentation. In other words, generative deep learning models play a supporting role in the brain signal field to enhance the training data quality and quantity. After the data augmentation, the discriminative models will be employed for the classification. This procedure is created to improve the robustness and effectiveness of the trained deep learning networks, especially when the training data is limited. In short, the generative models receive the input data and output a batch of similar data. In this section, we will introduce two typical generative deep learning models: variational Autoencoder (VAE) and Generative Adversarial Networks (GAN).

B.3.1. Variational Autoencoder (VAE) Variational Autoencoder, proposed in 2013 [46], is an important variant of AE, and one of the most powerful generative algorithms. The standard AE and its other variants can be used for representation but fail in generation for the reason that the learned code (or representation) may not be continuous. Therefore, we cannot generate a random sample which is similar to the input sample. In other words, the standard AE does not allow interpolation. Thus, we can replicate the

input sample but cannot generate a similar one. VAE has one fundamentally unique property that separates it from other AEs, and it is this property that makes VAE so useful for generative modeling: the latent spaces are designed to be continuous which allows easy random sampling and interpolation. Next, we will introduce how VAE works.

Similar to the standard AE, VAE can be divided into an encoder and decoder where the former embeds the input data to a latent space and the latter transfers the data from the latent space to the original space. However, the learned representation in the latent space is forced to approximate a prior distribution $p(\bar{z})$ which is generally set as Standard Gaussian distribution. Based on the reparameterization trick [46], the first hidden layer of VAE is designed to have two parts where one denotes the expectation $\boldsymbol{\mu}$ and another denotes the standard deviation $\boldsymbol{\sigma}$, thus we have

$$\boldsymbol{\mu} = \sigma(\mathcal{T}(\boldsymbol{x})) \tag{32}$$

$$\boldsymbol{\sigma} = \sigma(\mathcal{T}(\boldsymbol{x})) \tag{33}$$

Then, the latent code in the hidden layer is not directly calculated but sampled from a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$. The statistic code

$$\boldsymbol{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} * \boldsymbol{\varepsilon} \tag{34}$$

where $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. The representation $\boldsymbol{z}$ is forced to a prior distribution, and the distance $error_{KL}$ is measured by Kullback–Leibler divergence,

$$error_{KL} = D_{KL}(z, \boldsymbol{p}(\bar{\boldsymbol{z}})) \tag{35}$$

where $\boldsymbol{p}(\bar{\boldsymbol{z}})$ denotes the prior distribution. In the decoder, $\boldsymbol{z}$ is decoded into the output $\boldsymbol{y}'$,

$$\boldsymbol{y}' = \sigma(\mathcal{T}(\boldsymbol{z})) \tag{36}$$

and the reconstruction error is

$$error_{recon} = \|\boldsymbol{y}' - \boldsymbol{x}\|_2 \tag{37}$$

The overall error for VAE is combined by the DL divergence and the reconstruction error,

$$error = error_{KL} + error_{recon} \tag{38}$$

The key point of VAE is that all the latent representations $\boldsymbol{z}$ are forced to obey the normal distribution. Thus, we can randomly sample a representation $\boldsymbol{z}' \in \boldsymbol{p}(\bar{\boldsymbol{z}})$ from the prior distribution and then reconstruct a sample based on $\boldsymbol{z}'$. This is why VAE is so powerful in generation.

B.3.2. Generative Adversarial Networks (GAN) Generative Adversarial Networks [47] is proposed in 2014 and achieved great success in a wide range of research areas (e.g., computer vision and natural language processing). GAN is composed of two simultaneously trained neural networks with a generator and a discriminator. The generator captures the distribution of the input data, and the discriminator is used to estimate the probability that a sample came from the training data. The generator aims to generate fake samples while the discriminator aims to distinguish whether the sample is genuine. The functions of the generator and the discriminator are opposite; that's why GAN is called 'adversarial.' After the convergence of both the generator and the discriminator, the discriminator ought to be unable to recognize the generated samples. Thus, the pre-trained generator can be used to create a batch of samples and use them for further operations such as as classification.

Figure 12b shows the procedure of a standard GAN. The generator receives a noise signal s which is randomly sampled from a multimodal Gaussian distribution and outputs the fake brain signals x_F. The distributor receives the real brain signals x_R and the generated fake sample x_F, and then it predicts whether the received sample is real or fake. The internal architecture of the generator and discriminator are designed depending on the data types and scenarios. For instance, we can build the GAN by convolutional layers on fMRI images since CNN has an excellent ability to extract spatial features. The discriminator and the generator are trained jointly. After the convergence, numerous brain signals x_G can be created by the generator. Thus, the training set is enlarged from x_R to $\{x_R, x_G\}$ to train a more effective and robust classifier.

B.4. Hybrid Model

Hybrid deep learning models refers to models which are composed of at least two deep basic learning models where the basic model is a discriminative, representative, or generative deep learning model. Hybrid models comprise two subcategories based on their targets: classification-aimed (CA) hybrid models and the non-classification-aimed (NCA) hybrid models.

Most of the deep learning related studies in brain signal area are focused on the first category. Based on the existing literature, the representative and generative models are employed to enhance the discriminative models. The representative models can provide more informative and low dimensional features for the discrimination while the generative models can help to augment the training data quality and quantity which supply more information for the classification. The CA hybrid models can be further subdivided into: 1) several discriminative models combined to extract more distinctive and robust features (e.g., CNN+RNN); 2) representative model followed by a discriminative model (e.g., DBN+MLP); 3) generative + representative model followed by a discriminative model; 4) generative + representative model followed by a non-deep learning classifier. In which, a representative model followed by a non-deep learning classifier is regarded as a representative deep learning model.

A few NCA hybrid models aim for brain signal reconstruction. For example, St-yves et al. [259] adopted GAN to reconstruct visual stimuli based on fMRI images.