

# Deep Learning in EEG: Advance of the Last Ten-Year Critical Period

Shu Gong<sup>1</sup>, Kaibo Xing, Andrzej Cichocki<sup>2</sup>, *Fellow, IEEE*, and Junhua Li<sup>3</sup>, *Senior Member, IEEE*

**Abstract**—Deep learning has achieved excellent performance in a wide range of domains, especially in speech recognition and computer vision. Relatively less work has been done for electroencephalogram (EEG), but there is still significant progress attained in the last decade. Due to the lack of a comprehensive and topic widely covered survey for deep learning in EEG, we attempt to summarize recent progress to provide an overview, as well as perspectives for future developments. We first briefly mention the artifacts removal for EEG signal and then introduce deep learning models that have been utilized in EEG processing and classification. Subsequently, the applications of deep learning in EEG are reviewed by categorizing them into groups, such as brain-computer interface, disease detection, and emotion recognition. They are followed by the discussion, in which the pros and cons of deep learning are presented and future directions and challenges for deep learning in EEG are proposed. We hope that this article could serve as a summary of past work for deep learning in EEG and the beginning of further developments and achievements of EEG studies based on deep learning.

**Index Terms**—Brain-computer interface (BCI), classification, deep learning, disease, electroencephalogram (EEG), emotion, mental state, sleep.

## I. INTRODUCTION

MACHINE learning technology has benefited to diverse domains in our modern society [1], [2]. Deep learning, a subcategory of machine learning technology, has been showing excellent performance in pattern recognition [3], dramatically improving classification accuracy. It is worth noting that new world records were created by using deep learning in many competitions, such as ImageNet Competition [4]. The research outcomes of deep learning in speech recognition [5] and computer vision [6] have been successfully utilized to

develop practical application systems, which are remarkably influencing our life and even changing our lifestyle.

Deep learning is an enhanced variant of traditional neural network, which is thought to be established based on the inspiration of hierarchical structure existing in visual cortex of the human brain. The adjective “deep” in the term of deep learning describes the attribute of multiple processing layers forming a long-cascaded architecture. The extracted information becomes more and more abstract from the lowest layer to the highest layer. This is one of the advantages for the deep learning as information expression could be more meaningful when passing onto a higher layer. Meanwhile, deep learning suffers from the issues of slow convergence and high computation demand. These disadvantages have been released by introducing training strategies, such as dropout [7] and batch normalization [8], and the availability of high-performance computers. The high performance is not only due to the capacity improvement of central processing units (CPUs), but also new computing units, such as graphic processing unit and tensor processing unit. These new computing units are designed to suit matrix manipulation, which greatly reduce computational time in deep learning. Moreover, the availability of large scale of data and increased capacity of data storage also promote the use of deep learning.

Electroencephalogram (EEG) signal was first recorded by Hans Berger in the year of 1924 [9], which manifests underlying brain activity. Multiple electrodes can be set to record EEG signal by placing them on different locations of the scalp and temporal fluctuations in voltage can be captured in a high resolution (e.g., in milliseconds) by using a high sampling rate. With the advantages of multichannel recording and high temporal resolution, EEG has been applied to numerous domains from brain-computer interface [10]–[13], to emotion [14], [15], to cognition [16], to brain diseases [17]. EEG processing methodology is evolved from simple methods, such as mean and amplitude comparison to complicated methods, such as connectivity topology and deep learning. In particular, deep learning exhibits better performance in EEG classification (also known as, recognition or identification) compared to conventional methods [e.g., support vector machine (SVM)]. By using deep learning, discriminative features could be extracted without handcraft, which requires specific knowledge and expertise. It could avoid the low performance derived from unsuitable handcrafted features. However, deep learning is not a destination because model architecture and parameters have to be set manually. A good classification performance is usually not obtained by

Manuscript received December 18, 2020; revised March 20, 2021; accepted May 2, 2021. Date of publication May 13, 2021; date of current version June 10, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61806149; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020A1515010991; and in part by the Ministry of Education and Science of the Russian Federation under Grant 14.756.31.0001. (*Corresponding author: Junhua Li.*)

Shu Gong and Kaibo Xing are with the College of Life Sciences, Sichuan University, Chengdu 610065, China.

Andrzej Cichocki is with the Computational and Data Science and Engineering, Skolkovo Institute of Science and Technology, 121205 Moscow, Russia, and also with the Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland.

Junhua Li is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K., and also with the Laboratory for Brain-Bionic Intelligence and Computational Neuroscience, Wuyi University, Jiangmen 529020, China (e-mail: juhalee.bcmi@gmail.com).

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/TCDS.2021.3079712>.

Digital Object Identifier 10.1109/TCDS.2021.3079712

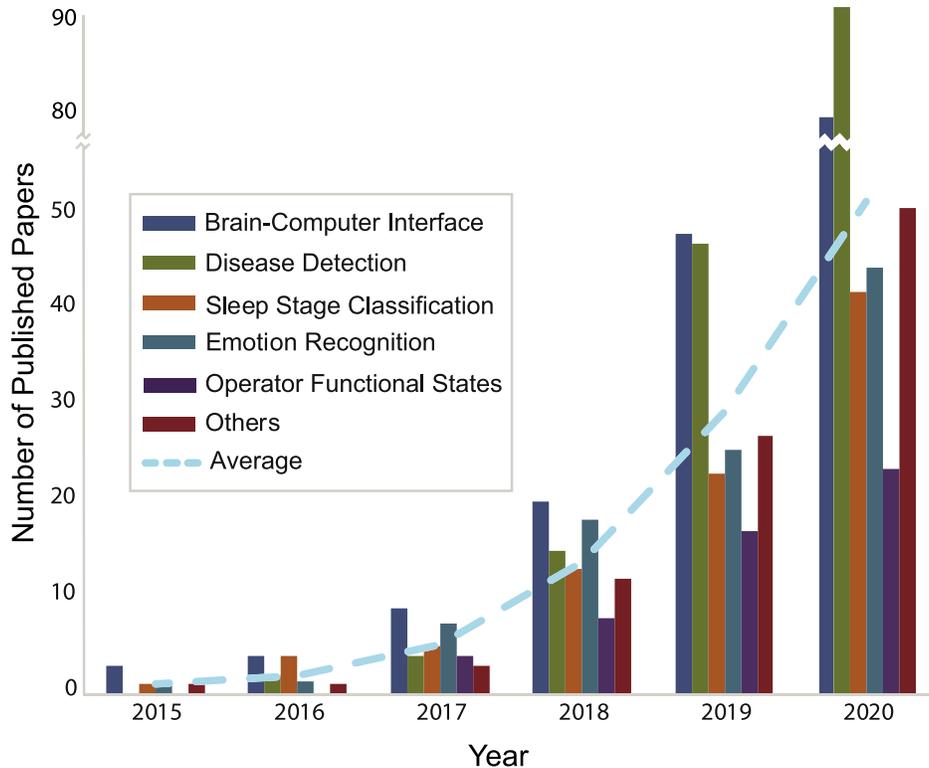


Fig. 1. Numbers of the published papers in each year. Note that numbers before 2015 are omitted because of rare papers.

just feeding data into a deep learning model. This is because the target signal is much weaker than the background signal and noise, resulting in a low signal-to-noise ratio (SNR). Therefore, artifacts removal is commonly adopted to remove artifacts so that the SNR can be improved before feeding into a deep learning model. This is quite different compared to image or video processing, where image or video is directly fed into a deep learning model. To date, different kinds of deep learning models have been employed to process and classify EEG signal. Cecotti and Graeser [18] used convolutional neural network (CNN) to extract features from steady-state visual evoked potential in 2008. Li and Cichocki [19] employed denoising autoencoder (AE) to classify two classes of motor imagery (MI) using EEG recorded from 14 electrodes on the sensorimotor cortex. Tsiouris *et al.* [20] applied recurrent neural network (RNN) to capture sequential relationships for seizure detection. A survey covering six EEG-based applications was done in 2019, where studies were reviewed separately for task type, model type and so on [21]. A more specialized survey on MI classification can be found in [22]. A distribution summary showing which disease is dominantly targeted in the studies of deep learning-based disease diagnosis can be found in [23]. If you want to read a survey on brain-computer interface (more beyond MI), it can be found in [24, Sec. 5]. If a wide range of topics of deep learning in EEG is sought, this survey can be an option.

Although EEG domain is far behind compared to the domains, such as computer vision [25] and speech recognition [26] in terms of adopting deep learning, significant progress has been achieved in the last decade. It is time to

summarize the achievements of deep learning in EEG for the past ten years and discuss current existing issues and future directions. The searching criterion [“Deep Learning” AND “EEG” AND “Classification” OR “Recognition” OR “Identification”] was used for literature retrieval in the Web of Science in March 2020. After manual selection, 193 papers were included in this survey. During the revision in February 2021, we applied the same searching criterion to find newly published literature after the previous searching and selected 20 papers to be included in this survey. After the acceptance, seven more papers were further included, but they were not used to update the figures and the supplementary tables due to the constrained time.

As shown in Fig. 1, the majority of these papers were published after 2017 while there was a rapid increase from the year of 2019. In 2019, the number of papers in the topic of brain-computer interface and disease detection are significantly more than the other topics. In 2020, the numbers of the published papers in more topics are rapidly increased, although disease detection is still a leading topic. The rapid increase of the published papers about deep learning in EEG is continued in 2021. The remainder of the survey is organized as follows. In Section II, artifacts removal is briefly introduced. This is followed by the detailed descriptions of all deep learning models which have been applied to EEG in Section III. In this section, we also mention the advantages and limitations of each deep learning model. Subsequently, the applications of deep learning in EEG are detailed along with publicly available EEG data sets used in these applications in Section IV. Finally, discussions are given and future directions are drawn at the end

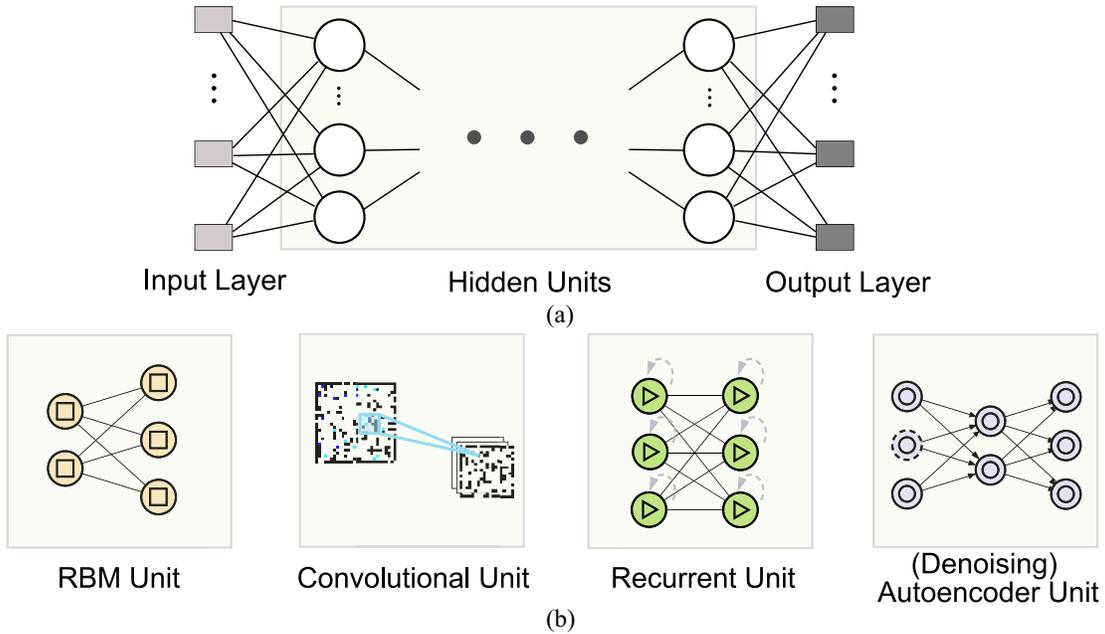


Fig. 2. (a) Generic framework of a deep learning model. (b) Classical units that are employed in a deep learning model.

of the survey. All abbreviations used in this survey are listed in the supplementary Table I.

## II. ARTIFACTS REMOVAL

In general, artifacts are larger than that we intend to extract from EEG signal in terms of scale, leading to a low SNR. In order to improve SNR, EEG signal is preprocessed to remove or mitigate the effect of artifacts on the signal before the signal is further processed. For example, a notch filter [16] is effective for eliminating the interference of power line. Independent component analysis [27] is usually utilized to remove eye movements-related and muscular activity-related artifacts. Classical methods of artifacts removal and their targeted artifacts are summarized in the supplementary Table II.

When deep learning emerges, the step of artifacts removal is kept. EEG signal is preprocessed as usual to remove artifacts before inputting into a deep learning model. This is an effective way as all artifacts removal methods can be applied with deep learning models to be of both benefits inherited from the artifacts removal methods and deep learning models. This is also a natural and straightforward way that researchers are able to easily implement. However, an independent step of artifacts removal is not always necessary. The first several layers in a deep learning model could be functioned as artifacts removal, where noise is removed through the layers. To this end, a few attempts were done. For example, Supratak *et al.* [28] inputted raw EEG data into a CNN for the classification of sleep stages. Their study showed that an acceptable performance can be achieved without an independent step of artifacts removal. In addition, Bahador *et al.* mapped the correlation of EEG channels into a 2-D space and used a CNN model to learn representations related to particular artifacts. With respect to artifact detection, this method

outperformed spectrogram-based CNNs [29]. Moreover, no auxiliary reference signal was required in their method.

## III. DEEP LEARNING MODELS

In this section, we describe each fundamental deep learning model. Their variants and combinations are not included as they share the similar rationale with fundamental models. A deep learning model is a hierarchical structure, comprising layers through which data are mapped into more and more abstract. Whatever a deep learning model is, there are an input layer, an output layer, and one or more hidden units [see Fig. 2(a)]. The hidden unit might be one of the layer structures illustrated in Fig. 2(b) or their combinations. In the following sections, we introduce classical deep learning models where typical units illustrated in Fig. 2(b) are embedded.

### A. Restricted Boltzmann Machine and Deep Belief Networks

A restricted Boltzmann machine (RBM) [30] is an undirected graph model (see Fig. 2(b): RBM Unit), which has a visible layer  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  and a hidden layer  $\mathbf{h} = (h_1, h_2, \dots, h_n)$ . Connections exist only between visible layer  $\mathbf{v}$  and hidden layer  $\mathbf{h}$  and there are no connections between nodes within the visible layer or hidden layer. The energy function for an RBM is defined as

$$E(\mathbf{v}, \mathbf{h}) = -\mathbf{v}^T \mathbf{W} \mathbf{h} - \mathbf{a}^T \mathbf{v} - \mathbf{b}^T \mathbf{h} \quad (1)$$

where  $\mathbf{W}$  is the weight matrix,  $\mathbf{a}$  and  $\mathbf{b}$  are bias vectors. The joint probability of  $\mathbf{v}$  and  $\mathbf{h}$  is constructed in terms of  $E$

$$P(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v}, \mathbf{h})} \quad (2)$$

where  $Z$  is a normalizing constant defined as

$$Z = \sum_{\mathbf{v}, \mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h})}. \quad (3)$$

The marginal distribution over the visible variables is obtained as

$$P(\mathbf{v}) = \frac{1}{Z} \sum_{\mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h})}. \quad (4)$$

The conditional probabilities can be described as

$$P(h_j = 1 | \mathbf{v}) = \sigma(\mathbf{W}_j \mathbf{v} + b_j) \quad (5)$$

$$P(v_i = 1 | \mathbf{h}) = \sigma(\mathbf{W}_i \mathbf{h} + a_i) \quad (6)$$

where  $\sigma$  is logistic function defined as

$$\sigma(x) = (1 + e^{-x})^{-1}. \quad (7)$$

A deep belief network (DBN) is constructed by stacking multiple RBMs [31]. Each RBM in the DBN is trained using an unsupervised manner at first. Then, the output of previous RBM is inputted into the next RBM. All RBMs are fine-tuned together by supervised optimization.

### B. Convolutional Neural Network

CNN [32] is good at capturing spatial information of data [see Fig. 2(b): convolutional unit]. Most CNNs consist of two types of layers: 1) convolutional layer and 2) pooling layer.

In specific, a convolutional layer has filters  $k_{ij}^l$ , the size of which is usually much smaller than the dimension of input data and forms a locally connected structure. Filter at layer  $l$  can produce feature maps  $\mathbf{X}_j^l$  by convolving with the input  $\mathbf{X}_i^{l-1}$  plus biases  $b_j^l$ . These features are subjected to a nonlinear transformation  $f(\cdot)$  and can be mathematically expressed as

$$\mathbf{X}_j^l = f\left(\sum_{i=1}^{M^{l-1}} \mathbf{X}_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (8)$$

where  $M^{l-1}$  represents the number of feature maps in layer  $l-1$ , and  $*$  denotes convolution operation.

A pooling layer is responsible for feature selection and information filtering. Two kinds of pooling operations are widely used: max pooling and average pooling. In max pooling, maximum value is mapped from a subregion by pooling operator. In average pooling, the average value of a subregion is selected as the result. A fully connected layer is usually added at the last part of a CNN in the case of classification. It transforms a long 1-D vector and outputs to the next layer (usually softmax).

Weight sharing and sparse connections are two basic strategies in CNN models, which lead to dramatic reduction in the number of parameters. These strategies are helpful to reduce training time and enhance training effectiveness. Moreover, they also mitigate the overfitting problem while retaining a good capability of complex feature extraction.

### C. Recurrent Neural Networks

RNN [33] was developed to deal with sequential data because of its unique recurrent structure [see Fig. 2(b): recurrent unit], which allows previous outputs to be used as inputs while having hidden states. It is widely used in applications that need to extract sequential information, such as natural language processing, speech recognition, and EEG classification.

1) *GRU*: Gated recurrent unit (GRU) [34] has two gates, reset  $\mathbf{r}_t$  and update  $\mathbf{z}_t$ . Let  $\mathbf{x}_t$  be the input at time step  $t$  to a GRU layer and  $\mathbf{h}_t$  be the output vector. The output activation is a linear interpolation between the activation from the previous time step and a candidate activation  $\hat{\mathbf{h}}_t$

$$\mathbf{h}_t = \mathbf{z}_t \odot \mathbf{h}_{t-1} + (1 - \mathbf{z}_t) \odot \hat{\mathbf{h}}_t \quad (9)$$

where  $\mathbf{z}_t$  decides the interpolation weight, which is computed by

$$\mathbf{z}_t = f(\mathbf{W}_z \mathbf{x}_t + \mathbf{U}_z \mathbf{h}_{t-1} + \mathbf{b}_z) \quad (10)$$

where  $\mathbf{W}$  and  $\mathbf{U}$  are weight matrices for the update gate,  $\mathbf{b}$  is a bias vector, and  $f(\cdot)$  is a nonlinear function (usually sigmoid function). The candidate activation is also controlled by an additional reset gate and computed as follows:

$$\hat{\mathbf{h}}_t = g(\mathbf{W}_h \mathbf{x}_t + \mathbf{U}_h (\mathbf{r}_t \odot \mathbf{h}_{t-1}) + \mathbf{b}_h) \quad (11)$$

where  $\odot$  represents an elementwise multiplication and  $g(\cdot)$  is often a nonlinear tanh function. The reset gate is computed in a similar manner as the update gate

$$\mathbf{r}_t = f(\mathbf{W}_r \mathbf{x}_t + \mathbf{U}_r \mathbf{h}_{t-1} + \mathbf{b}_r). \quad (12)$$

2) *LSTM*: Different from GRU, long short-term memory (LSTM) [35] has three gates, input  $\mathbf{i}_t$ , output  $\mathbf{o}_t$ , and forget gates  $\mathbf{f}_t$ . Each LSTM cell has an additional memory component  $\mathbf{c}_t$ . The gates are calculated in a similar manner as the GRU but LSTM has additional memory components

$$\mathbf{i}_t = f(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i) \quad (13)$$

$$\mathbf{o}_t = f(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o) \quad (14)$$

$$\mathbf{f}_t = f(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f). \quad (15)$$

A memory component is updated by forgetting the existing content and adding a new memory component as

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{c}}_t \quad (16)$$

where  $\hat{\mathbf{c}}_t$  can be computed by

$$\hat{\mathbf{c}}_t = g(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c). \quad (17)$$

The updated equation for the memory component is controlled by the forget and input gates. Then, the output of the LSTM unit is computed from the memory modulated by the output gate according to the following equation:

$$\mathbf{h}_t = \mathbf{o}_t \odot g(\mathbf{c}_t). \quad (18)$$

### D. Autoencoder and Stacked Autoencoder

AE is a symmetrical structure with two layers [36] [see Fig. 2(b): AE unit].

An encoder learns latent representation from the input data while a decoder restores the latent representation as close to the input data as possible. The goal of an AE is to minimize the reconstruction error between the input and the output.

Given the inputs  $\mathbf{x} \in \mathbf{R}$ , the encoding process first maps it into a latent representation  $\mathbf{h} \in \mathbf{R}$  through a weight matrix  $\mathbf{W}_v$ , bias  $\mathbf{b}_v$ , and an activation function  $f(\cdot)$

$$\mathbf{h} = f(\mathbf{W}_v \mathbf{x} + \mathbf{b}_v). \quad (19)$$

Then the decoding process transforms the latent representation  $\mathbf{h}$  into the reconstruction  $\mathbf{y}$  through a weight matrix  $\mathbf{W}_h$ , bias  $\mathbf{b}_h$ , and an activation function  $g(\cdot)$

$$\mathbf{y} = g(\mathbf{W}_h \mathbf{h} + \mathbf{b}_h). \quad (20)$$

To simplify the network architecture, the tied weights strategy  $\mathbf{W}_v = \mathbf{W}_h = \mathbf{W}$  are usually employed. The parameters to be determined are  $\{\mathbf{W}, \mathbf{b}_v, \mathbf{b}_h\}$ . The training of an AE is to minimize the loss

$$\arg \min_{\mathbf{W}, \mathbf{b}_v, \mathbf{b}_h} \mathcal{J}(\mathbf{W}, \mathbf{b}_v, \mathbf{b}_h). \quad (21)$$

Given the training samples  $\mathbf{D}_n$ , the loss function is defined as

$$\mathcal{J}(\mathbf{W}, \mathbf{b}_v, \mathbf{b}_h) = \frac{1}{N_{D_n}} \sum_{\mathbf{x} \in \mathbf{D}_n} L(\mathbf{x}, \mathbf{y}) \quad (22)$$

where  $L$  is the error of the reconstruction and  $N_{D_n}$  is the number of the training samples.

Stacked AE (SAE) is a neural network, where AEs are connected one another to form a cascade.

### E. Others

In addition to the aforementioned models, there are other models aiming to solve particular shortcomings existing in the above models. For example, capsule network (CapsNet) was proposed to overcome the shortcoming that CNN does not well capture the relationships between the parts of an image [37]. When it applied to fMRI [38] and EEG [15], it is expected to capture comprehensive relationships among brain regions, channels, or frequencies, and so on. To shorten training time, extreme learning machine (ELM) was proposed, where the weights of hidden layers are randomly assigned and fixed during the training [39]. Weight randomization is also implemented in echo state network (ESN) [40]. ESN is a RNN where the weights of hidden layers are randomly and sparsely assigned and fixed while the weights of output layer can be tuned. Spiking neural network (SNN) is a biologically inspired model and has been used to explore brain activity patterns in [41]. Deep polynomial network (DPN) uses a quadratic function to process its inputs and is able to learn features between different samples or dimensions. It was implemented in [42] to utilize features from multiple views for MI classification, including common spatial pattern, power spectral density, and wavelet packet transform. In addition, some variants of deep learning models were proposed by using different training strategies, such as generative adversarial network.

## IV. APPLICATIONS

We summarized applications, in which deep learning was utilized for EEG processing and classification, in this section. For your convenience, we group diverse applications into six topics, which are brain-computer interface (see the supplementary Table III for the details of studies), disease detection (see the supplementary Table IV for the details of studies), emotion recognition (see the supplementary Table V for the details of studies), operator functional states (OFSs) (see the supplementary Table VI for the details of studies), sleep stage

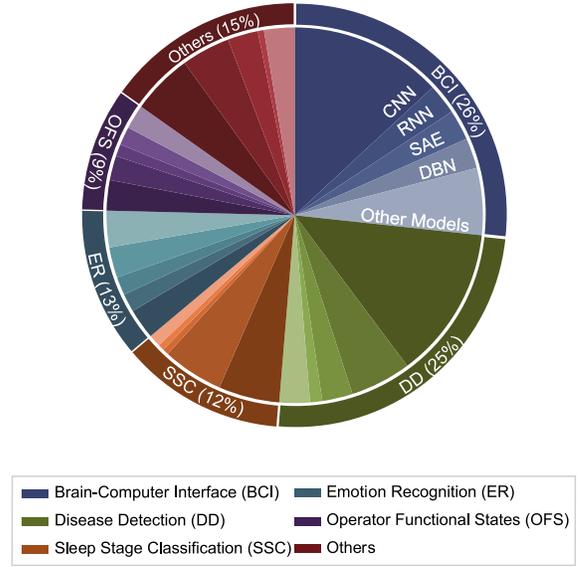


Fig. 3. Percentages of application topics and deep learning models. The outer ring represents paper percentages for each topic. The models within each topic are distinguished from the darkest to lightest colors, which stand for CNN, RNN, SAE, DBN, and other models in order.

classification (see the supplementary Table VII for the details of studies), as well as the applications other than above topics (see the supplementary Table VIII for the details of studies). According to statistics, the majority of selected papers belong to the topics of brain-computer interface (account for 26%) and disease detection (account for 25%). The percentages of each topic and the percentages of each model used in each topic are illustrated in Fig. 3. In addition, we collected the information of the publicly available data sets which had been used in the studies and listed them in the supplementary Table IX.

### A. Brain-Computer Interface

A brain-computer interface (BCI) can be defined as a system that decodes brain activity and translate user's intentions into messages or commands for the purposes of communication or the control of external devices, and more. In this topic, deep learning was mainly applied to establish MI - and P300-based BCIs (see Fig. 4).

Transfer learning is utilized to mitigate the cost of retraining or solve the problem of data lack in the target domain. A deep learning model trained on the data collected from a session or a subject can be transferred to classify/recognise the data of another session or another subject with a fine-tuning. In some cases, the fine-tuning is omitted. In general, the fine-tuning positively contributes to the performance. The extent of fine-tuning was investigated in a recent study [43]. It shows that the best performance of MI classification was achieved when all layers were tuned except the first hidden layer under the condition of a low learning rate. Another study comparing cross-session transferring and cross-subject transferring demonstrated that the cross-session transferring was feasible and the cross-subject transferring was inefficient [44]. With the combination of transfer learning and CNN, Hang *et al.* [45]

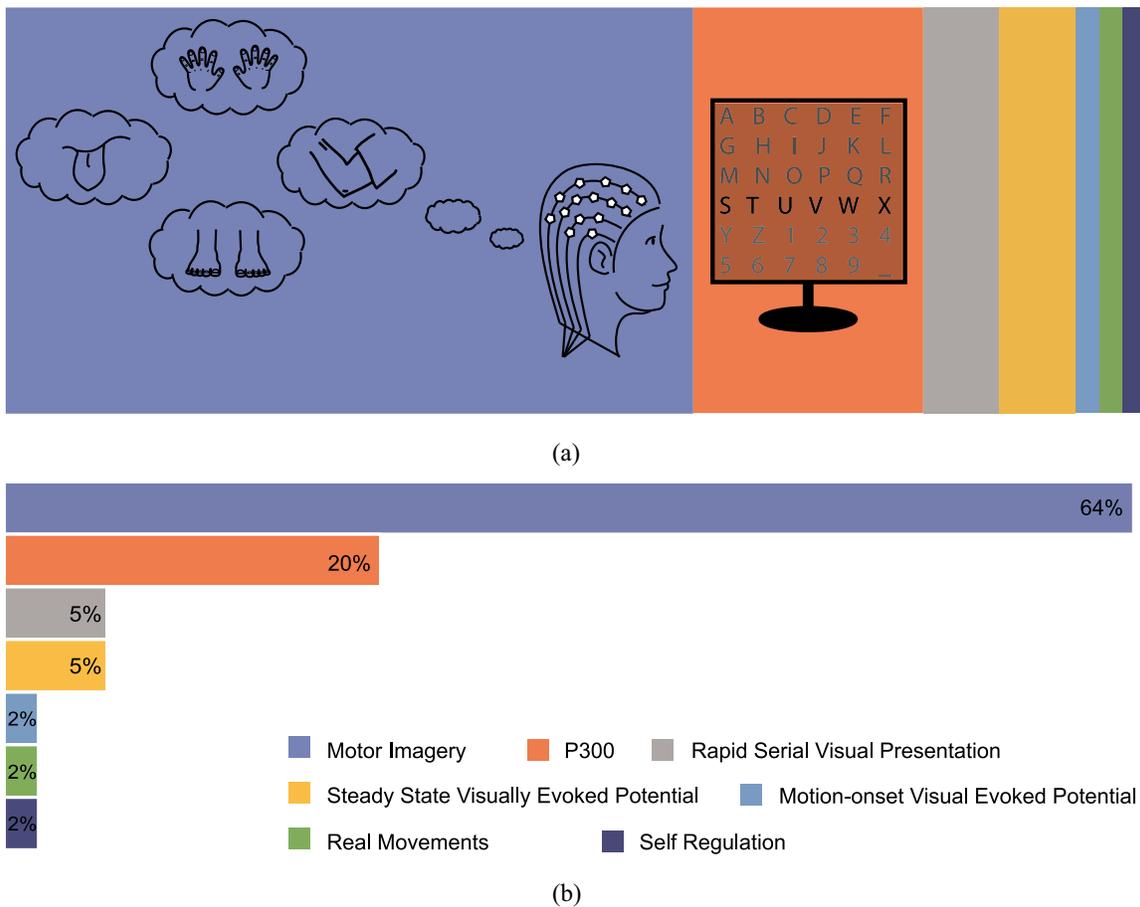


Fig. 4. (a) Paradigms of brain-computer interface. (b) Percentages of the selected papers for each paradigm by the year of 2020.

proposed a deep domain adaption network. They used maximum mean discrepancy to minimize the distribution discrepancy between target and source subjects and used the center-based discriminative feature learning method to make deep features closer to corresponding class centers. The evaluation on BCI Competition data sets (i.e., Data set IVa of Competition III and Data set IIa of Competition IV) demonstrated a good classification performance. In the study of cross-subject transferring [46], network weights were transferred. Dose *et al.* [47] used a pool of data to obtain a universal model of CNN. This model was then adapted based on a small amount of data from a subject before applying to this subject. Their results showed that an average improvement of 6%–9% was achieved for MI classification in terms of classification accuracy.

Transferring can also be conducted between domains. A CNN-based model (VGG-16) trained on image data (the data from ImageNet) was transferred to recognize EEG data by freezing the parameters in the first several layers and fine-tuning the parameters in the last several layers using an EEG data set [48]. The performance was better than that of SVM. Similar to the domain of image recognition, the amount of EEG data can also be increased by augmentation procedure. Li *et al.* [49] produced new samples by adding noise into EEG data. They claimed that adding noise into amplitudes of power spectra was superior to that adding noise into EEG time series in terms of classification accuracy. Zhang *et al.* [50]

used intrinsic mode functions derived from empirical mode decomposition to generate new EEG samples so that the total number of samples was increased.

Classical models, such as CNN and RNN were originally developed for image or speech recognition, so they did not well match the characteristics of EEG signal. They should be adapted before applying to EEG recognition. Li *et al.* [49] designed a CNN-based network consisted of three blocks to capture spatial and temporal dependencies. Multichannel raw EEG signals were fed into temporal convolutional layer and spatial convolutional layer successively in the first block. In the second block, a standard convolutional layer and a dilated convolutional layer were utilized to extract temporal information at different scales while reducing the number of parameters. The extracted features were finally used for MI classification in the third block. In another CNN-based network [51], a layer was fed by all outputs from previous layers and its output was inputted to all following layers. By using such dense interlayer connections, information loss could be reduced. In [50], EEG signals were transformed into tensors and fed into a CNN-like network where convolution were replaced with complex Morlet wavelets, resulting in parameter reduction. Wavelet kernel was also used to learn time-frequency features [46]. Their results demonstrated that wavelet kernels can provide faster convergence rate and higher classification accuracy compared to plain CNN. Alazrai *et al.* [52] used CNN to extract

features from time-frequency images, which were transformed using a quadratic time-frequency distribution. The methods were compared to a SVM, and it suggested that CNN can achieve good performance in MI tasks of the same hand.

In order to accelerate the training course and alleviate the overfitting problem, Liu *et al.* [8] adjusted the number and position of batch normalization layers in a CNN-based network for P300 detection. Kshirsagar and Londhe [53] employed leaky rectified linear unit activation function at each convolutional layer. To evaluate whether the number of convolutional layers needs to be adjusted for different BCI tasks and find out an optimal structure, Lawhern *et al.* [54] compared networks with different numbers of convolutional layers. Their results showed that deep CNN (i.e., five convolutional layers) tended to perform better on the oscillatory BCI data set than on the event-related potential BCI data set, while shallow CNN (i.e., two convolutional layers) achieved better performance on the event-related potential BCI data set. Apart from CNN, Lu *et al.* [44] used a DBN (i.e., three RBMs and an output layer) to extract features of MI. Some studies aimed to compare performances of different deep learning models. For example, Pei *et al.* [55] compared SAE and CNN in the classification of reaching movements. They found that SAE was better than CNN and suggested that poorer performance in CNN might be due to the lack of training data. One year later, another study comparing between these two models showed that SAE had satisfactory performance in some trials, but inefficient to those trials of the subjects who were less attentive in P300 detection, while CNN performed well in terms of accuracy and information transfer rate [53].

The combination of deep learning model and traditional model or the mixture of two or more types of deep learning models is applied to EEG classification. For example, SAE was combined with SVM to classify EEG signal [56]. SAE was also combined with CNN to develop a new model [57], where CNN layers were used to extract features from 2-D time-frequency images (obtained by Fourier transform over EEG signals) and SAE was further used to extract features. In [58], the features extracted by CNN were fed into an AE for cross-subject MI classification. This combination achieved a better accuracy for the cross-subject classification, but worse for the subject-specific classification, compared to the combination of CNN and multilayer perceptron (MLP). Zhang *et al.* [59] presented a hybrid network comprised of CNN and LSTM, in which EEG signals were sequentially processed through common spatial pattern, CNN, and LSTM. The idea of using CNN and LSTM to extract spatial and temporal features was also conceived by Yang *et al.* [60]. However, they inserted a discrete wavelet transformation (DWT) between CNN and LSTM, which led to better performance in the MI classification compared to that of pure combination of CNN and LSTM.

In addition to P300- and MI-based BCIs, deep learning models also applies to the other BCIs, including motion-onset visual evoked potentials [61] and self-paced reaching movements [55]. Nguyen and Chung [62] developed a steady state visually evoked potential (SSVEP)-based BCI speller system, in which only one channel was used. They used fast

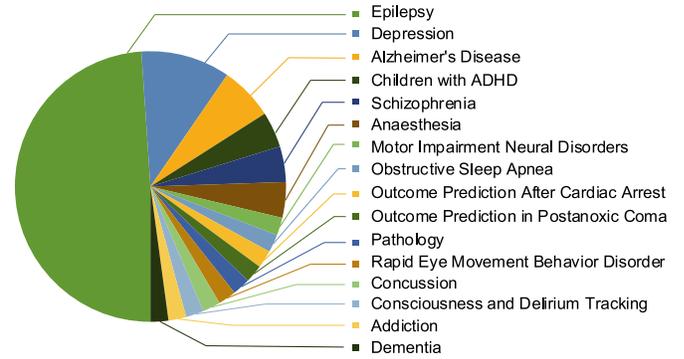


Fig. 5. Percentages of the selected papers across diseases.

Fourier transform to extract features from this channel and then fed the features into a CNN model. According to their results, frequency resolution and time window length influence classification performance. The frequency resolution of 0.0625 Hz and time window of 2s were optimal for the five-class classification [62]. Waytowich *et al.* [63] proposed a compact CNN to deal with asynchronous problem in SSVEP classification. It outperformed canonical correlation analysis (CCA) and combined-CCA.

### B. Disease Detection

Machine learning could benefit disease diagnosis by providing assistant information and preliminary diagnostic results. In this topic, deep learning models were also widely employed to detect a variety of diseases (see the distribution of the selected papers over diseases in Fig. 5). In this section, commonly used models and model designing strategies were introduced at first, including the examples of single or hybrid models, as well as the detailed architecture (e.g., layer settings). Afterwards, we described other techniques that have an influence on the performance of deep learning.

CNN is a deep learning model, which has been widely adopted for the detection of brain diseases (e.g., seizure detection [64] and schizophrenia identification [65]). Cao *et al.* stacked multiple CNNs to classify epileptic signals. In this study, the proposed model was compared to a few classification algorithms (i.e., SVM, k-nearest neighbors (kNNs), ELM) under different conditions (i.e., 1) two-class, seizure/nonseizure; 2) three-class, interictal/preictal/ictal; and 3) five-class, interictal/three preictal states/ictal) [66]. To enhance the performance of epilepsy classification, original binary labels, namely, interictal epileptiform discharge (IED) and non-IED, were converted into multiple labels used for model training [67]. Specifically, samples were further divided into five subclasses according to spatial distribution and morphology of EEG waveforms and were then fed into a CNN model for the training. A new sample was first classified to one of these subclasses and then the final classification result (IED versus non-IED) was obtained by applying a threshold at the last layer. Compared to the CNN model training with binary labels, the training with further finer tags could enhance the discriminative power of the model and led to better performance in the most subjects.

When CNN is combined with other models, classification performance can be improved. In [68], CNN and AE were combined to learn robust features in an unsupervised way. The integrated network had an encoder consisting of convolution and down-sampling and a decoder consisting of deconvolution and up-sampling. Their results demonstrated that CNN+AE is superior to principal component analysis (PCA) and sparse random projection (SRP) in epilepsy related feature extraction. In [69], a hybrid model combining CNN, AE, and LSTM achieved remarkable prediction of seizure. Combined deep learning model was used for pretraining and latent representation learning. By this, the accuracy of focal and nonfocal classification was improved [70]. However, model combination is not always positive to the performance improvement. Some studies showed that performance may decline in some cases. For instance, Mumtaz and Qayyum [71] combined CNN and LSTM to detect unipolar depression. Their results showed that the hybrid model did not outperform single model of CNN.

Beyond the selection of deep learning models, model settings also vary across studies. Tsiouris *et al.* [20] found that overfitting problem can be mitigated by shuffling input EEG segments, which could replace the dropout role partially. Qiu *et al.* [72] applied data corruption in the SAE for seizure detection. Specifically, they designed a denoising sparse AE, in which some of the input data were set to zero. This improved model robustness and reduced overfitting problem. In addition, performance is also influenced by the condition of data recording. Mumtaz and Qayyum [71] found that unipolar depression can be more accurately detected using the EEG recorded under the condition of eyes open compared to that of eyes closed. In the study of attention deficit hyperactivity disorder (ADHD) detection using a CNN model, EEG signals at different channels were rearranged to make adjacent channels together in the connectivity matrix to improve accuracy [73]. Moreover, Tsiouris *et al.* [20] shuffled interictal and preictal segments of EEG to avoid the overfitting in seizure detection. Yuan *et al.* [74] used a channel-aware module to enhance the capability of feature learning and concentrate on important and relevant EEG channels. Daoud and Bayoumi [69] computed the statistical variance and entropy of the channels, and selected those with the highest variance entropy product for seizure prediction.

The performance of deep learning for disease detection is affected by EEG data arrangement. For example, EEG data are reshaped into 2-D format before inputting into a deep learning model. In [75], EEG data were transformed into 2-D images of spectral powers. Then, these images were fed into a CNN network for distinguishing Alzheimer's disease and mild cognitive impairment from healthy controls. To differentiate patients with schizophrenia [76], Pearson correlation coefficients were calculated between channels and assembled as a correlation matrix. Correlation matrices of each subject were fed into a CNN network. Moreover, fast Fourier transform [77] and continuous wavelet transform [78] were used to transform EEG data into 2-D images for motor impairment neural disorders and epilepsy classification, respectively. Wei *et al.* [79] furthered converted 2-D images into 3-D stacked images according to the mutual correlation intensity

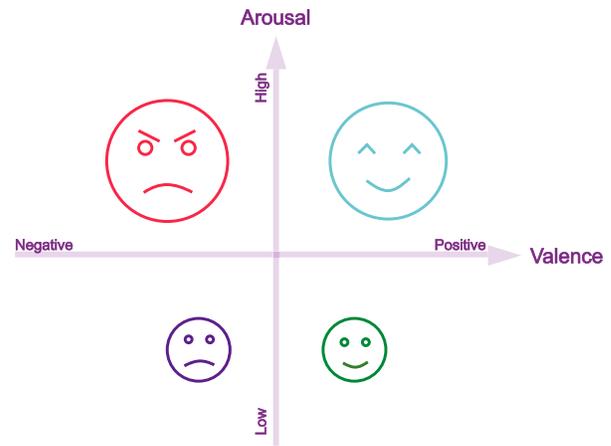


Fig. 6. Four illustrative emotions classified based on the scores of arousal and valence.

between channels. To utilize comprehensive information from different data forms, Tian *et al.* [80] used three CNNs to, respectively, obtain features existing in the time, frequency, and time-frequency domain, and then utilized these features for seizure detection. By comparing with the methods that utilizing features from only one domain, the proposed method exhibited better performance. According to the study comparing among raw EEG signal, Fourier transform, wavelet transform, and empirical mode decomposition, raw signals and empirical mode decomposition were better than the others in distinguishing focal EEG from nonfocal EEG, while Fourier transform was best in ictal and nonictal classification [81]. To handle the problem of inadequate data, sliding time window was used to split continuous EEG signal into segments with partial overlapping to increase the data amount in [82]. Cao *et al.* [17] developed an interactive system to help experts label the new data, and the data can be added to fine-tune the deep learning model to gradually improve the interictal-ictal continuum classification accuracy.

### C. Emotion Recognition

Emotion conveys lots of underlying information during conversations and is part of communication between people. People can understand emotion by reading facial expression, voice tone, and gestures. From the perspective of artificial intelligence, emotion can be recognized based on the data of facial expression [83], eye movement measures [84], EEG [85], or galvanic skin response signal [86]. According to the arousal and valence, emotion can be categorized into different classes (see Fig. 6). Based on the statistics of the included papers in this survey, the studies mainly aimed to classify three classes (i.e., positive, neutral, and negative) or more classes (partitioned based on the scores of arousal and valence). Within these papers, the data sets named “SEED” [87] and “DEAP” [88] were frequently used to evaluate deep learning models for emotion recognition.

SEED data set was published by the BCMI laboratory at the Shanghai Jiao Tong University [87]. For this data set, 62 channels were used to collect EEG data from 15 subjects when they were watching positive, negative, and neutral

video clips. The data were collected from the subjects three times with an interval of one week or longer. Thus, it enables cross-session investigations. Zheng *et al.* [89] demonstrated the stable patterns of EEG signals over time for emotion recognition. Besides, they found that differential entropy could provide better performance than other features, such as differential asymmetry and rational asymmetry. Using this data set, Yang *et al.* [90] proposed a hierarchical network which consists of subnetwork node, and this method boosted 5%–10% accuracy. Li *et al.* [91] trained a CNN and accomplished around 88% of recognition accuracy based on features of the gamma band. Zhang *et al.* proposed a two-layer RNN model to extract spatial and temporal features, respectively. The first layer of their model is an RNN layer that takes EEG signals from electrodes as inputs. The outputs of the first layer were concatenated along the time dimension and fed into the second RNN layer. The performance evaluated on the SEED data set was 89.5% [83]. Zeng *et al.* [92] used an architecture that adapted from SincNet (a CNN-based network proposed for speaker recognition [93]) to classify emotion. Their results demonstrated that the adapted SincNet (i.e., three convolutional layers and three fully connected layers) was promising for emotion classification, reaching an accuracy of around 95% as evaluated on the SEED data set.

Another data set named DEAP [88], was collected from 32 subjects when they watched 40 one-minute-long music videos. Perceptual emotion was assessed in terms of arousal, valence, liking, and dominance. Studies using this data set have shown that deep learning was successful and effective to classify emotion categories based on EEG. [85], [94]. Even using raw EEG as the input, LSTM achieved an acceptable accuracy of around 85% in the emotion classification [95]. In [96], various hand-crafted EEG features (e.g., sample entropy, mean, and power spectral density) were fed into three SAEs in a parallel way for voting. Chao *et al.* [97] also designed a parallel architecture to process EEG signal. However, they used DBN as the basic unit. To improve the classification performance and utilize strengths of different models. Li *et al.* [98] combined CNN and LSTM to extract representations from multichannel EEG, in which CNN was used to learn interchannel and interfrequency correlation while LSTM was used to extract contextual information. The model combination was also used in [99], where feature extraction was done by graph convolutional networks, temporal information was memorized by LSTM, and classification was done by a SVM. The same idea of model combination was also used in [100], where CNN was used for feature extraction.

Besides the two commonly used data sets (i.e., SEED and DEAP), Serap Aydin used affective video clips to induce nine emotional states (fear, anger, happiness, sadness, amusement, surprise, excitement, calmness, and disgust) and investigated gender effect on emotion recognition [101]. This article revealed that emotion is more affected by individual experience than gender. Zhu *et al.* designed an experiment to explore the emotion in the scenario of two-person interaction. In their experiment, two person need to rate their emotions induced by the same picture one by one. They extracted the intrabrain and interbrain phase synchronization features from

emotional EEG signals and applied a CNN model to evaluate [102]. As we know, deep learning needs parameter tuning and it is time consuming. To mitigate this problem, various strategies were proposed. Hemantha *et al.* [103] modified the back-propagation neural network by arranging layers in a circular manner that the output can access the parameters of the input and hidden layers. This modification reduced convergence time by around 20%. Jirayucharoensak *et al.* [104] used PCA for dimension reduction to lower computation cost. Gao *et al.* [105] utilized gradient priority particle swarm optimization to optimize parameters of a CNN model.

#### D. Operator Functional States

The OFSs describe the mental states of operators in specific working conditions [106]. Two of them are mental workload and mental fatigue. In specific, mental workload is a measure of cognitive resources consumed in the human working memory while mental fatigue is identified by an accumulated process of a disinclination of effort and drowsiness. To date, deep learning was used to identify mental states based on EEG signal. For example, drivers' [107]–[110] and pilots' [111] fatigue was monitored for the purposes of preventing fatigued operation.

Generalization is one of the important metrics to evaluate a model. In the classification of OFSs, large variance across subjects is challenging. Many studies employed subject-specific classifiers. For example, Tao *et al.* [112] fused multiple ELMs and Naive Bayesian model to build a subject-specific classifier. This ensemble model with fine-tuned hyperparameters was of the higher subject-specific accuracy in mental workload assessment. In the study of [113], Zhang *et al.* selected the most relevant EEG channels for each subject and used these subject-specific channels for calculating weights between the input layer and the first hidden layer in the DBN. In contrast to the subject-specific models, the cross-subject model aims to have a general model for tolerating variance of subjects. For example, Heron *et al.* used multipath convolutional layers and bi-directional LSTM layers to learn frequency and temporal features over subjects. This model achieved low variance in performance across subjects and showed better generalization compared to subject-specific models [114]. Another cross-subject model was proposed using an adaptive DBN with the weights of the first hidden layer iteratively updated to track the EEG changes in a new subject [115]. When different tasks were used to induce mental workload, the induced workload might be variable across tasks. The cross-task workload classification was made by using a CNN+RNN model [116]. Another study used transfer learning strategy to improve model generalization for the classification of mental workload [117].

Multiple kinds of features can be fused to improve assessment performance of mental workload. Gao *et al.* presented a temporal convolutional block to extract sequential information of EEG. The block orderly consists of a 1-D convolution, a rectified linear activation, and a batch normalization. Temporal convolutional blocks and dense layers for spatial feature fusion were combined to form a novel network. Their results showed that this architecture can achieve higher accuracy

for fatigue classification, when compared to these networks that replace convolutional block by 1-D convolution [109]. Zhang *et al.* [118] proposed a two-stream CNN network to learn spectral and temporal features. One stream of CNN was fed by power spectral density topographic maps and the other was fed by topographic maps of amplitude distributions. At the same year (2019), they designed another network for the same propose of learning spectral and temporal features for mental workload classification. In this network, CNN with 3-D kernels were first applied to EEG cubes, then extracted features from CNN were flatten to 1-D vectors and fed to a bidirectional LSTM for further processing and classification [116]. Both models (i.e., two-stream CNN and CNN+LSTM) showed a significant improvement in mental workload classification.

### E. Sleep Stage Classification

Sleep stage classification helps us understand the course of sleep to assess sleep quality and diagnose sleep-related disorders. The supplementary Table X briefly summarized the characteristics of each sleep stage. With the aid of EEG recording, sleep quality can be assessed objectively. In the processing of sleep quality, sleep staging is a precedent step. To date, deep learning has been applied to sleep staging. For instance, LSTM model was used for sleep stage classification based on a single channel EEG [119]. CNN+LSTM model was proposed to classify sleep stages [120], [28] and detect sleep spindles [121].

Sleep consists of a sequence of stages. Therefore, temporal information should be useful for sleep stage classification. Morlet wavelets [122] and time-frequency representations [119], [123] were applied to retain temporal information in the extraction of spectral features. These extracted features were then learned by deep learning models for sleep stage classification, showing promising performance. Using the time-frequency representation of EEG, CNN model achieved good performance [124]. In another study, the CNN was combined with LSTM to capture both temporal and spatial information for sleep stage classification [125]. The CNN was also combined with attention mechanism for sleep stage classification [126]. In contrast to the supervised learning, unsupervised learning can perform with unlabeled data, which is preferable when the data labeling is expensive or very time-consuming. Zhang and Wu [127] presented a CNN model with a greedy layerwise training strategy, in which complex-valued k-means was utilized to train filters used in the convolution with unlabeled EEG data. In [128], unsupervised sparse DBN was used to extract features. Subsequent classifiers (e.g., kNN or SVM) performed well on sleep stage classification by using these unsupervised-extracted features. Jaoude *et al.* demonstrated that a large training data can help validate classification performance. They trained a deep learning model (CNN+RNN) on sleep data from more than six thousand participants and tested on several publicly available data sets. The model achieved as good as human experts in sleep staging accuracy [129]. Usually, the numbers of samples for each sleep stage are unbalanced. To date, several methods have been proposed to release this issue, including the class-balanced random sampling [122], data augmentation [130], class-balance

training set design [28], and synthetic minority oversampling technique [131].

### F. Others

Those studies that cannot be grouped into the above topics are presented in this section. A summary table with key information of those studies is prepared (see the supplementary Table VIII). On the one hand, EEG with deep learning can be used for person identification [132], [133], age and gender prediction [134]. On the other hand, it can also be used to decode brain activity related to vision, audio [135], and pain [136]. In a study of image classification [137], LSTM was used to extract EEG features while CNN was used to extract image features. This study claimed that features extracted from EEG could help image classification so that classification performance was improved. In [138], a CNN+LSTM hybrid network was used to extracted visual representations from EEG, and a generative adversarial network was applied to reconstruct images from the learnt EEG representations. Deep learning and EEG were also applied to understand brain functions and structure. These studies aimed to understand functional brain connectivity [139], speech laterality [140], as well as memory under specific conditions. For example, Baltatzis *et al.* [141] investigated the brain's activity of different people (ever experienced school bullying or not) to different stimuli (2-D videos or Virtual Reality). Dobarjeh *et al.* [142] used EEG and SNN to decode how the brain react to various commercial brands (locally familiar or not). Arora *et al.* [143] studied the memory loss after seizure surgery.

## V. DISCUSSION

In this survey, we reviewed the researches of deep learning in EEG for the last ten years, which is a critical period for the development of deep learning used in EEG. An introduction about deep learning in EEG was first presented in the first section. Subsequently, we presented classical methods of artifacts removal which is an important step in EEG processing. We detailed prevalent deep learning models, followed by the comprehensive reviews on different applications that used deep learning to process and classify EEG signals. These applications were categorized into several topics for presentation. The increase in the number of published papers suggested that the research of deep learning in EEG are expanding over time. Although remarkable achievements were obtained, challenges and limitations still exist, which need to be addressed. We discuss them below and provide our perspectives.

The performance of deep learning-based classification should be further improved. Although the published papers showed the advantages of deep learning in EEG classification and demonstrated that deep learning is superior to conventional methods, the performance is much lower compared to the performance achieved by deep learning in image or speech classification [25], [26]. The reasons for the lower performance are mainly due to two aspects: 1) EEG signal itself and 2) deep learning models. On the one hand, EEG signal is nonstationary and much variable over time, which makes the extraction of robust features difficult. An effective solution for this problem

is to partition continuous EEG signal into short segments, which can be seen as a stationary signal. However, this is only an approximation but not a final solution. When performing cross-subject classification or cross-session classification, EEG over subjects or sessions is largely variable, making the above problem more dominant. On the other hand, most deep models are originally proposed to process other signals (e.g., images) rather than EEG. Although certain adaptations of the models have been done, the performance is still not ideal because of mismatch between the models and EEG characteristics. Taking CNN as an example, it is more suitable for image processing. Raw images can be directly fed into the CNN. However, this is not the case when applying to EEG signals. Although we have seen some studies, in which raw EEG was fed into CNN directly without preprocessing, it is not mainstream. The mainstream is still to preprocess EEG before feeding into a deep learning model because the preprocessing is very effective for removing noises to improve SNR. Another advantage of the preprocessing step is that EEG data can be transformed into other representations and/or reorganized to facilitate the following processing in the deep learning model. For instance, spectral power density is one of the most widely used feature for EEG signal. Without a separate preprocessing step, this kind of feature cannot be obtained because temporal EEG signal cannot be transformed into spectral domain within the deep learning model.

Available data size in EEG studies is significantly smaller than that available in image or speech studies [25], [26]. As we know, the deep learning model requires extensive training and a large data size can benefit model training to a great extent. Compared to the millions of training data in image or speech recognition, the scale of training data is much less in EEG classification, only from tens, hundreds, or at most thousands of participants. One potential solution for the lack of EEG data in the model training is the use of transfer learning. Deep learning model can be trained by the data which are not collected at the moment and the trained model can be used for recognition or classification on the new collected EEG data after fine-tuning or even without fine-tuning [44]–[46]. Unlike image classification, for which there are mature existing pre-trained models (e.g., ImageNet pretrained VGG model), there is no publicly available pretrained model for EEG classification. If VGG model is directly applied to EEG, reorganization of EEG has to be done in order to meet the input data format of VGG model. This reorganization might lead to information loss and give detrimental effect on the EEG classification. In addition, there is no idea how well a model trained on images can be tuned to classify EEG signal.

Based on the effectiveness comparison of transfer learning, greater performance improvement was observed in image classification compared to EEG classification. This might be due to the lack of effective training framework and strategies that are suitable for transferring EEG patterns. There was an attempt to transfer the model trained on images to EEG classification [48]. This transferring is across distinct modalities. It is likely to have a better performance when transferring across relevant modalities. As we know, there are different modalities (e.g., functional near-infrared spectroscopy (fNIRS) and

EEG) that can be used to measure underlying brain activity. A deep learning model can be trained on one modality and then fine-tuned by the other modality to classify signals of that modality. Or, different modalities can be used together to train a deep learning model so that the training can be benefited from the complementary information existing in the different modalities. It is a fusion of modalities. It has been seen that classification performance was elevated by feature fusion in the case of using conventional classifiers [144]. The fusion could be done at the different stages of the classification process (e.g., at the beginning of initial feature fusion or at the later stage of decision fusion [145], [146]). Wu *et al.* [147] utilized both EEG and electrooculogram (EOG) to classify the level of vigilance by fusing the features extracted from EEG and EOG. In the future, more extensive research should be carried out to elevate the development of fusion in deep learning models. Especially, to address how to effectively fuse multiple modalities in deep learning models for neurophysiological signal classification and analysis. Of course, collecting adequate data is a straightforward solution for the lack of EEG data. However, this results in new issues, such as cost increase and time delay. If data collection involves different institutes, extra communication effort should be paid to coordinate the data collection. Meanwhile, computation demand will be increased with the increase of data size, which requires to upgrade computational hardware or replace with the new generation hardware (e.g., CPU and graphics processing unit (GPU)). As mentioned in [148], cloud computing service is an effective way to share hardware resources so that the hardware cost in individual institutes will be reduced. Using the cloud computing service, data protection and privacy have to be considered, especially for clinical data.

When applying a deep learning model to EEG, we need to adapt the deep learning model in compliance with the characteristics of EEG. For example, how to arrange the input data or how to set kernel size should be considered. EEG signal is usually not directly used and commonly transformed before feeding into a deep learning model. There are strong relationships among temporal domain, spectral domain, and spacial domain. It is important these relationships should be kept as much as possible when arranging the input data. When EEG channels are stacked along a dimension, their spacial layout is distorted. In this case, kernels, such as square kernel, that usually used in image recognition are no longer effective for EEG classification. A column kernel (covering all channels) is a better choice, which has been supported by the study in [149]. Furthermore, Wang *et al.* [38] extended the column kernel by considering brain anatomic structure to develop multiple kernels with the sizes matching brain region sizes, achieving a better performance in schizophrenia identification compared to the usually used kernels, such as square kernel.

We believe deep learning models should be changed to be more flexible. The trained model can be adapted dynamically in real time as needed. This is not limited to dynamic parameter tuning. Ideally, model architecture can also be adjusted when needed. Also, we hope the newly developed deep learning model could perform multiple tasks at the same time in the future. Please see the detailed description in [150].

Apart from the purposes of deep learning-based EEG classification, deep learning may also be a useful tool to reveal neural mechanisms of the brain. When a deep learning model achieves a satisfactory classification performance, it captures essential differences existing between the classes. Therefore, we can look at what information the deep learning model focuses on to roughly infer the underlying associated brain activity. For example, Goh *et al.* [149] presented spatial distribution of brain activations associated with lower limb movements by probing into the model of spatio-spectral representation learning. We expect that advanced deep learning models developed in the future could reversely decompose EEG signal back into the representation in the brain to reveal underlying brain mechanisms. It is unrealistic at the current stage, but paying efforts to make progress toward to this target.

A prominent advance we need to mention is the EEGNet [54], which is proven effective for different BCI paradigms. Another promising model is SincNet, which was initially proposed for speaker recognition and also well for the classification of EEG signal [92]. New deep learning architectures, such as CapsNet [38], are also required to enhance the chance of success of EEG applications.

Finally, a mix of different deep learning units has been increasingly seen, which integrates the characteristics of these units to benefit data learning. Because there is not definite guidance to set optimal deep learning architecture (e.g., model depth and model width) currently, model complexity might be considered to determine the model architecture. The model should have enough capacity for learning information in accordance with classification tasks while its complexity should be kept as low as possible to minimize computational cost.

## VI. CONCLUSION

Our survey is a glimpse of what have been done for the deep learning in EEG over the past ten years. There are still many researches currently on-going at laboratories and hospitals, dealing with challenges we mentioned above and beyond. We hope that our survey can provide the researchers who are working in this field with a summary and facilitate their researches.

## REFERENCES

- [1] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
- [2] D. Michie, D. J. Spiegelhalter, C. Taylor, and J. Campbell, "Machine learning," *Neural and Statistical Classification*, vol. 13. New York, NY, USA: Horwood, 1994, pp. 1–298.
- [3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [4] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
- [5] D. Amodei *et al.*, "Deep Speech 2: End-to-end speech recognition in English and Mandarin," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 173–182.
- [6] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.
- [7] W. P. He, G. Wang, J. Hu, C. Li, B. L. Guo, and F. P. Li, "Simultaneous human health monitoring and time-frequency sparse representation using EEG and ECG signals," *IEEE Access*, vol. 7, pp. 85985–85994, 2019.
- [8] M. F. Liu, W. Wu, Z. H. Gu, Z. L. Yu, F. F. Qi, and Y. Q. Li, "Deep learning based on batch normalization for P300 signal detection," *Neurocomputing*, vol. 275, pp. 288–297, Jan. 2018.
- [9] L. F. Haas, "Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography," *J. Neurol. Neurosurgery Psychiatry*, vol. 74, no. 1, p. 9, 2003.
- [10] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–240, Jun. 2006.
- [11] J. Li, Y. Wang, L. Zhang, and T.-P. Jung, "Combining ERPs and EEG spectral features for decoding intended movement direction," in *Proc. Annu. IEEE Int. Conf. Eng. Med. Biol. Soc.*, 2012, pp. 1769–1772.
- [12] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Sparse Bayesian classification of EEG for brain-computer interface," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2256–2267, Nov. 2016.
- [13] J. Li, C. Li, and A. Cichocki, "Canonical polyadic decomposition with auxiliary information for brain-computer interface," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 263–271, Jan. 2017.
- [14] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "EEG-based emotion classification using deep belief networks," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, 2014, pp. 1–6.
- [15] H. Chao, L. Dong, Y. L. Liu, and B. Y. Lu, "Emotion recognition from multiband EEG signals using CapsNet," *Sensors*, vol. 19, no. 9, p. 2212, 2019.
- [16] F. Li *et al.*, "Deep models for engagement assessment with scarce label information," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 4, pp. 598–605, Aug. 2017.
- [17] L. Cao *et al.*, "Smile: A system to support machine learning on EEG data at scale," *Proc. VLDB Endowment*, vol. 12, no. 12, pp. 2230–2241, 2019.
- [18] H. Cecotti and A. Graeser, "Convolutional neural network with embedded fourier transform for EEG classification," in *Proc. 19th Int. Conf. Pattern Recognit.*, 2008, pp. 1–4.
- [19] J. Li and A. Cichocki, "Deep learning of multifractal attributes from motor imagery induced EEG," in *Proc. Int. Conf. Neural Inf. Process.*, 2014, pp. 503–510.
- [20] 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," *Comput. Biol. Med.*, vol. 99, pp. 24–37, Aug. 2018.
- [21] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, 2019, Art. no. 31001.
- [22] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review," *Biomed. Signal Process. Control*, vol. 63, Jan. 2021, Art. no. 102172.
- [23] M. J. Rivera, M. A. Teruel, A. Maté, and J. Trujillo, "Diagnosis and prognosis of mental disorders by means of EEG and deep learning: A systematic mapping study," *Artif. Intell. Rev.*, pp. 1–43, Mar. 2021.
- [24] X. Gu *et al.*, "EEG-based brain-computer interfaces (BCIs): A survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Jan. 19, 2021, doi: [10.1109/TCBB.2021.3052811](https://doi.org/10.1109/TCBB.2021.3052811).
- [25] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [26] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019.
- [27] 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 Trans. Big Data*, vol. 3, no. 4, pp. 392–404, Dec. 2017.
- [28] A. Supratak, H. Dong, C. Wu, and Y. K. Guo, "DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1998–2008, Nov. 2017.
- [29] N. Bahador, K. Erikson, J. Laurila, J. Koskenkari, T. Ala-Kokko, and J. Kortelainen, "A correlation-driven mapping for deep learning application in detecting artifacts within the EEG," *J. Neural Eng.*, vol. 17, Oct. 2020, Art. no. 056018.
- [30] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 791–798.

- [31] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [33] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," 2015. [Online]. Available: arXiv:1506.00019.
- [34] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014. [Online]. Available: arXiv:1412.3555.
- [35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [36] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," *Inst. Cogn. Sci., California Univ., San Diego, CA, USA, Rep. 8506*, 1985.
- [37] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3856–3866.
- [38] T. Wang, A. Bezerianos, A. Cichocki, and J. Li, "Multikernel capsule network for schizophrenia identification," *IEEE Trans. Cybern.*, early access, Dec. 1, 2020, doi: [10.1109/TCYB.2020.3035282](https://doi.org/10.1109/TCYB.2020.3035282).
- [39] S. F. Ding, N. Zhang, X. Z. Xu, L. L. Guo, and J. Zhang, "Deep extreme learning machine and its application in EEG classification," *Math. Problems Eng.*, vol. 2015, May 2015, Art. no. 129021.
- [40] L. Bozhkov, P. Koprinkova-Hristova, and P. Georgieva, "Reservoir computing for emotion valence discrimination from EEG signals," *Neurocomputing*, vol. 231, pp. 28–40, Mar 2017.
- [41] M. G. Dobarjeh, G. Y. Wang, N. K. Kasabov, R. Kydd, and B. Russell, "A spiking neural network methodology and system for learning and comparative analysis of EEG data from healthy versus addiction treated versus addiction not treated subjects," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1830–1841, Sep. 2016.
- [42] B. Y. Lei *et al.*, "Walking imagery evaluation in brain computer interfaces via a multi-view multi-level deep polynomial network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 497–506, Mar. 2019.
- [43] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," *Neural Netw.*, vol. 136, pp. 1–10, Apr. 2021.
- [44] N. Lu, T. F. Li, X. D. Ren, and H. Y. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, Jun. 2017.
- [45] W. L. Hang *et al.*, "Cross-subject EEG signal recognition using deep domain adaptation network," *IEEE Access*, vol. 7, pp. 128273–128282, 2019.
- [46] D. Y. Zhao, F. Z. Tang, B. L. Si, and X. S. Feng, "Learning joint space-time-frequency features for EEG decoding on small labeled data," *Neural Netw.*, vol. 114, pp. 67–77, Jun. 2019.
- [47] H. Dose, J. S. Moller, H. K. Iversen, and S. Puthusserypady, "An end-to-end deep learning approach to MI-EEG signal classification for BCIs," *Expert Syst. Appl.*, vol. 114, pp. 532–542, Dec. 2018.
- [48] G. W. Xu *et al.*, "A deep transfer convolutional neural network framework for EEG signal classification," *IEEE Access*, vol. 7, pp. 112767–112776, 2019.
- [49] Y. Li, X. R. Zhang, B. Zhang, M. Y. Lei, W. G. Cui, and Y. Z. Guo, "A channel-projection mixed-scale convolutional neural network for motor imagery EEG decoding," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1170–1180, Jun. 2019.
- [50] Z. W. Zhang *et al.*, "A novel deep learning approach with data augmentation to classify motor imagery signals," *IEEE Access*, vol. 7, pp. 15945–15954, 2019.
- [51] D. L. Li, J. H. Wang, J. C. Xu, and X. K. Fang, "Densely feature fusion based on convolutional neural networks for motor imagery EEG classification," *IEEE Access*, vol. 7, pp. 132720–132730, 2019.
- [52] R. Alazrai, M. Abuhijleh, H. Alwanni, and M. I. Daoud, "A deep learning framework for decoding motor imagery tasks of the same hand using EEG signals," *IEEE Access*, vol. 7, pp. 109612–109627, 2019.
- [53] G. B. Kshirsagar and N. D. Londhe, "Improving performance of devanagari script input-based P300 speller using deep learning," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 11, pp. 2992–3005, Nov. 2019.
- [54] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, 2018, Art. no. 056013.
- [55] D. Y. Pei, M. Burns, R. Chandramouli, and R. Vinjamuri, "Decoding asynchronous reaching in electroencephalography using stacked autoencoders," *IEEE Access*, vol. 6, pp. 52889–52898, 2018.
- [56] S. Kundu and S. Ari, "P300 based character recognition using sparse autoencoder with ensemble of SVMs," *Biocybern. Biomed. Eng.*, vol. 39, no. 4, pp. 956–966, 2019.
- [57] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017, Art. no. 16003.
- [58] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain, "Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion," *Future Gener. Comput. Syst.*, vol. 101, pp. 542–554, Dec. 2019.
- [59] R. L. Zhang, Q. Zong, L. Q. Dou, and X. Y. Zhao, "A novel hybrid deep learning scheme for four-class motor imagery classification," *J. Neural Eng.*, vol. 16, no. 6, 2019, Art. no. 66004.
- [60] J. Yang, S. W. Yao, and J. Wang, "Deep fusion feature learning network for MI-EEG classification," *IEEE Access*, vol. 6, pp. 79050–79059, 2018.
- [61] T. Ma *et al.*, "The extraction of motion-onset VEP BCI features based on deep learning and compressed sensing," *J. Neurosci. Methods*, vol. 275, pp. 80–92, Jan. 2017.
- [62] T. H. Nguyen and W. Y. Chung, "A single-channel SSVEP-based BCI speller using deep learning," *IEEE Access*, vol. 7, pp. 1752–1763, 2019.
- [63] N. Waytowich *et al.*, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *J. Neural Eng.*, vol. 15, no. 6, 2018, Art. no. 66031.
- [64] 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," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2018.
- [65] S. L. Oh, J. Vicsnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, "Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals," *Appl. Sci.*, vol. 9, no. 14, p. 2870, 2019.
- [66] J. Cao, J. Zhu, W. Hu, and A. Kummer, "Epileptic signal classification with deep EEG features by stacked CNNs," *IEEE Trans. Cogn. Devel. Syst.*, vol. 12, no. 4, pp. 709–722, Dec. 2020.
- [67] A. Antoniadis *et al.*, "Detection of interictal discharges with convolutional neural networks using discrete ordered multichannel intracranial EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2285–2294, Dec. 2017.
- [68] T. X. Wen and Z. N. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.
- [69] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 804–813, Oct. 2019.
- [70] H. Daoud and M. Bayoumi, "Deep learning approach for epileptic focus localization," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 2, pp. 209–220, Apr. 2020.
- [71] W. Mumtaz and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," *Int. J. Med. Inform.*, vol. 132, Dec. 2019, Art. no. 103983.
- [72] Y. Qiu, W. D. Zhou, N. N. Yu, and P. D. Du, "Denoising sparse autoencoder-based Ictal EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1717–1726, Sep. 2018.
- [73] H. Chen, Y. Song, and X. L. Li, "A deep learning framework for identifying children with ADHD using an EEG-based brain network," *Neurocomputing*, vol. 356, pp. 83–96, Sep. 2019.
- [74] Y. Yuan, G. X. Xun, K. B. Jia, and A. D. Zhang, "A multi-view deep learning framework for EEG seizure detection," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 1, pp. 83–94, Jan. 2019.
- [75] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, "A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings," *Neurocomputing*, vol. 323, pp. 96–107, Jan. 2019.
- [76] C. A. T. Naira and C. J. L. Del Alamo, "Classification of people who suffer schizophrenia and healthy people by EEG signals using deep learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 10, pp. 511–516, 2019.
- [77] G. Vrbancic and V. Podgorelec, "Automatic classification of motor impairment neural disorders from EEG signals using deep convolutional neural networks," *Elektronika Ir Elektrotehnika*, vol. 24, no. 4, pp. 3–7, 2018.

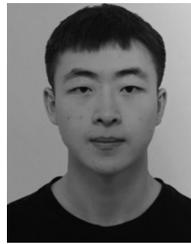
- [78] O. Turk and M. S. Ozerdem, "Epilepsy detection by using scalogram based convolutional neural network from EEG signals," *Brain Sci.*, vol. 9, no. 5, p. 115, 2019.
- [79] X. Y. Wei, L. Zhou, Z. Y. Chen, L. J. Zhang, and Y. Zhou, "Automatic seizure detection using three-dimensional CNN based on multi-channel EEG," *BMC Med. Inform. Decis. Making*, vol. 18, no. 5, p. 111, 2018.
- [80] X. B. Tian *et al.*, "Deep multi-view feature learning for EEG-based epileptic seizure detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 1962–1972, Oct. 2019.
- [81] R. San-Segundo, M. Gil-Martin, L. F. D'Haro-Enriquez, and J. M. Pardo, "Classification of epileptic EEG recordings using signal transforms and convolutional neural networks," *Comput. Biol. Med.*, vol. 109, pp. 148–158, Jun. 2019.
- [82] I. Ullah, M. Hussain, E. U. Qazi, and H. Aboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," *Expert Syst. Appl.*, vol. 107, pp. 61–71, Oct. 2018.
- [83] T. Zhang, W. M. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 839–847, Mar. 2019.
- [84] W. L. Zheng, W. Liu, Y. F. Lu, B. L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019.
- [85] A. Al-Nafjan, A. Al-Wabil, M. Hosny, and Y. Al-Ohali, "Classification of human emotions from electroencephalogram (EEG) signal using deep neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 9, pp. 419–425, 2017.
- [86] Y. H. Kwon, S. B. Shin, and S. D. Kim, "Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system," *Sensors*, vol. 18, no. 5, p. 1383, 2018.
- [87] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.
- [88] S. Koelstra *et al.*, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 18–31, Jan.–Mar. 2011.
- [89] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Trans. Affective Comput.*, vol. 10, no. 3, pp. 417–429, Jul.–Sep. 2019.
- [90] Y. Yang, Q. M. J. Wu, W.-L. Zheng, and B.-L. Lu, "EEG-based emotion recognition using hierarchical network with subnetwork nodes," *IEEE Trans. Cogn. Devel. Syst.*, vol. 10, no. 2, pp. 408–419, Jun. 2018.
- [91] J. P. Li, Z. X. Zhang, and H. G. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," *Cogn. Comput.*, vol. 10, no. 2, pp. 368–380, 2018.
- [92] H. Zeng *et al.*, "EEG emotion classification using an improved SincNet-based deep learning model," *Brain Sci.*, vol. 9, no. 11, p. 326, 2019.
- [93] M. Ravanelli and Y. Bengio, "Speaker recognition from raw waveform with SincNet," in *Proc. IEEE Spoken Lang. Technol. Workshop (SLT)*, 2018, pp. 1021–1028.
- [94] E. J. Choi and D. K. Kim, "Arousal and valence classification model based on long short-term memory and DEAP data for mental healthcare management," *Healthcare Inform. Res.*, vol. 24, no. 4, pp. 309–316, 2018.
- [95] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 355–358, 2017.
- [96] S. Bagherzadeh, K. Maghooli, J. Farhadi, and M. Z. Soroush, "Emotion recognition from physiological signals using parallel stacked autoencoders," *Neurophysiology*, vol. 50, no. 6, pp. 428–435, 2018.
- [97] H. Chao, H. L. Zhi, L. Dong, and Y. L. Liu, "Recognition of emotions using multichannel EEG data and DBN-GC-based ensemble deep learning framework," *Comput. Intell. Neurosci.*, vol. 2018, Dec. 2018, Art. no. 9750904.
- [98] X. Li, D. W. Song, P. Zhang, Y. X. Hou, and B. Hu, "Deep fusion of multi-channel neurophysiological signal for emotion recognition and monitoring," *Int. J. Data Min. Bioinform.*, vol. 18, no. 1, pp. 1–27, 2017.
- [99] Y. Yin, X. Zheng, B. Hu, Y. Zhang, and X. Cui, "EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106954.
- [100] A. Topic and M. Russo, "Emotion recognition based on EEG feature maps through deep learning network," *Int. J. Eng. Sci. Technol.*, to be published.
- [101] S. Aydin, "Deep learning classification of neuro-emotional phase domain complexity levels induced by affective video film clips," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 6, pp. 1695–1702, Jun. 2020.
- [102] L. Zhu *et al.*, "EEG-based approach for recognizing human social emotion perception," *Adv. Eng. Inform.*, vol. 46, Oct. 2020, Art. no. 101191.
- [103] D. J. Hemanth, J. Anitha, and L. H. Son, "Brain signal based human emotion analysis by circular back propagation and deep kohonen neural networks," *Comput. Electr. Eng.*, vol. 68, pp. 170–180, May 2018.
- [104] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation," *Sci. World J.*, vol. 2014, Sep. 2014, Art. no. 627892.
- [105] Z. K. Gao, Y. L. Li, Y. X. Yang, X. M. Wang, N. Dong, and H. D. Chiang, "A GPSO-optimized convolutional neural networks for EEG-based emotion recognition," *Neurocomputing*, vol. 380, pp. 225–235, Mar. 2020.
- [106] Z. Yin and J. H. Zhang, "Task-generic mental fatigue recognition based on neurophysiological signals and dynamical deep extreme learning machine," *Neurocomputing*, vol. 283, pp. 266–281, Mar. 2018.
- [107] Y. L. Ma *et al.*, "Driving fatigue detection from EEG using a modified PCANet method," *Comput. Intell. Neurosci.*, vol. 2019, no. 3, pp. 1–9, 2019.
- [108] H. Zeng, C. Yang, G. J. Dai, F. W. Qin, J. H. Zhang, and W. Z. Kong, "EEG classification of driver mental states by deep learning," *Cogn. Neurodyn.*, vol. 12, no. 6, pp. 597–606, 2018.
- [109] Z. K. Gao *et al.*, "EEG-based spatio-temporal convolutional neural network for driver fatigue evaluation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2755–2763, Sep. 2019.
- [110] R. F. Chai *et al.*, "Improving EEG-based driver fatigue classification using sparse-deep belief networks," *Front. Neurosci.*, vol. 11, p. 103, Mar. 2017.
- [111] E. Q. Wu, X. Y. Peng, C. Z. Z. Zhang, J. X. Lin, and R. S. F. Sheng, "Pilots' fatigue status recognition using deep contractive autoencoder network," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 3907–3919, Oct. 2019.
- [112] J. D. Tao, Z. Yin, L. Liu, Y. Tian, Z. Q. Sun, and J. H. Zhang, "Individual-specific classification of mental workload levels via an ensemble heterogeneous extreme learning machine for EEG modeling," *Symmetry*, vol. 11, no. 7, p. 944, 2019.
- [113] J. H. Zhang and S. Li, "A deep learning scheme for mental workload classification based on restricted Boltzmann machines," *Cogn. Technol. Work*, vol. 19, no. 4, pp. 607–631, 2017.
- [114] R. Hefron, B. Borghetti, C. S. Kabban, J. Christensen, and J. Estep, "Cross-participant EEG-based assessment of cognitive workload using multi-path convolutional recurrent neural networks," *Sensors*, vol. 18, no. 5, p. 1339, 2018.
- [115] Z. Yin and J. H. Zhang, "Cross-subject recognition of operator functional states via EEG and switching deep belief networks with adaptive weights," *Neurocomputing*, vol. 260, pp. 349–366, Oct. 2017.
- [116] P. B. Zhang, X. Wang, W. H. Zhang, and J. F. Chen, "Learning spatial-spectral-temporal EEG features with recurrent 3D convolutional neural networks for cross-task mental workload assessment," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 1, pp. 31–42, Jan. 2019.
- [117] Z. Yin, M. Y. Zhao, W. Zhang, Y. X. Wang, Y. G. Wang, and J. H. Zhang, "Physiological-signal-based mental workload estimation via transfer dynamical autoencoders in a deep learning framework," *Neurocomputing*, vol. 347, pp. 212–229, Jun. 2019.
- [118] P. B. Zhang, X. Wang, J. F. Chen, W. You, and W. H. Zhang, "Spectral and temporal feature learning with two-stream neural networks for mental workload assessment," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1149–1159, Jun. 2019.
- [119] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 324–333, Feb. 2018.
- [120] E. Bresch, U. Grosseckthofer, and G. Garcia-Molina, "Recurrent deep neural networks for real-time sleep stage classification from single channel EEG," *Front. Comput. Neurosci.*, vol. 12, p. 85, Oct. 2018.
- [121] P. M. Kulkarni *et al.*, "A deep learning approach for real-time detection of sleep spindles," *J. Neural Eng.*, vol. 16, no. 3, p. 36004, 2019.
- [122] O. Tsinialis, P. M. Matthews, and Y. K. Guo, "Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders," *Ann. Biomed. Eng.*, vol. 44, no. 5, pp. 1587–1597, May 2016.

- [123] S. Hartmann and M. Baumert, "Automatic a-phase detection of cyclic alternating patterns in sleep using dynamic temporal information," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 9, pp. 1695–1703, Sep. 2019.
- [124] J. M. Zhang, R. X. Yao, W. G. Ge, and J. F. Gao, "Orthogonal convolutional neural networks for automatic sleep stage classification based on single-channel EEG," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 105089.
- [125] Y. Jeon *et al.*, "Pediatric sleep stage classification using multi-domain hybrid neural networks," *IEEE Access*, vol. 7, pp. 96495–96505, 2019.
- [126] E. Eldele *et al.*, "An attention-based deep learning approach for sleep stage classification with single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 809–818, Apr. 2021.
- [127] J. M. Zhang and Y. Wu, "Complex-valued unsupervised convolutional neural networks for sleep stage classification," *Comput. Methods Programs Biomed.*, vol. 164, pp. 181–191, Oct. 2018.
- [128] J. M. Zhang, Y. Wu, J. Bai, and F. Q. Chen, "Automatic sleep stage classification based on sparse deep belief net and combination of multiple classifiers," *Trans. Inst. Meas. Control*, vol. 38, no. 4, pp. 435–451, 2016.
- [129] M. A. Jaoude *et al.*, "Expert-level automated sleep staging of long-term scalp electroencephalography recordings using deep learning," *Sleep*, vol. 43, no. 11, Nov. 2020, Art. no. zsaal12.
- [130] Z. Mousavi, T. Y. Rezaii, S. Sheykhiwand, A. Farzamia, and S. N. Razavi, "Deep convolutional neural network for classification of sleep stages from single-channel EEG signals," *J. Neurosci. Methods*, vol. 324, Aug. 2019, Art. no. 108312.
- [131] P. Chriskos, C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, and C. Kourtidou-Papadeli, "Automatic sleep staging employing convolutional neural networks and cortical connectivity images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 113–123, Jan. 2020.
- [132] O. Ozdenizci, Y. Wang, T. Koike-Akino, and D. Erdoğmuş, "Adversarial deep learning in EEG biometrics," *IEEE Signal Process. Lett.*, vol. 26, no. 5, pp. 710–714, May 2019.
- [133] M. Wang, H. El-Fiqi, J. K. Hu, and H. A. Abbass, "Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states," *IEEE Trans. Inf. Forensics Security*, vol. 14, pp. 3259–3272, 2019.
- [134] P. Kaushik, A. Gupta, P. P. Roy, and D. P. Dogra, "EEG-based age and gender prediction using deep BLSTM-LSTM network model," *IEEE Sensors J.*, vol. 19, no. 7, pp. 2634–2641, Apr. 2019.
- [135] N. Huang, M. Slaney, and M. Elhilali, "Connecting deep neural networks to physical, perceptual, and electrophysiological auditory signals," *Front. Neurosci.*, vol. 12, p. 532, Aug. 2018.
- [136] M. X. Yu *et al.*, "Diverse frequency band-based convolutional neural networks for tonic cold pain assessment using EEG," *Neurocomputing*, vol. 378, pp. 270–282, Feb. 2020.
- [137] J. M. Jiang, A. Fares, and S. H. Zhong, "A context-supported deep learning framework for multimodal brain imaging classification," *IEEE Trans. Human-Mach. Syst.*, vol. 49, no. 6, pp. 611–622, Dec. 2019.
- [138] X. Zheng, W. Z. Chen, M. Y. Li, T. Zhang, Y. You, and Y. Jiang, "Decoding human brain activity with deep learning," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101730.
- [139] C. C. Hua, H. Wang, H. Wang, S. W. Lu, C. Liu, and S. M. Khalid, "A novel method of building functional brain network using deep learning algorithm with application in proficiency detection," *Int. J. Neural Syst.*, vol. 29, no. 1, 2019, Art. no. 1850015.
- [140] S. Toraman, S. A. Tuncer, and F. Balgetir, "Is it possible to detect cerebral dominance via EEG signals by using deep learning?" *Med. Hypotheses*, vol. 131, Oct. 2019, Art. no. 109315.
- [141] V. Baltatzis, K. M. Binti, 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," *Sci. Rep.*, vol. 7, Dec. 2017, Art. no. 17292.
- [142] Z. G. Dobarjeh, N. Kasabov, M. G. Dobarjeh, and A. Sumich, "Modelling peri-perceptual brain processes in a deep learning spiking neural network architecture," *Sci. Rep.*, vol. 8, p. 8912, Jun. 2018.
- [143] A. Arora *et al.*, "Comparison of logistic regression, support vector machines, and deep learning classifiers for predicting memory encoding success using human intracranial EEG recordings," *J. Neural Eng.*, vol. 15, no. 6, 2018, Art. no. 66028.
- [144] T. K. K. Ho, J. Gwak, C. M. Park, and J. I. Song, "Discrimination of mental workload levels from multi-channel fNIRS using deep learning-based approaches," *IEEE Access*, vol. 7, pp. 24392–24403, 2019.
- [145] J. Harvy, E. Sigalas, N. Thakor, A. Bezerianos, and J. Li, "Performance improvement of driving fatigue identification based on power spectra and connectivity using feature level and decision level fusions," in *Proc. 40th IEEE Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, 2018, pp. 102–105.
- [146] Z. Pei, H. Wang, A. Bezerianos, and J. Li, "EEG-based multiclass workload identification using feature fusion and selection," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–8, Aug. 2020.
- [147] W. Wu *et al.*, "Multimodal vigilance estimation using deep learning," *IEEE Trans. Cybern.*, early access, Oct. 7, 2020, doi: 10.1109/TCYB.2020.3022647.
- [148] J. Li, A. Bezerianos, and N. Thakor, "Cognitive state analysis, understanding, and decoding from the perspective of brain connectivity," 2020. [Online]. Available: arXiv:2005.12191.
- [149] S. K. Goh *et al.*, "Spatio-spectral representation learning for electroencephalographic gait-pattern classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1858–1867, Sep. 2018.
- [150] J. Li, "Thoughts on neurophysiological signal analysis and classification," *Brain Sci. Adv.*, vol. 6, no. 3, pp. 210–223, 2020.
- [151] X. L. Ma, S. Qiu, W. Wei, S. P. Wang, and H. G. He, "Deep channel-correlation network for motor imagery decoding from the same limb," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 297–306, Jan. 2020.
- [152] X. Y. Zhu, P. Y. Li, C. B. Li, D. Z. Yao, R. Zhang, and P. Xu, "Separated channel convolutional neural network to realize the training free motor imagery BCI systems," *Biomed. Signal Process. Control*, vol. 49, pp. 396–403, Mar. 2019.
- [153] A. M. Chiarelli, P. Croce, A. Merla, and F. Zappasodi, "Deep learning for hybrid EEG-fNIRS brain-computer interface: Application to motor imagery classification," *J. Neural Eng.*, vol. 15, no. 3, 2018, Art. no. 36028.
- [154] Z. Tayeb *et al.*, "Validating deep neural networks for online decoding of motor imagery movements from EEG signals," *Sensors*, vol. 19, no. 1, p. 210, 2019.
- [155] M. X. Dai, D. Z. Zheng, R. Na, S. Wang, and S. L. Zhang, "EEG classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, no. 3, p. 551, 2019.
- [156] K.-W. Ha and J.-W. Jeong, "Motor imagery EEG classification using capsule networks," *Sensors*, vol. 19, no. 13, p. 2854, 2019.
- [157] T. W. Shi, L. Ren, and W. H. Cui, "Feature recognition of motor imaging EEG signals based on deep learning," *Pers. Ubiquitous Comput.*, vol. 23, nos. 3–4, pp. 499–510, 2019.
- [158] Z. J. Wang, L. Cao, Z. Zhang, X. L. Gong, Y. R. Sun, and H. R. Wang, "Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition," *Concurrency Comput. Pract. Exp.*, vol. 30, no. 23, 2018, Art. no. e4413.
- [159] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Bencherif, and M. S. Hossain, "Multilevel weighted feature fusion using convolutional neural networks for EEG motor imagery classification," *IEEE Access*, vol. 7, pp. 18940–18950, 2019.
- [160] R. T. Schirrmeyer *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Map.*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [161] A. Hassanpour, M. Moradikia, H. Adeli, S. R. Khayami, and P. Shamsinejadbabaki, "A novel end-to-end deep learning scheme for classifying multi-class motor imagery electroencephalography signals," *Expert Syst.*, vol. 36, no. 6, 2019, Art. no. e12494.
- [162] Q. S. She, B. Hu, Z. Z. Luo, T. Nguyen, and Y. C. Zhang, "A hierarchical semi-supervised extreme learning machine method for EEG recognition," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 147–157, 2019.
- [163] L. F. S. Uribe *et al.*, "A correntropy-based classifier for motor imagery brain-computer interfaces," *Biomed. Phys. Eng. Exp.*, vol. 5, no. 6, 2019, Art. no. 65026.
- [164] L. J. Duan, M. H. Bao, S. Cui, Y. H. Qiao, and J. Miao, "Motor imagery EEG classification based on kernel hierarchical extreme learning machine," *Cogn. Comput.*, vol. 9, no. 6, pp. 758–765, 2017.
- [165] H. Wu *et al.*, "A parallel multiscale filter bank convolutional neural networks for motor imagery EEG classification," *Front. Neurosci.*, vol. 13, p. 1275, Nov. 2019.
- [166] I. Majidov and T. Whangbo, "Efficient classification of motor imagery electroencephalography signals using deep learning methods," *Sensors*, vol. 19, no. 7, p. 1736, 2019.
- [167] X. L. Tang, W. C. Ma, D. S. Kong, and W. Li, "Semisupervised deep stacking network with adaptive learning rate strategy for motor imagery EEG recognition," *Neural Comput.*, vol. 31, no. 5, pp. 919–942, 2019.

- [168] B. G. Xu *et al.*, "Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification," *IEEE Access*, vol. 7, pp. 6084–6093, 2019.
- [169] O.-Y. Kwon, M.-H. Lee, C. Guan, and S.-W. Lee, "Subject-independent brain-computer interfaces based on deep convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 3839–3852, Oct. 2020.
- [170] N. Mammone, C. Ieracitano, and F. C. Morabito, "A deep CNN approach to decode motor preparation of upper limbs from time-frequency maps of EEG signals at source level," *Neural Netw.*, vol. 124, pp. 357–372, Apr. 2020.
- [171] D. Zhang, K. Chen, D. Jian, and L. Yao, "Motor imagery classification via temporal attention cues of graph embedded EEG signals," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 9, pp. 2570–2579, Sep. 2020.
- [172] J. Chen, Z. Yu, Z. Gu, and Y. Li, "Deep temporal-spatial feature learning for motor imagery-based brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 11, pp. 2356–2366, Nov. 2020.
- [173] J.-H. Jeong, K.-H. Shim, D.-J. Kim, and S.-W. Lee, "Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1226–1238, May 2020.
- [174] W. Gao, J. A. Guan, J. F. Gao, and D. Zhou, "Multi-ganglion ANN based feature learning with application to P300-BCI signal classification," *Biomed. Signal Process. Control*, vol. 18, pp. 127–137, Apr. 2015.
- [175] A. Farahat, C. Reichert, C. M. Sweeney-Reed, and H. Hinrichs, "Convolutional neural networks for decoding of covert attention focus and saliency maps for EEG feature visualization," *J. Neural Eng.*, vol. 16, no. 6, 2019, Art. no. 66010.
- [176] A. J. Solon, V. J. Lawhern, J. Touryan, J. R. McDaniel, A. J. Ries, and S. M. Gordon, "Decoding P300 variability using convolutional neural networks," *Front. Human Neurosci.*, vol. 13, p. 201, Jun. 2019.
- [177] L. Vareka and P. Mautner, "Stacked autoencoders for the P300 component detection," *Front. Neurosci.*, vol. 11, p. 302, May 2017.
- [178] S. Morabbi, M. Keyvanpour, and S. V. Shojadini, "A new method for P300 detection in deep belief networks: Nesterov momentum and drop based learning rate," *Health Technol.*, vol. 9, no. 4, pp. 615–630, 2019.
- [179] A. Ditthaporn, N. Banluesombatkul, S. Kettrat, E. Chuangsuwanich, and T. Wilaiprasitporn, "Universal joint feature extraction for P300 EEG classification using multi-task autoencoder," *IEEE Access*, vol. 7, pp. 68415–68428, 2019.
- [180] L. Citi, R. Poli, and C. Cinel, "Documenting, modelling and exploiting p300 amplitude changes due to variable target delays in Donchin's speller," *J. Neural Eng.*, vol. 7, no. 5, 2010, Art. no. 56006.
- [181] M. Schreuder, B. Blankertz, and M. Tangermann, "A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue," *PLoS One*, vol. 5, no. 4, 2010, Art. no. e9813.
- [182] L. Acqualagna and B. Blankertz, "Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP)," *Clin. Neurophysiol.*, vol. 124, no. 5, pp. 901–908, 2013.
- [183] M. S. Treder, H. Purwins, D. Miklody, I. Sturm, and B. Blankertz, "Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification," *J. Neural Eng.*, vol. 11, no. 2, 2014, Art. no. 026009.
- [184] B. Boloukian and F. Safi-Esfahani, "Recognition of words from brain-generated signals of speech-impaired people: Application of autoencoders as a neural Turing machine controller in deep neural networks," *Neural Netw.*, vol. 121, pp. 186–207, Jan. 2020.
- [185] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [186] J. X. Chen, Z. J. Mao, R. Zheng, Y. F. Huang, and L. F. He, "Feature selection of deep learning models for EEG-based RSVP target detection," *IEICE Trans. Inf. Syst.*, vol. E102D, no. 4, pp. 836–844, 2019.
- [187] J. Touryan, G. Apker, S. Kerick, B. Lance, A. J. Ries, and K. McDowell, "Translation of EEG-based performance prediction models to rapid serial visual presentation tasks," in *Proc. Int. Conf. Augmented Cogn.*, 2013, pp. 521–530.
- [188] R. Manor and A. B. Geva, "Convolutional neural network for multi-category rapid serial visual presentation BCI," *Front. Comput. Neurosci.*, vol. 9, p. 146, Dec. 2015.
- [189] R. Manor, L. Mishali, and A. B. Geva, "Multimodal neural network for rapid serial visual presentation brain computer interface," *Front. Comput. Neurosci.*, vol. 10, p. 130, Dec. 2016.
- [190] Q. Q. Liu *et al.*, "Efficient representations of EEG signals for SSVEP frequency recognition based on deep multiset CCA," *Neurocomputing*, vol. 378, pp. 36–44, Feb. 2020.
- [191] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PLoS One*, vol. 10, no. 10, 2015, Art. no. e0140703.
- [192] X. J. Bi and H. B. Wang, "Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning," *Neural Netw.*, vol. 114, pp. 119–135, Jun. 2019.
- [193] F. C. Morabito *et al.*, "Deep learning representation from electroencephalography of early-stage Creutzfeldt-Jakob disease and features for differentiation from rapidly progressive dementia," *Int. J. Neural Syst.*, vol. 27, no. 2, 2017, Art. no. 1650039.
- [194] K. Hayase, K. Hayashi, and T. Sawa, "Hierarchical Poincaré analysis for anaesthesia monitoring," *J. Clin. Monitor. Comput.*, vol. 34, pp. 1321–1330, Dec. 2020.
- [195] Q. Liu *et al.*, "Spectrum analysis of EEG signals using CNN to model patient's consciousness level based on anesthesiologists' experience," *IEEE Access*, vol. 7, pp. 53731–53742, 2019.
- [196] Y. Park, S.-H. Han, W. Byun, J.-H. Kim, H.-C. Lee, and S.-J. Kim, "A real-time depth of anesthesia monitoring system based on deep neural network with large EDO tolerant EEG analog front-end," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 825–837, Aug. 2020.
- [197] S. Kim, J. Kim, and H. W. Chun, "Wave2Vec: Vectorizing electroencephalography bio-signal for prediction of brain disease," *Int. J. Environ. Res. Public Health*, vol. 15, no. 8, p. 1750, 2018.
- [198] H. Chen, Y. Song, and X. L. Li, "Use of deep learning to detect personalized spatial-frequency abnormalities in EEGs of children with ADHD," *J. Neural Eng.*, vol. 16, no. 6, 2019, Art. no. 066046.
- [199] R. Boshra, K. I. Ruitter, C. DeMatteo, J. P. Reilly, and J. F. Connolly, "Neurophysiological correlates of concussion: Deep learning for clinical assessment," *Sci. Rep.*, vol. 9, Nov. 2019, Art. no. 17341.
- [200] H. Q. Sun *et al.*, "Automated tracking of level of consciousness and delirium in critical illness using deep learning," *NPJ Digit. Med.*, vol. 2, p. 89, Sep. 2019.
- [201] B. Ay *et al.*, "Automated depression detection using deep representation and sequence learning with EEG signals," *J. Med. Syst.*, vol. 43, no. 7, p. 205, 2019.
- [202] 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," *Comput. Methods Programs Biomed.*, vol. 161, pp. 103–113, Jul. 2018.
- [203] U. R. Acharya, S. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, and D. P. Subha, "Automated EEG-based screening of depression using deep convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 161, pp. 103–113, Jul. 2018.
- [204] X. W. Li *et al.*, "Depression recognition using machine learning methods with different feature generation strategies," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101696.
- [205] J. Zhu *et al.*, "Multimodal mild depression recognition based on EEG-EM synchronization acquisition network," *IEEE Access*, vol. 7, pp. 28196–28210, 2019.
- [206] G. Bouallegue, R. Djemal, S. A. Alshebeili, and H. Aldhalaan, "A dynamic filtering DF-RNN deep-learning-based approach for EEG-based neurological disorders diagnosis," *IEEE Access*, vol. 8, pp. 206992–207007, 2020.
- [207] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Nonlinear Soft Matter Phys.*, vol. 64, no. 6, 2001, Art. no. 061907.
- [208] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Phys. Rev. E, Stat. Nonlinear Soft Matter Phys.*, vol. 86, no. 4, 2012, Art. no. 046206.
- [209] A. M. Karim, M. S. Guzel, M. R. Tolun, H. Kaya, and F. V. Celebi, "A new framework using deep auto-encoder and energy spectral density for medical waveform data classification and processing," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 148–159, 2019.
- [210] D. K. Thara, B. G. PremaSudha, and F. Xiong, "Epileptic seizure detection and prediction using stacked bidirectional long short term memory," *Pattern Recognit. Lett.*, vol. 128, no. 2, pp. 529–535, 2019.
- [211] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "Neuro-Detect: A machine learning-based fast and accurate seizure detection system in the IoMT," *IEEE Trans. Consum. Electron.*, vol. 65, no. 3, pp. 359–368, Aug. 2019.

- [212] M. Stead *et al.*, "Microseizures and the spatiotemporal scales of human partial epilepsy," *Brain*, vol. 133, no. 9, pp. 2789–2797, 2010.
- [213] B. H. Brinkmann, M. R. Bower, K. A. Stengel, G. A. Worrell, and M. Stead, "Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data," *J. Neurosci. Methods*, vol. 180, no. 1, pp. 185–192, 2009.
- [214] R. Hussein, H. Palangi, R. K. Ward, and Z. J. Wang, "Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals," *Clin. Neurophysiol.*, vol. 130, no. 1, pp. 25–37, 2019.
- [215] A. M. Abdelhameed and M. Bayoumi, "Semi-supervised EEG signals classification system for epileptic seizure detection," *IEEE Signal Process. Lett.*, vol. 26, no. 12, pp. 1922–1926, Dec. 2019.
- [216] R. Akut, "Wavelet based deep learning approach for epilepsy detection," *Health Inf. Sci. Syst.*, vol. 7, p. 8, Apr. 2019.
- [217] A. Emami, N. Kunii, T. Matsuo, T. Shinozaki, K. Kawai, and H. Takahashi, "Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images," *Neuroimage-Clinical*, vol. 22, Jan. 2019, Art. no. 101684.
- [218] U. B. Baloglu and O. Yildirim, "Convolutional long-short term memory networks model for long duration EEG signal classification," *J. Mech. Med. Biol.*, vol. 19, no. 1, 2019, Art. no. 1940005.
- [219] A. O'Shea, G. Lightbody, G. Boylan, and A. Temko, "Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture," *Neural Netw.*, vol. 123, pp. 12–25, Mar. 2020.
- [220] C. Jansen, S. Hodel, T. Penzel, M. Spott, and D. Krefting, "Feature relevance in physiological networks for classification of obstructive sleep apnea," *Physiol. Meas.*, vol. 39, no. 12, 2018, Art. no. 124003.
- [221] G. Klosh *et al.*, "The SIESTA project polygraphic and clinical database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 51–57, May/June 2001.
- [222] S. Jonas, A. O. Rossetti, M. Oddo, S. Jenni, P. Favaro, and F. Zubler, "EEG-based outcome prediction after cardiac arrest with convolutional neural networks: Performance and visualization of discriminative features," *Human Brain Map.*, vol. 40, no. 16, pp. 4606–4617, 2019.
- [223] M. C. Tjepkema-Cloostermans *et al.*, "Outcome prediction in postanoxic coma with deep learning," *Crit. Care Med.*, vol. 47, no. 10, pp. 1424–1432, 2019.
- [224] S. U. Amin, M. S. Hossain, G. Muhammad, M. Alhussein, and M. A. Rahman, "Cognitive smart healthcare for pathology detection and monitoring," *IEEE Access*, vol. 7, pp. 10745–10753, 2019.
- [225] G. Ruffini *et al.*, "Deep learning with EEG spectrograms in rapid eye movement behavior disorder," *Front. Neurol.*, vol. 10, p. 806, Jul. 2019.
- [226] A. Piryatinska, B. Darkhovsky, and A. Kaplan, "Binary classification of multichannel-EEG records based on the  $\epsilon$ -complexity of continuous vector functions," *Comput. Methods Programs Biomed.*, vol. 152, pp. 131–139, Dec. 2017.
- [227] C.-R. Phang, F. Noman, H. Hussain, C.-M. Ting, and H. Ombao, "A multi-domain connectome convolutional neural network for identifying schizophrenia from EEG connectivity patterns," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 5, pp. 1333–1343, May 2020.
- [228] Z. Yin, M. Y. Zhao, Y. X. Wang, J. D. Yang, and J. H. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Comput. Methods Programs Biomed.*, vol. 140, pp. 93–110, Mar. 2017.
- [229] T. Chen *et al.*, "Emotion recognition using empirical mode decomposition and approximation entropy," *Comput. Electr. Eng.*, vol. 72, pp. 383–392, Nov. 2018.
- [230] S.-K. Kim and H.-B. Kang, "An analysis of smartphone overuse recognition in terms of emotions using brainwaves and deep learning," *Neurocomputing*, vol. 275, pp. 1393–1406, Jan. 2018.
- [231] J. Teo, L. H. Chew, J. T. Chia, and J. Mountstephens, "Classification of affective states via EEG and deep learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 5, pp. 132–142, 2018.
- [232] J. X. Chen, D. M. Jiang, and N. Zhang, "A hierarchical bidirectional GRU model with attention for EEG-based emotion classification," *IEEE Access*, vol. 7, pp. 118530–118540, 2019.
- [233] O. Balan, G. Moise, A. Moldoveanu, M. Leordeanu, and F. Moldoveanu, "Fear level classification based on emotional dimensions and machine learning techniques," *Sensors*, vol. 19, no. 7, p. 1738, 2019.
- [234] Y. Cimtay and E. Ekmekcioglu, "Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset EEG emotion recognition," *Sensors*, vol. 20, p. 2034, Apr. 2020.
- [235] Y. Kim and A. Choi, "EEG-based emotion classification using long short-term memory network with attention mechanism," *Sensors*, vol. 20, no. 23, p. 6727, Nov. 2020.
- [236] B. H. Kim and S. Jo, "Deep physiological affect network for the recognition of human emotions," *IEEE Trans. Affective Comput.*, vol. 11, no. 2, pp. 230–243, Apr.–Jun. 2020.
- [237] J. H. Jeong, B. W. Yu, D. H. Lee, and S. W. Lee, "Classification of drowsiness levels based on a deep spatio-temporal convolutional bidirectional LSTM network using electroencephalography signals," *Brain Sci.*, vol. 9, no. 12, p. 348, 2019.
- [238] Z. Yin and J. H. Zhang, "Cross-session classification of mental workload levels using EEG and an adaptive deep learning model," *Biomed. Signal Process. Control*, vol. 33, pp. 30–47, Mar. 2017.
- [239] Z. C. Jiao, X. B. Gao, Y. Wang, J. Li, and H. J. Xu, "Deep convolutional neural networks for mental load classification based on EEG data," *Pattern Recognit.*, vol. 76, pp. 582–595, Apr. 2018.
- [240] S. Yang, Z. Yin, Y. G. Wang, W. Zhang, Y. X. Wang, and J. H. Zhang, "Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders," *Comput. Biol. Med.*, vol. 109, pp. 159–170, Jun. 2019.
- [241] O. Yildirim, U. B. Baloglu, and U. R. Acharya, "A deep learning model for automated sleep stages classification using PSG signals," *Int. J. Environ. Res. Public Health*, vol. 16, no. 4, p. 599, 2019.
- [242] A. Patanaik, J. L. Ong, J. J. Gooley, S. Ancoli-Israel, and M. W. L. Chee, "An end-to-end framework for real-time automatic sleep stage classification," *Sleep*, vol. 41, no. 5, 2018, Art. no. zsy041.
- [243] Y. Yuan *et al.*, "A hybrid self-attention deep learning framework for multivariate sleep stage classification," *BMC Bioinform.*, vol. 20, p. 586, Dec. 2019.
- [244] L. D. Zhang, D. Fabbri, R. Uppender, and D. Kent, "Automated sleep stage scoring of the Sleep Heart Health Study using deep neural networks," *Sleep*, vol. 42, no. 11, 2019, Art. no. zsz159.
- [245] F. Chapotot and G. Becq, "Automated sleep-wake staging combining robust feature extraction, artificial neural network classification, and flexible decision rules," *Int. J. Adapt. Control Signal Process.*, vol. 24, no. 5, pp. 409–423, 2010.
- [246] A. Malafeev *et al.*, "Automatic human sleep stage scoring using deep neural networks," *Front. Neurosci.*, vol. 12, p. 781, Nov. 2018.
- [247] H. Phan, F. Andreotti, N. Cooray, O. Y. Chen, and M. De Vos, "Joint classification and prediction CNN framework for automatic sleep stage classification," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1285–1296, May 2019.
- [248] 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 Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 4, pp. 758–769, Apr. 2018.
- [249] C. L. Rosen *et al.*, "A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: The HomePAP study," *Sleep*, vol. 35, no. 6, pp. 757–767, 2012. [Online]. Available: <https://academic.oup.com/sleep/article-pdf/35/6/757/26619641/aasm.35.6.757.pdf>
- [250] J. P. Bakker *et al.*, "Gastric banding surgery versus continuous positive airway pressure for obstructive sleep apnea: A randomized controlled trial," *Amer. J. Respir. Crit. Care Med.*, vol. 197, pp. 1080–1083, Apr. 2018.
- [251] S. Biswal, H. Q. Sun, B. Goparaju, M. B. Westover, J. M. Sun, and M. T. Bianchi, "Expert-level sleep scoring with deep neural networks," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 12, pp. 1643–1650, 2018.
- [252] 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," *Biomed. Signal Process. Control*, vol. 42, pp. 107–114, Apr. 2018.
- [253] J. B. Blank *et al.*, "Overview of recruitment for the osteoporotic fractures in men study (MrOS)," *Contemp. Clin. Trials*, vol. 26, no. 5, pp. 557–568, 2005.
- [254] S. Redline *et al.*, "The childhood adenotonsillectomy trial (CHAT): Rationale, design, and challenges of a randomized controlled trial evaluating a standard surgical procedure in a pediatric population," *Sleep*, vol. 34, no. 11, pp. 1509–1517, 2011.
- [255] M. H. AlMeer, H. Hassen, and N. Nawaz, "ROM-based inference method built on deep learning for sleep stage classification," *J. Technol. Educ. Manag. Inform.*, vol. 8, no. 1, pp. 28–40, 2019.
- [256] W. Qu *et al.*, "A residual based attention model for EEG based sleep staging," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 10, pp. 2833–2843, Oct. 2020.
- [257] S. Chambon, V. Thorey, P. J. Arnal, E. Mignot, and A. Gramfort, "DOSED: A deep learning approach to detect multiple sleep micro-events in EEG signal," *J. Neurosci. Methods*, vol. 321, pp. 64–78, Jun. 2019.

- [258] O. Andlauer *et al.*, “Nocturnal rapid eye movement sleep latency for identifying patients with narcolepsy/hypocretin deficiency,” *JAMA Neurol.*, vol. 70, no. 7, pp. 891–902, 2013.
- [259] T. Young *et al.*, “Sleep disordered breathing and mortality: Eighteen-year follow-up of the Wisconsin Sleep Cohort,” *Sleep*, vol. 31, no. 8, pp. 1071–1078, 2008.
- [260] B. Kaur, D. Singh, and P. P. Roy, “Age and gender classification using brain–computer interface,” *Neural Comput. Appl.*, vol. 31, no. 10, pp. 5887–5900, 2019.
- [261] D. F. Wulsin, J. R. Gupta, R. Mani, J. A. Blanco, and B. Litt, “Modeling electroencephalography waveforms with semi-supervised deep belief nets: Fast classification and anomaly measurement,” *J. Neural Eng.*, vol. 8, no. 3, 2011, Art. no. 036015.
- [262] J. Anem, G. S. Kumar, and R. Madhu, “Cat swarm fractional calculus optimization-based deep learning for artifact removal from EEG signal,” *J. Exp. Theor. Artif. Intell.*, vol. 32, no. 6, pp. 939–958, 2020.
- [263] S. Jacob, V. G. Menon, F. Al-Turjman, P. G. Vinoj, and L. Mostarda, “Artificial muscle intelligence system with deep learning for post-stroke assistance and rehabilitation,” *IEEE Access*, vol. 7, pp. 133463–133473, 2019.
- [264] B. H. Yang, K. W. Duan, C. C. Fan, C. X. Hu, and J. L. Wang, “Automatic ocular artifacts removal in EEG using deep learning,” *Biomed. Signal Process. Control*, vol. 43, pp. 148–158, May 2018.
- [265] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly, and M. Shah, “Deep learning human mind for automated visual classification,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 6809–6817.
- [266] P. Croce, F. Zappasodi, L. Marzetti, A. Merla, V. Pizzella, and A. M. Chiarelli, “Deep convolutional neural networks for feature-less automatic classification of independent components in multi-channel electrophysiological brain recordings,” *IEEE Trans. Biomed. Eng.*, vol. 66, no. 8, pp. 2372–2380, Aug. 2019.
- [267] Y. R. Ming *et al.*, “Subject adaptation network for EEG data analysis,” *Appl. Soft Comput.*, vol. 84, Nov. 2019, Art. no. 105689.
- [268] P. Nagabushanam, S. T. George, and S. Radha, “EEG signal classification using LSTM and improved neural network algorithms,” *Soft Comput.*, vol. 24, no. 12, pp. 9981–10003, 2020.
- [269] A. Fares, S. H. Zhong, and J. M. Jiang, “EEG-based image classification via a region-level stacked bi-directional deep learning framework,” *BMC Med. Inform. Decis. Making*, vol. 19, p. 268, Dec. 2019.
- [270] H. Akbari, B. Khalighinejad, J. L. Herrero, A. D. Mehta, and N. Mesgarani, “Towards reconstructing intelligible speech from the human auditory cortex,” *Sci. Rep.*, vol. 9, p. 874, Jan. 2019.
- [271] A. Antoniadou *et al.*, “Deep neural architectures for mapping scalp to intracranial EEG,” *Int. J. Neural Syst.*, vol. 28, no. 8, 2018, Art. no. 1850009.
- [272] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekart, and C. D. Buckingham, “A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction,” *Complexity*, vol. 2019, no. 1, 2019, Art. no. 4316548.
- [273] V. Singhal, A. Majumdar, and R. K. Ward, “Semi-supervised deep blind compressed sensing for analysis and reconstruction of biomedical signals from compressive measurements,” *IEEE Access*, vol. 6, pp. 545–553, 2018.
- [274] A. Gogna, A. Majumdar, and R. Ward, “Semi-supervised stacked label consistent autoencoder for reconstruction and analysis of biomedical signals,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2196–2205, Sep. 2017.
- [275] H. J. Jang and K. O. Cho, “Dual deep neural network-based classifiers to detect experimental seizures,” *Korean J. Physiol. Pharmacol.*, vol. 23, no. 2, pp. 131–139, 2019.
- [276] T. Ogawa, Y. Sasaka, K. Maeda, and M. Haseyama, “Favorite video classification based on multimodal bidirectional LSTM,” *IEEE Access*, vol. 6, pp. 61401–61409, 2018.
- [277] A. B. Said *et al.*, “A deep learning approach for vital signs compression and energy efficient delivery in mhealth systems,” *IEEE Access*, vol. 6, pp. 33727–33739, 2018.



**Shu Gong** is currently pursuing the B.E. degree in software engineering and the B.S. degree in biology from Sichuan University, Chengdu, China.

His research interests cover a wide range of neural interface, brain–machine interface, and deep learning.



**Kaibo Xing** is currently pursuing the B.E. degree in computational biology from the College of Life Sciences, Sichuan University, Chengdu, China.

His current research interests include brain–computer interface and machine learning.



**Andrzej Cichocki** (Fellow, IEEE) received the M.Sc. (Hons.), Ph.D., and Dr.Sc. (Habilitation) degrees in electrical engineering from Warsaw University of Technology, Warszawa, Poland.

He spent several years with the University of Erlangen-Nuremberg, Erlangen, Germany, as an Alexander-von-Humboldt Research Fellow and a Guest Professor. He was a Senior Team Leader and the Head of the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Saitama, Japan. He is currently a Professor with the Skolkovo Institute of Science and Technology, Moscow, Russia. He has authored more than 500 technical journal papers and six monographs in English (two of them translated to Chinese). His publications currently report over 46 000 citations according to Google Scholar, with an H-index of 97. His current research focus on multiway blind source separation, tensor decompositions, tensor networks, deep learning, human–robot interactions, and brain–computer interface.

Prof. Cichocki served as an Associate Editor for IEEE TRANSACTION ON SIGNALS PROCESSING, IEEE TRANSACTION ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTION ON CYBERNETICS, and *Journal of Neuroscience Methods*. He is the Founding Editor-in-Chief of *Journal Computational Intelligence and Neuroscience*.



**Junhua Li** (Senior Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2013.

He is currently a Lecturer with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, U.K. He was a Senior Research Fellow with the National University of Singapore, Singapore, before moving to the U.K. His research interests include computational neuroscience, brain–computer interface, machine learning, neurophysiological data analytics, and their practical applications.

Dr. Li is an Associate Editor for IEEE ACCESS, a Review Editor for *Frontiers in Human Neuroscience*, and served as a Guest Associate Editor for several journal special issues related to his research interests.