Received: 25 October 2018 | Revised: 14 December 2018 | Accepted: 16 January 2019
DOI: 10.1002/cpe.5199

WILEY

SPECIAL ISSUE PAPER

Deep learning for EEG data analytics: A survey

Gen Li[1] | Chang Ha Lee[1] | Jason J. Jung[1] | Young Chul Youn[2] | David Camacho[3]

[1]Department of Computer Engineering, Chung-Ang University, Seoul, South Korea
[2]Neurology, Chung-Ang University Hospital, Seoul, South Korea
[3]Department of Computer Engineering, Universidad Autonoma de Madrid, Madrid, Spain

Correspondence
Jason J. Jung, Department of Computer Engineering, Chung-Ang University, Seoul 06974, South Korea.
Email: j2jung@gmail.com

Funding information
Chung-Ang University, Grant/Award Number: Research Grants in 2017; National Research Foundation of Korea (NRF); Korea Government (MSIP), Grant/Award Number: NRF-2018K1A3A1A38056595

Summary

In this work, we conducted a literature review about deep learning (DNN, RNN, CNN, and so on) for analyzing EEG data for decoding the activity of human's brain and diagnosing disease and explained details about various architectures for understanding the details of CNN and RNN. It has analyzed a word, which presented a model based on CNN and LSTM methods, and how these methods can be used to both optimize and set up the hyper parameters of deep learning architecture. Later, it is studied how semi-supervised learning on EEG data analytics can be applied. We review some studies about different methods of semi-supervised learning on EEG data analytics and discussing the importance of semi-supervised learning for analyzing EEG data. In this paper, we also discuss the most common applications for human EEG research and review some papers about the application of EEG data analytics such as Neuromarketing, human factors, social interaction, and BCI. Finally, some future trends of development and research in this area, according to the theoretical background on deep learning, are given.

KEYWORDS

clinical AI-based diagnosis, deep learning, EEG signal processing

1 | INTRODUCTION

In the era of "big data," transformation of large quantities of data into valuable knowledge has become increasingly important in various domains,[1,2] such as the image recognition,[3] speech recognition,[4] and the EEG signals are also included in it. With the current exponential growth of the amount of data available, the large number of different formats, and the increasing computational power, and taking into account the expectations generated by Artificial Intelligence, as a new powerful tool to the service of humans and companies, many studies began focusing on EEG's research. For example, IBM designed a platform for giving some treatment options for clinicians by analyzing the medical information of patient.[5,6] Before the deep learning become popular, most of researchers prefer scripting the algorithm to extract valuable information from EEG signals and computing these information using machine learning techniques; this is due to the machine learning algorithms demonstrating an excellent performance when dealing with different kind of real, complex, and dynamic problems (using different techniques, as those based on regression, classification, or unsupervised learning such as clustering). Until recently, in the big environment of deep learning popularity, the advantage of accuracy of deep learning on some fields of industry or academic, studies on deep learning for various fields are receiving extensive attention on all countries in the world.

Deep learning was proposed by Hinton et al,[7] inspired by human's way of thinking. Deep learning forms more abstract high levels representations by combining low layer features to represent attribute categories or features to discover distributed characteristics over data. It based on deep belief network (DBN); an unsupervised greedy layer-by-layer training algorithm based on Deep Confidence Network (DBN) is proposed to solve the optimization problems related to deep structure. With the development of big data, the advantage of deep learning, due to its flexibility to fit complex models and its high accuracy, has become one of the best methods in several Big Data-based problems. From Figure 1,[8] it is not difficult to find that the interest of researchers has increased rapidly since 2000s. Furthermore, deep learning is responsible for major advances in diverse fields where the Artificial Intelligence (AI) community has struggled for many years,[2,9] such as image and speech recognition[10-12] or some fields related to natural language processing like language translation,[13] or sentiment analysis,[14] amongst many others. Currently, deep learning can also benefit to decoding the EEG signals.

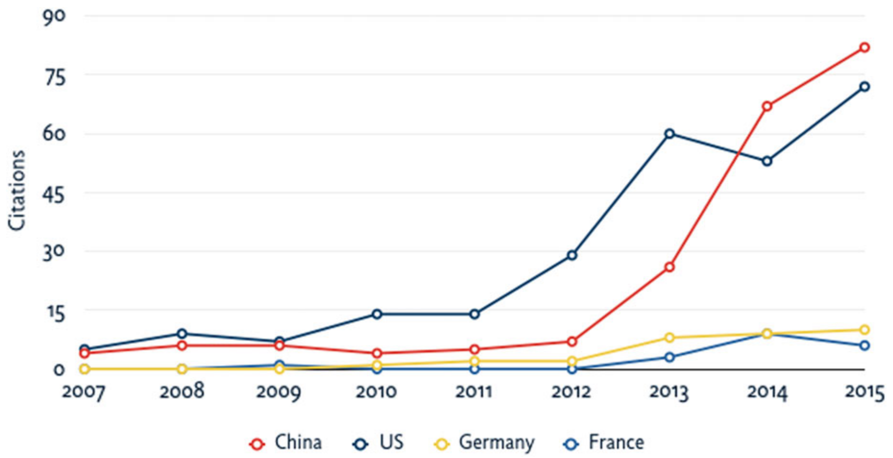

FIGURE 1 Number of publications about deep learning for reference of countries: China, United States, Germany, and France[8]

EEG signals processing is the recording of electrical signals from human brain to decode human's behavior based on their brain activity; most of studies focus on EEG processing so far, while it can help us to understand the relationships between brain activity and electrical signals. We can use traditional algorithms from machine learning, like deep learning, to analyze, learn, and extract complex patterns from these complex signals. Because the signals recorded are usually some kind of mixed noise and artifacts combinations, people usually transfer this raw data into a wavelet or frequency before using it as input data. However, with the continuous improvement of the Convolutional Neural Networks (CNN), raw EEG signals have been used for anomaly classification issue and brain activity's decoding, such as the work of Stober et al[15] classified the rhythm and genre of music, which performed to experimenter, and the work of Cecotti and Graser[16] detected the characters that experimenter viewed by CNN for EEG signal analysis. Tang et al,[17] Lawhern et al,[18] and Sun et al[19] have discussed CNN for EEG to understand the behavior of human's brain. In this paper, we will will analyze the basics on deep neural networks architectures, whereas some relevant papers will be revised to show up the current state of the art in this area; finally, we will discuss the futures trends of deep learning for EEG signal processing to finally discuss about some limitations and weakness of deep learning.

The main contributions of this work are summarized as follows.

1. To understand the current research trends on deep learning for EEG signal processing and analytics by conducting literature review. The different available architectures and applications will revised.
2. How non-fixed EEG data can be analyzed by deep learning. It will be described a new model, built by combining CNN and LSTM models, developed by authors and it will be discussed how this model could be employed over other time-serial data.
3. Considering the application of EEG analysis and EEG signal processing by reviewing some studies about the most popular applications like Neuromarketing, human factors, social interaction, and BCI. We also surveyed some papers about the application of EEG analysis on clinical such as Alzheimer's disease and epileptic seizure.

The rest of this paper is organized as follows. In Section 2, we conduct a literature review on deep learning for EEG data analytics. In Section 3, we discuss about a new model created by CNN and LSTM techniques and consider how it could be trained by using EEG data. In Section 4, we review some current applications and research on EEG processing. Some discussions and future works are concluded in Section 5.

2 | INTRODUCTION AND METHODOLOGIES FOR EEG DATA ANALYSIS

2.1 | Introduction to EEG data analysis

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain; research on brain signal processing and analysis have been a popular topic from the last years, and a large amount of methods and techniques have been proposed to handle with EEG data. In the initial stages of this area, the researchers directly extracted the information by recording the EEG signals to later decode the brain activity. The machine learning methods raised their relevance when they were applied to specific application problems, such as the brain-computer interface (BCI) systems for clinical domains. The current state of research on deep learning has shown outstanding results on both computer vision and image processing.[20] Due to the nature of EEG problems and the basic mathematical and computer-based features of deep learning, it would be expected that, in the near future, these kind of methods would be the mainstream research technique on EEG signal processing.

TABLE 1 Survey papers on CNN for EEG signal processing

Band	Frequency(Hz)	Means
Delta	Less than 4	Deep sleep without dreamming
Theta	4-7	When adults are emotionally stressed, especially disappointment or frustration.
Alpha	8-15	Relax, calm, and close your eyes but wake up
Beta(Low Range)	12.5-16	Relax but concentrate
Beta(Middle Range)	16.5-20	Thinking, processing and receiving information from the outside world
Beta(High Range)	20.5-28	Excitement and anxiety
Gamma	25-100	Raise awareness, happiness, stress, and meditation
Lambda	Evoked potential	When the eyes are stimulated by light, they are induced after 100 ms (also known as P100)
P300	Evoked potential	When you see or hear what is imagined in your brain, 300 ms induces it.

2.1.1 | Type of EEG signal processing and analysis method

Modern scientific research shows that the human brain works with spontaneous electrophysiological activity. This activity can be expressed in form of brain waves by a special EEG recoder. There are at least four important bands in the study of EEG. EEG is a spontaneous rhythmic electrical activity with frequency range of 1 to 30 times per second; it can be divided into four bands, ie, δ (1 to 4Hz), θ (4 to 7Hz), α (8 to 15Hz), and, β (12.5 to 28Hz). In addition, when awaking or focusing on something, it is often seen that a γ wave with higher frequency than β wave, which has a frequency of 30 to 80 Hz and the amplitude is uncertain. While sleeping, there are other normal brain waves with special waveforms, such as hump wave, α wave, λ wave, κ-complex wave, and μ wave. Table 1 shows the Band, Frequencies, and some other details related to the EEG signal and what it could mean related to the human brain behaviour.

Therefore, it is possible to induce different human behaviors and states from the analysis of the waveforms; the most common approach is based on the observation of peak values and frequencies on EEG signals to finally decode the brain activity or to diagnosing a disease. Therefore, and before training any machine learning algorithm by using EEG data, the researchers usually generate a specific code (in form of a script) to extract the peak values, the frequencies of the wave, or to transform raw EEG data into a spectrum diagram based on Fourier transform, or a wavelet transform. Later, this data is used to train a classification model. However, deep learning techniques can directly extract features from the data by automatically set up and updating weights according to back propagation approach and reducing the effect of noise in raw data. This is one of the main reasons why deep learning and CNN methods have been widely used for brain decoding or diagnosing based on EEG data.

2.2 | Machine learning methods for EEG data

The less amount of data and lower performance computer lead to the deep learning has not improved on data analytics earlier; at that time, most people prefer using the algorithms of machine learning to training data, such like naive Bayes or support machine learning. Early approaches attempted to explicitly program the required knowledge for given tasks; however, these faced difficulties in dealing with complex real-world problems because designing all the detail required for an AI system to accomplish satisfactory results by band is such a demanding job.[21] As we know, machine learning has a good capacity on train AI system to have a cognitive ability through by experience, which is learned by training large number data; however, there are limitations on feature extraction from raw data; sometimes people have to spent lots of time to script a complex algorithm with capacity of extracting feature from raw data by hand. Although these algorithms have some limit, many studies were used machine learning to analysis EEG data. Wang et al[22] have used SVM to make a classification for classify human's emotion based on three kinds of EEG features and monitor the changes of emotional states in real time. They found (1) that the spectrum feature is superior than other two features, (2) using the feature smoothing method to improve a linear dynamic system could improve the accuracy of classification, and (3) the changes of emotion could be monitored by reducing the subject-independent features with manifold learning. Müller et al[23] have used machine learning method to train the EEG data when analyzing single-trial data in real-time and made a application for monitoring the BCI and mental state. Shoeb and Guttag[24] have used machine learning method for seizure detection.

2.3 | Deep neural networks for EEG data

Since mixing with noise and artifacts in recorded EEG signals, used features are more frequent than raw signals. In brain decoding, the works of Jia et al[25] and Jirayucharoensak et al[26] have used DBN and SAE to model a EEG signal's classification for emotion detection. Zheng et al[27] have presented the critical channels and frequency bands, which related with emotion recognition by decoding EEG signals with DBN. An et al[28] have classified left- and right-hand motor imagery skills by analyzing the frequency factor of EEG, which applied DBN. In the other classification of deep neural networks for EEG,[29-32] Wulsin et al[29] have used semi-supervised deep belief nets to make a fast classification and detection anomaly

TABLE 2 Summarize of deep learning for different EEG type

Study	EEG type	Architecture	Decoding problem
Wang et al[22]	Frequency spectrum	Deep belief network (DBN)	Emotion detection
Müller et al[23]	Frequency spectrum	SAE	Emotion detection
Shoeb and Guttag[24]	Frequency spectrum	DBN	Emotion recognition
Jia et al[25]	Frequency spectrum	Deep neural network	Left- and right- hand motor
Jirayucharoensak et al[26]	Frequency spectrum	Semi-supervised DBN	Detection of anomaly measurement
Zheng et al[27]	Raw data	DBN	Seizure detection
An et al[28]	Raw data	DNN	Alzheimer's disease
Wulsin et al[29]	Frequency spectrum	Unsupervised learning	Classification of sleep stage
Turner et al[30]	Frequency spectrum	DBN	Prediction of driver's states

measurement, and Längkvist et al[32] have used unsupervised feature learning to classify the human's sleep stage; Hajinoroozi et al[33] have used DBN to extract the feature from EEG signals for cognitive the driver's states.

Since the influence of the noise, as shown in Table 2, a few study used raw EEG data in deep neural networks. Turner et al[30] have computed high resolution multichannel EEG data for seizure detection by DBN; Zhao and He[31] have built a system for diagnosis Alzheimer's disease by deep learning for EEG signals.

2.4 | Deep recurrent networks for EEG data

EEG signals are also sequential data, and RNN is one of the architecture to train the sequential processing and has a good performance on sequential processing, especially on NLP(nature language processing); it plays a important role. For training EEG data, RNN could better to extract the feature information from EEG data. As usual, we have to transform the raw EEG data to the frequency spectrogram features for the input of RNN. Many studies focus on diagnosis of disease or prediction and brain decoding such as emotion detection.

Petrosian et al[34] have predicted seizures by applied RNN to raw EEG data and corresponding wavelet features. Davidson et al[35] have transformed the EEG data to spectra features and used LSTM to detect lapses. Minasyan et al[36] have developed a method for detection of seizures prior to or immediately after clinical onset using features derived from scalp EEG data. Naderi and Mahdavi-Nasab[37] have proposed a three-stage technique for seizure detection in EEG signals; the first method used Welch method power spectrum density estimation to extract the features from EEG signals, second, using statistics to reduced the dimensionality of features and time series signals samples, and the third stage is make a classification by RNN.

Some studies about brain decoding and anomaly classification are presented. Soleymani et al[38] and Fomey and Anderson[39] have proposed the Elman Recurrent neural networks a new architecture of RNN to classify EEG signals during imagined mental tasks. Li et al[40] have recognized MI-EEG combined with LSTM based network employing. Patnaik et al[41] have used wavelet transform to deal with extraction of EEG features for predict the human's brain state. Ni et al[42] have used LSTM for EEG data for disentangling brain activity of human, like prediction of statement of confused or not confused. In summary, most of studies about Recurrent neural networks for EEG always focus on emotion recognition, disease diagnosis, and brain decoding.

2.5 | Deep convolutional networks for EEG data

Convolutional neural networks have a good performance on image processing tasks recently, since it has good capacity on feature extraction from images by convolution kernel; they can extract the information features through passing every part of the images with number of kernels. For EEG signals, it not only works well on raw EEG data, but also used on frequency spectrum diagram. There, we categorized the research studies in EEG signals processing into two groups, namely, brain decoding and diagnose by anomaly classification.

2.5.1 | Brain decoding

Using CNN for training EEG data can reduce the effect from noisy; thus, we can use raw EEG signals data for CCN's input, which can reduce the complexity of training; therefore, most studies used CNN for EEG signals. For instance, the work of Cecotti and Graser[16] has classified characters that are viewed by participants. Tang et al[17] have presented a new approach to extract the feature and classify the single-trial MI EEG. Sun et al[19] have proposed a computational method to detect memory performance of remembered or forgotten by training EEG data while memory processing. Thodoroff et al[43] have combined CNN and RNN to train robust features to automatically detect seizures. Shamwell et al[44] have explored a new CNN architecture with 4 convolution layers and 3 full connect layers to generalize multi-class single-trial EEG classification across subjects, aimed to increase human-autonomy classification performance. Manor and Geva[45] presented a CNN model for the use for classify single trail EEG in RSVP (Rapid Serial Visual Presentation), and in order to reduce the overfitting of model, they approached a novel spatio-temporal regularization; finally, they compared the feature extraction by CNN and manually designing feature extraction algorithms. Sakhavi et al[46] have presented a Parallel convulutional-linear network, which is an architecture that can make EEG data as a dynamic energy representation for input and utilizes CNN for imagery classification. Ren and Wu[47] have applied convolutional deep belief networks learning features from EEG data and evaluated it based on comparison with the datasets from BCI. Ruffini et al[48] have collected data from idiopathic RBD patients and healthy controls and proposed CNN for classifying rapid eye movement behavior disorder prognosis. Hajinoroozi et al[49] have used covariance learning to train EEG data for driver's fatigue prediction. Jiao et al[50] have proposed improved CNNs methods for mental load classification task.

2.5.2 | Diagnose diseases

EEG can also reflect the health of the many human diseases; the detection of lots of diseases need to be diagnosed by anomaly EEG signals analysis, which is recorded when the diseases occurred. Thus, a system, which can provide a good accuracy of diagnose, has become a direction for variety fields. CNN for training EEG data can reduce the effect from noise; therefore, most studies used CNN for EEG signals to diagnose diseases by anomaly signals classification.

Mirowski et al[51] have extracted features as phase-locking synchrony and wavelet coherence and coded them as pixel colors to formulate two-dimensional patterns.[2] Liang et al[52] have adopted EEG datasets, which are not directly related to seizure prediction and training by deep learning to detected seizure. Antoniades et al[53] have considered generating feature automatically from epileptic intracranial EEG data in time domain by deep leaning. Page et al[54] have made an end-to-end learning by max-pooling convolutional neural networks (MPCNN) and demonstrated that transfer-learning can be used to teach MPCNNs generalized features of raw EEG data. Acharya et al[55] have presented a deep convolutional neural networks with 5 layers to detected normal, preictal, and seizure classes.

Thus, it can be seen that the development of deep learning for EEG signals processing are increasing, and the research studies of deep learning are roughly divided into disease diagnosis, such as epileptic seizures prediction or Alzheimer prediction, and emotion recognition and brain decoding, like driver statement prediction or brain statement prediction. Compared with traditional algorithm of feature extraction, deep learning can automatically extract features and reduce the effect from noise. That is the reason why many studies begin to focus on deep learning for EEG signals processing. There, we make Table 3 to list some details of CNN architectures, decoding problems, and input domain by survey from part of papers.

2.6 | Semi-supervised learning for EEG data

Semi-supervised learning is the key problem in the field of pattern recognition and machine learning. It is a learning method by combing supervised learning with unsupervised learning. In some practical problems, there are only a few labeled data, because the cost of marking data is sometimes very high. These data are like EEG and so on, medical data, or some biological data. However, semi-supervised learning learning uses a large number of unlabeled data and uses labeled data at the same time for pattern recognition. The usually used methods are like Self training, Generative model, S3VMs, and so on. Li et al[40] have presented a self-training semi-supervised SVM method for classifying P300 data, and using this algorithm would reduce the training effort of the P300 data. Wulsin et al[56] have applied DBNs in a semi-supervised learning to model EEG data for classification and prediction. Shi et al[57] have used semi-supervised clustering method to analyze vigilance based on EEG data. Wulsin et al[56] have used semi-supervised learning for anomaly detection based on EEG waveforms.

3 | DEEP LEARNING ARCHITECTURE

Many studies have presented different architectures of deep learning for EEG data analytics that including deep believe network, deep convolutional neural network, or recurrent neural network. In this section, we will explain the CNN model and LSTM model and talking about some details by studying the paper of Bashivan et al.[58]

3.1 | Convolutional neural networks

On account of a great ability on feature extraction from multidimensional dataset, CNN has achieved great success in many recognition issue like image recognition or behavior recognition through extracting information from photo. Recently, researchers found that the CNN also has a good capacity on sequential data like sounds wave, and it is not worse than recurrent neural networks on sequential information processing, and it has become one of the popular architecture on nature language processing, that is because it can extract a detail feature better from raw data of sound wave by convolution operation. Being different with CNN, for RNN, one of the neuron output value can not only pass to the next neuron, but also act on itself; that means that the output of the current moment has been combined the experience of this time and the history. That is the reason why RNN has a greatly success for sequential data. However, for CNN, the essence of convolution operation uses a filter with weights to pass the whole picture for extracting the features of each part of pictures; in nature language processing, it just like a window of N-gram model; this feature information are useful for nature language processing. Thus, many researchers have begun to use CNN to analyze

TABLE 3 Survey papers of CNN for EEG signals

Study	Decoding problem	Input domain	Conv/dense layers
Luong et al[13]	Type of music rhythm	Time 0.5-39Hz	3/3
Stober et al[15]	Imaged movement classes	Time,8-30Hz	2/2
Tang et al[17]	Memory performance	Time,0.05-15Hz	2/2
Jirayucharoensak et al[26]	Dirver performance	Time, 1-50Hz	1/3
Davidson et al[35]	Start of epileptic seizure	Frequency, mean amplitude for 0-7Hz,7-14Hz,14-49Hz	3/1(+LSTM)
Minasyan et al[36]	Oddball response	Time, 05-50Hz	4/3
Naderi and Mahdavi-Nasab[37]	Oddball response using RSVP and image	Time,0.3-20Hz	3/2
Soleymani et al[38]	Imagined movement classes	Frequency 4-40Hz	2/2
Forney and Anderson[39]	Imagined movement classes	Frequency 8-30Hz	2/0
Patnaik et al[41]	Eye movement classes	Stacked multi-channel spectrograms	4/1
Ni et al[42]	Driver performance	Frequency 256Hz	2/2
Manor and Geva[45]	Seizure prediction	Frequency, 0-200Hz	1/2
Sakhavi et al[46]	Epileptic discharge	Time, 1-50Hz	1/3
Ren and Wu[47]	Seizure detection	Time, 0-128Hz	1-3/1-3
Ruffini et al[48]	Seizure detection	Frequency, 4-7Hz,8-13Hz,13-30Hz	5/1

the EEG signals and achieved some results. The type of EEG signals are same as speech data performed by a wave; in the same way, we believe that we can decode the sequential information of EEG signals by using the CNN for deal with problems of human's behavior or disease diagnosis.

3.2 | Recurrent neural network

Although the traditional multilayer perception based on the above networks structure has excellent performance and applied in many fields, these always have some defects, which consists none of the above models that can analyze the overall logical sequence between the input information. These logical sequences are rich in content and have a complex time relationship with each other. In order to solve these sequence problems, recurrent neural networks arise at the historic moment; the key is that hidden state of the current network will retain the previous input information, and it is used for the next current network. Compared with DNN, RNN has a closed loop; in other words, after passing the value to the next hidden layer, it always give this value to itself; it makes this network as a memory. We say recurrent neural networks have the ability to remember, and this ability is to sum up the past input state through weight value, as the auxiliary of the next input. It is possible to understand hidden state in this way; the hidden state is equal to a function, which combine with existing input and the summary of past memory.

For the training of recurrent neural networks, the spread of network is based on input, and the longer the input, the deeper the network is. Thus, for the training of RNN, we always face the problems of gradient explode and gradient vanish; in order to refrain this problem, the work Hochreiter and Schmidhuber[59] has proposed a new architecture of RNN named Long Short-Term Memory (LSTM). Being different from traditional RNN model, LSTM increased the forget gate and update gate. The forget gate can let model know which information should be saved and which information should be lost; the update gate will learn whether there is information worth using and saving. Therefore, while new data is inputing, the model will forget no useful informations; then, learning valuable informations from new input data and save on long memory, finally, model will learn which parts of the long-term memory can be used immediately.

3.3 | Problem formulation and architecture

EEG analyzing for detection could be generalized as time series classification issue; there, model should extract the useful information from varied length EEG data and classify to correct class. The input of this task is varied length signal $X = [x_1, x_2, \ldots x_k]$ and outputs the correct label. Thus, the objective of our model is to make minimization of cross entropy which between output label and given label. The object function given by

$$loss(X, g) = -log \frac{exp(p(X, g)}{\sum_j exp(p(X, t))}, \tag{1}$$

where $p(X, t)$ is the probability of the model training, the label t by given the input X, and g is given label.

Bashivan et al[58] have presented a model by combing LSTM and CNN for learning representations from EEG data; the model architecture is shown as Figure 2; they used 3 convolution layers to extracted the spectrogram features and got one EEG image set, and then pooling to get temporal features, and then training these by LSTM. Connected with the last pooling layer is LSTM model proposed by Hochreiter and Schmidhuber.[59] The LSTM model include some cells, these are like units of deep neural networks with some parameters; therefore, while new data is inputing, the LSTM model will forget no useful informations; then, learning valuable informations from new input data and save on long memory. That is the reason why the LSTM model can learning series time data better. The cells of LSTM behavior is concluded by four equations

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_{t-1}] + b_f\right) \tag{2}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_{t-1}] + b_i\right) \tag{3}$$

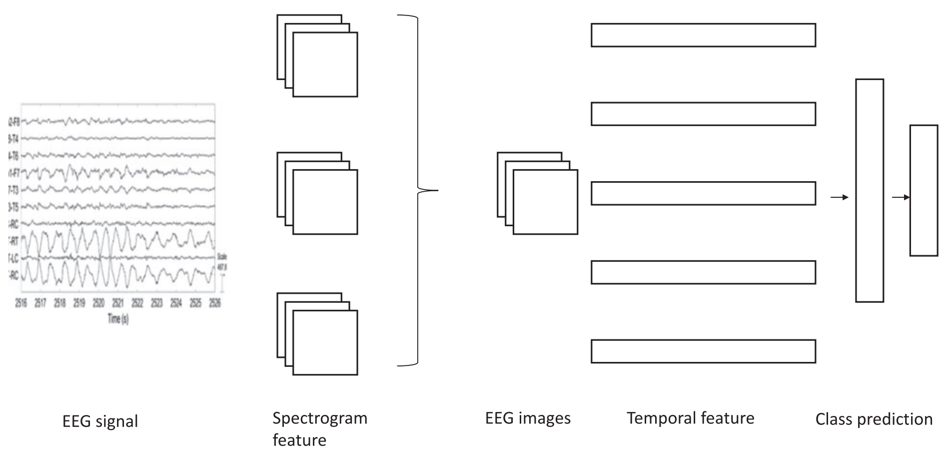

FIGURE 2 Deep learning model by combing with CNN and LSTM

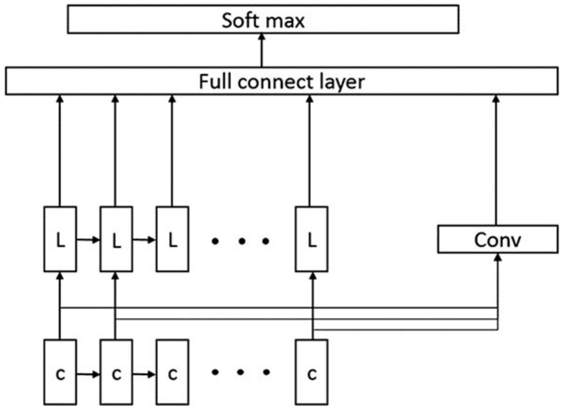

FIGURE 3 Connection between Conv layer and full connect layer. C: ConvNet, L: LSTM layer

$$C_t = f_t * C_{t-1} + i_t * tanh\left(W_c \cdot [h_{t-1}, x_{t-1}] + b_c\right) \tag{4}$$

$$h_t = \sigma\left(W_0[h_{t-1}, x_{t-1}] + b_0\right) * tanh(C_t), \tag{5}$$

where $h_t, x_t, C - t$ are the output, input, and private state at time t of cell. The W_f, W_c, W_0 are the parameters, which need to be trained. These parameters made the cell to decide whether to remember or forget the information. In this model, as shown in Figure 3, they used CNN model for extracting the feature from EEG data and used LSTM model to training EEG data by extracted features and learning the representations.

3.4 | Hyper parameters

Hyper parameter's initialization is importance before model training dataset; it could prevent falling into a local optimal and affect the result in the processing of optimization. For the convolution layer, we could use kaiming initialize,[60] and for the LSTM cells, the orthogonal initializer[61] could be used, because they have shown a great ability on improvement of parameters's converging speed in these papers. For the entire model, the usually used optimization of deep learning is gradient descent, owing to huge computation; SGD (stochastic gradient descent)[62] was proposed; compared with GD, SGD has three advantages.

1. Intuitive motivation,
2. Practical motivation, and
3. Theoretical motivation.

On intuitive motivation, SGD can make use of information more effectively, especially when information is more redundant. On practical motivation, comparing with GD, SGD is excellent; in the previous iteration, there we can see some results in related papers.[63] On theoretical motivation, if the sample size is large, SGD computational complexity still has advantages by comparing with GD. In addition, Adam optimizer also is a better approach; Adam is a first-order optimization algorithm, which can replace the traditional SGD process; it can update neural network weights iteratively based on training data. Adam was proposed on 2015.[64] This paper lists some advantages of Adam.

1. A straightforward implementation,
2. The efficient calculation,
3. Less memory required,
4. Invariance of gradient diagonal scaling,
5. Suitable for solving optimization problems with large data and parameters,
6. Applicable to non-stationary targets,
7. Suitable for solving problems involving very high noise or sparse gradient, and
8. Super parameters can be interpreted intuitively and basically require only a very small number of parameters.

Therefore, Adam sometimes will be a better choice. It is combined the advantages of AdaGrad and RMSProp, and the implementation is simple; the calculation is efficient.

4 | EEG DATA ANALYTICS APPLICATIONS

Reviewing by some studies, the EEG data analysis can be applied for neuromarketing, human factors, social interaction, and brain computer interfaces. Researching for neuromarketing could understand the customer purchase psychology and the impact of advertising on customers.

Research of human factors is the decoding for brain statement like judging whether a driver is a tired driver or judging what a man sees. Social interaction is that understanding how other people's behavior will affect themselves. Brain computer interfaces (BCI) are the most popular research for analysis EEG data to build a direct connection between human brain and external devices; it allows the signals transformation between human's brain and external devices. This technology is often used for restoring hearing, visual, and limb movement abilities. With the development of EEG researching on medical science, EEG signal gets a lot of applications on clinical application such as Alzheimer's disease (AD) diagnosing, epilepsy diagnosing, and other brain disease diagnosing.

4.1 | Neuromarketing, human factors social interaction, and BCI

In the field of neuromarketing, economists always detect brain process by EEG research that includes the driving consumer decisions, brain areas that are active when consumer purchase a product or service, and so on. Yadava et al[65] have proposed a predictive model for classify the consumer's choice towards in "like" or "dislike" by EEG data analytics and used this model for understanding the decision of the consumer to judge the profit of products. Murugappan et al[66] have used KNN and Possibility neural network (PNN) methods to training EEG data for classifying the subject intention on advertisements for identifying the most preferred brand.

In this field of human factors, EEG data analytics is always used for identifying the brain processes related to specific personality traits such as intro-/extroversion or social anxiety. Gevins et al[67] have used neural networks pattern recognition applied to EEG spectrogram features for assessing the load of working memory and discussed the feasibility of the memory load monitoring. Smith et al[68] have applied the multivariate EEG methods for task loading monitoring during the naturalistic computer-based work.

In social interaction research, brain processing related to social perception, self-evaluation, and social behavior is investigated. Thus, EEG data analytics on social interaction plays an important role and many studies have researched the social interaction based on EEG data analytics. In Perry et al[69]'s paper, Mu rhythms are EEG oscillations in 8 to 13 Hz recorded at sites located roughly over the sensory-motor cortex; this paper found that Mu suppression not only responses to actual activities but also when the participant observes actions executed by other people.

A relatively new but emergent field for EEG is brain-computer interfaces. Today, we know in much more detail which brain areas are active when we perceive stimuli, when we prepare and execute bodily movements, or when we learn and memorize things. This gives rise to very powerful and targeted EEG applications to steer devices using brain activity, for instance, helping paralyzed patients steer their wheelchairs or move a cursor on a screen. Lotte et al[70] have surveyed some papers, which are about the classification algorithms used to the BCI design system-based EEG. This paper explained the feature extraction for BCI and reviewed some papers to introduced the methods of BCI features extraction like band powers (BP),[71] power spectral density (PSD),[72,73] and time-frequency features.[22] In addition, this paper surveyed classification methods for BCI research like linear discriminant analysis (LDA), support vector machine (SVM), Nonlinear bayesian classifiers, and some architectures of neural networks.

4.2 | Clinical application

EEG is more sensitive objective index, which not only can be used in the basic theory of brain science, but also is more important in the application of its clinical practice, which is closely related to human life and health. Thus, as shown in Table 4, EEG is the necessary basis for the diagnosis of epilepsy and AD. EEG also has great diagnostic help for various intracranial lesions, such as cerebral apoplexy, encephalitis, brain tumor, and metabolic encephalopathy. For the EEG processing by computer science, most studies are about the prediction of epilepsy or AD by machine learning algorithms such as Naive bayes and SVM. Recently, most studies focus on EEG processing by deep learning.

TABLE 4 Survey papers of EEG analysis on Clinical application

Study	Clinical problem	Methodologies
Goodfellow et al[21]	Epileptic seizure	Machine learning algorithms to make classification
An et al[28]	Alzheimer's disease	Deep learning to achieve unsupervised learning to extract the feature
Pfurtscheller et al[71]	Alzheimer's disease	PC LDA, LDA and various machine learning algorithms with linear and non-linear
Chiappa and Bengio[72]	Alzheimer's disease	Complementary biomaker algorithm
Millan and Mouriño[73]	Alzheimer's disease	Elman neural networks to draw feature and using LDA and SVM to make 2 classifications
Lehmann et al[74]	Alzheimier's disease	SVM
Simpraga et al[75]	Epileptic seizure	Novel approach
Goli et al[76]	Epileptic seizure	SVM
Trambaiolli et al[77]	Epileptic seizure	Using hidden markov models to decode EEG data and give a classification by machine learning
Song et al[78]	Epileptic seizure	Using LSTM to predict seizure epileptic

4.2.1 | Clinical application of Alzheimer's disease

Lehmann et al[74] have explored various machine learning algorithms with the ability of linear and non-linear such as principal component linear discriminant analysis(PC LDA), partial least set squares LDA, principal component logistic regression, random forest, and SVM to discriminate between the EEGs of patients with different degree of Alzheimer's and their age-matched control subjects. Simpraga et al[75] have used complementary biomaker algorithms for EEG recording to get the signature of disease and pharmacological intervention, and then to improve the performance of distinguish, they used machine learning algorithms making a classification. Goli et al[76] have used Elman neural networks to draw out the optimal features and used linear discriminant analysis (LDA) and SVM to generate two classifications for mild Alzheimer's disease. Zhao and He[31] have used deep neural networks to archive the unsupervised learning to extracted the feature from 15 clinically diagnosed AD patients and 15 healthy people and train these features by SVM and get a 92% accuracy. Trambaiolli et al[77] have searched differentiation patients's patterns in EEG data by using SVM; they developed a new approach named quantitative EEG processing for diagnosing different patients with AD from generation individuals.

4.2.2 | Clinical application of epileptic seizure

Shoeb and Guttag[24] have used machine learning to construct patient classifier, which predict the onset of epileptic seizure by EEG analyzing. These studies point at the problem, which is that the brain's electrical activity is mixed with numerous classes with characteristics; they presented a new algorithm, which includes making a machine learning framework and identifying the useful features from other types of brain activity. Song et al[78] have presented a novel approach to detect the epileptic seizure automatically. This paper presented an optimized sample entropy algorithm and combined with extreme learning machine to judge the performance of EEG signals recording is the existence of normal or disease. However, the proposed method in this paper did not only achieve high detection accuracy but also have a high speed of computation. Tzimourta et al[79] have presented a method for automated seizure, which is based on detection discrete wavelet transform. This paper extracted five features from wavelet coefficients and training this features by SVM; for the result, they get a high accuracy about the seizure detection. Golmohammadi et al[80] gave a classification system with high performance and based on big data and machine learning environment. This method used hidden Markov models to decode the sequential data and post processing by deep neural networks. This system has three detections of clinical events; the first is spike and/or sharp waves, the second is periodic lateralized epileptiform discharges, and the third is generalized periodic epileptiform discharges. Tsiouris et al[81]'s paper used LSTM model to predicted the seizure and no seizures were missed with zero false predictions, and for the result, the proposed model has a better performance of seizure prediction. On the other hand, this paper also expands the CNN model for EEG data processing to predict the epileptic seizure.

5 | CONCLUSION AND EXPECTATION

As we enter environment of big data, deep learning plays a center stage for international academic and lots of interests. In EEG signals, especially, the deep learning models have shown various advantages to produce the expected results. In this paper, we have introduced some architecture of popular deep learning, type of EEG data, which is input data for deep learning, observed the advantage of deep learning for raw EEG data processing, and reviewed recent research papers about deep learning in various architecture for EEG signals processing. We study a model, which is presented on related paper by combing the CNN and LSTM. Then, we discussed the semi-supervised learning for EEG data such as some related papers of SSL for EEG data analytics and some methods of SSL. We list the most common applications for human EEG research and reviewed some studies, which researched applications for EEG data analytics, which includes neuromarketing, human factors, social interaction, and BCI. For the research of BCI, we discussed the feature extraction and training algorithms of BCI research based on EEG signal. We also reviewed some studies about the clinical application such as AD and epileptic seizures.

For the expectation of deep learning of EEG signals processing, incorporation of different deep learning architectures is a possible trends. For instance, the recent popular issues of like image captioning, video summarization, and image question answering are combined with CNN and RNN to be applied. Moreover, in this paper, we have already seen related papers, which combined these two architectures to learning representations. In addition, most of these papers, which we saw, have used supervised learning; in terms of EEG data processing, semi-supervised and reinforcement learning also have a good attention. Semi-supervised can learn unlabeled and labeled data in the area of EEG data processing; using semi-supervised learning not only requires the smallest number of personnel to do the work, but also at the same time, it can bring higher accuracy. Reinforcement learning also plays an important role in the artificial intelligence; it can adjust itself by reward function for completing the optimization if it is its own scheme. In the EEG signals, reinforcement can produce a kind of brain wave itself to simulate the brain waves produced by human activity in various environment and corresponding to different brain activity or statement; it may play an important role on BCI research in future.

Although deep learning has been used in many fields and got lots achievements, it does not always provide great results in EEG signals processing. The main reason is one of the prerequisites for maintaining high accuracy is using large amount of data to training deep learning model. However, the EEG data is rarely sometimes too little amount of dataset is not enough to fit a high accuracy model by deep learning; the other reason is the influence of noisy of EEG is obvious; although we could reduce these influences by CNN and LSTM or doing wavelet transform or FTT before training, but how to improve the model and mixing traditional algorithm to get a better result without a large of data and reduce the computation time will become one of the new issue in the now and future. We believe this paper will provide valuable information and as some new points for deep learning for EEG signals processing in future studies.

ACKNOWLEDGMENTS

This research was supported by the Chung-Ang University Research Grants in 2017. In addition, this work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2018K1A3A1A38056595).

ORCID

Jason J. Jung https://orcid.org/0000-0003-0050-7445

REFERENCES

1. Manyika J, Chui M, Brown B, et al. *Big data: the next frontier for innovation, competition, and productivity*. New York, NY: McKinsey Global Institute; 2011.
2. Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinform*. 2017;18(5):851-869.
3. Zhao G, Liu S, Wang Q, et al. Deep convolutional neural network for drowsy student state detection. *Concurrency Computat Pract Exper*. 2018;30(23):e4457.
4. Liu J-C, Leu F-Y, Lin G-L, Susanto H. An MFCC-based text-independent speaker identification system for access control. *Concurrency Computat Pract Exper*. 2018;30(2):e4255.
5. Ferrucci D, Brown E, Chu-Carroll J, et al. Building Watson: an overview of the deep QA project. *AI Mag*. 2010;31(3):59-79.
6. IBM. Oncology IBM Watson. 2018. http://www.ibm.com/
7. Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006;18(7):1527-1554.
8. van Etten Jaap. Datenna. 2017. http://www.datenna.com/industry/china-marching-forward-with-artificial-intelligence
9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444.
10. Farabet C, Couprie C, Najman L, LeCun Y. Learning hierarchical features for scene labeling. *IEEE Trans Pattern Anal Mach Intell*. 2013;35(8):1915-1929.
11. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. Paper presented at: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015; Boston, MA.
12. Chorowski JK, Bahdanau D, Serdyuk D, Cho K, Bengio Y. Attention-based models for speech recognition. In: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS); 2015; Montreal, Canada.
13. Luong M-T, Pham H, Manning CD. Effective approaches to attention-based neural machine translation. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing; 2015; Lisbon, Portugal.
14. Guan Z, Chen L, Zhao W, Zheng Y, Tan S, Cai D. Weakly-supervised deep learning for customer review sentiment classification. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI); 2016; New York, NY.
15. Stober S, Cameron DJ, Grahn JA. Using convolutional neural networks to recognize rhythm stimuli from electroencephalography recordings. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1 (NIPS); 2014; Montreal, Canada.
16. Cecotti H, Graser A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Trans Pattern Anal Mach Intell*. 2011;33(3):433-445.
17. Tang Z, Li C, Sun S. Single-trial EEG classification of motor imagery using deep convolutional neural networks. *Optik-Int J Light Electron Opt*. 2017;130:11-18.
18. Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. EEGNet: a compact convolutional network for EEG-based brain-computer interfaces. arXiv preprint arXiv:1611.08024. 2016.
19. Sun X, Qian C, Chen Z, Wu Z, Luo B, Pan G. Remembered or forgotten?– An EEG-based computational prediction approach. *PLoS One*. 2016;11(12):e0167497.
20. Chang J, Gu Y, Zhang Y. Chinese typeface transformation with hierarchical adversarial network. arXiv preprint arXiv:1711.06448. 2017.
21. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2016.
22. Wang X-W, Nie D, Lu B-L. Emotional state classification from EEG data using machine learning approach. *Neurocomputing*. 2014;129:94-106.
23. Müller, K-R, Tangermann M, Dornhege G, Krauledat M, Curio G, Blankertz B. Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. *J Neurosci Methods*. 2008;167(1):82-90.
24. Shoeb AH, Guttag JV. Application of machine learning to epileptic seizure detection. In: Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML); 2010; Haifa, Israel.
25. Jia X, Li K, Li X, Zhang A. A novel semi-supervised deep learning framework for affective state recognition on EEG signals. In: Proceedings of the 2014 IEEE International Conference on Bioinformatics and Bioengineering (BIBE); 2014; Boca Raton, FL.
26. Jirayucharoensak S, Pan-Ngum S, Israsena P. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *Sci World J*. 2014;2014.
27. Zheng W-L, Guo H-T, Lu B-L. Revealing critical channels and frequency bands for emotion recognition from EEG with deep belief network. In: Proceedings of the 7th International IEEE/EMBS Conference on Neural Engineering (NER); 2015; Montpellier, France.
28. An X, Kuang D, Guo X, Zhao Y, He L. A deep learning method for classification of EEG data based on motor imagery. In: *Intelligent Computing in Bioinformatics: 10th International Conference, ICIC 2014, Taiyuan, China, August 3-6, 2014. Proceedings*. Cham, Switzerland: Springer International Publishing; 2014:203-210.

29. Wulsin DF, Gupta JR, Mani R, Blanco JA, Litt B. Modeling electroencephalography waveforms with semi-supervised deep belief nets: fast classification and anomaly measurement. *J Neural Eng*. 2011;8(3):036015.
30. Turner JT, Page A, Mohsenin T, Oates T. Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection. Paper presented at: 2014 AAAI Spring Symposium Series; 2014; Palo Alto, CA. https://arxiv.org/abs/1708.08430
31. Zhao Y, He L. Deep Learning in the EEG diagnosis of Alzheimer's disease. In: *Computer Vision - ACCV 2014 Workshops*. Cham, Switzerland: Springer International Publishing; 2014:340-353.
32. Längkvist M, Karlsson L, Loutfi A. Sleep stage classification using unsupervised feature learning. *Adv Artif Neural Syst*. 2012;2012.
33. Hajinoroozi M, Jung T-P, Lin C-T, Huang Y. Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data. Paper presented at: 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP); 2015; Chengdu, China.
34. Petrosian A, Prokhorov D, Homan R, Dasheiff R, Wunsch II D. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial EEG. *Neurocomputing*. 2000;30(1-4):201-218.
35. Davidson PR, Jones RD, Peiris MTR. EEG-based lapse detection with high temporal resolution. *IEEE Trans Biomed Eng*. 2007;54(5):832-839.
36. Minasyan GR, Chatten JB, Chatten MJ, Harner RN. Patient-specific early seizure detection from scalp EEG. *J Clin Neurophysiol*. 2010;27(3):163-178.
37. Naderi MA, Mahdavi-Nasab H. Analysis and classification of EEG signals using spectral analysis and recurrent neural networks. Paper presented at: 2010 17th Iranian Conference of Biomedical Engineering (ICBME); 2010; Isfahan, Iran.
38. Soleymani M, Asghari-Esfeden S, Pantic M, Fu Y. Continuous emotion detection using EEG signals and facial expressions. Paper presented at: 2014 IEEE International Conference on Multimedia and Expo (ICME); 2014; Chengdu, China.
39. Forney EM, Anderson CW. Classification of EEG during imagined mental tasks by forecasting with Elman recurrent neural networks. Paper presented at: The 2011 International Joint Conference on Neural Networks; 2011; San Jose, CA.
40. Li Y, Guan C, Li H, Chin Z. A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. *Pattern Recogn Lett*. 2008;29(9):1285-1294.
41. Patnaik S, Moharkar L, Chaudhari A. Deep RNN learning for EEG based functional brain state inference. Paper presented at: 2017 International Conference on Advances in Computing, Communication and Control (ICAC3); 2017; Mumbai, India.
42. Ni Z, Yuksel AC, Ni X, Mandel MI, Xie L. Confused or not confused?: disentangling brain activity from EEG data using bidirectional LSTM recurrent neural networks. In: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics (ACM-BCB); 2017; Boston, MA.
43. Thodoroff P, Pineau J, Lim A. Learning robust features using deep learning for automatic seizure detection. In: Proceedings of the 1st Machine Learning for Healthcare Conference; 2016; Los Angeles, CA.
44. Shamwell J, Lee H, Kwon H, Marathe AR, Lawhern V, Nothwang W. Single-trial EEG RSVP classification using convolutional neural networks. In: Proceedings of the SPIE Volume 9836, Micro- and Nanotechnology Sensors, Systems, and Applications VIII; 2016; Baltimore, MD.
45. Manor R, Geva AB. Convolutional neural network for multi-category rapid serial visual presentation BCI. *Front Comput Neurosci*. 2015;9:146.
46. Sakhavi S, Guan C, Yan S. Parallel convolutional-linear neural network for motor imagery classification. Paper presented at: 2015 23rd European Signal Processing Conference (EUSIPCO); 2015; Nice, France.
47. Ren Y, Wu Y. Convolutional deep belief networks for feature extraction of EEG signal. Paper presented at: 2014 International Joint Conference on Neural Networks (IJCNN); 2014; Beijing, China.
48. Ruffini G, Soria DI, Dubreuil L, Gagnon J-F, Montplaisir J, Soria-Frisch A. Deep learning using EEG spectrograms for prognosis in idiopathic rapid eye movement behavior disorder (RBD). bioRxiv. 2018. https://doi.org/10.1101/240267
49. Hajinoroozi M, Zhang JM, Huang Y. Driver's fatigue prediction by deep covariance learning from EEG. Paper presented at: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 2017; Banff, Canada.
50. Jiao Z, Gao X, Wang Y, Li J, Xu H. Deep convolutional neural networks for mental load classification based on EEG data. *Pattern Recogn*. 2018;76:582-595.
51. Mirowski P, Madhavan D, LeCun Y, Kuzniecky R. Classification of patterns of EEG synchronization for seizure prediction. *Clin Neurophysiol*. 2009;120(11):1927-1940.
52. Liang J, Lu R, Wang F, Zhang C. Predicting seizures from electroencephalography recordings: a knowledge transfer strategy. Paper presented at: 2016 IEEE International Conference on Healthcare Informatics (ICHI); 2016; Chicago, IL.
53. Antoniades A, Spyrou L, Took CC, Sanei S. Deep learning for epileptic intracranial EEG data. Paper presented at: 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP); 2016; Vietri sul Mare, Italy.
54. Page A, Shea C, Mohsenin T. Wearable seizure detection using convolutional neural networks with transfer learning. Paper presented at: 2016 IEEE International Symposium on Circuits and Systems (ISCAS); 2016; Montreal, Canada.
55. Acharya UR, Oh SL, Hagiwara Y, Tan J-H, Adeli H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput Biol Med*. 2018;100:270-278.
56. Wulsin D, Blanco J, Mani R, Litt B. Semi-supervised anomaly detection for EEG waveforms using deep belief nets. Paper presented at: 2010 Ninth International Conference on Machine Learning and Applications; 2010; Washington, DC.
57. Shi L-C, Yu H, Lu B-L. Semi-supervised clustering for vigilance analysis based on EEG. Paper presented at: 2007 International Joint Conference on Neural Networks; 2007; Orlando, FL.
58. Bashivan P, Rish I, Yeasin M, Codella N. Learning representations from EEG with deep recurrent-convolutional neural networks. arXiv preprint arXiv:1511.06448. 2015.
59. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735-1780.
60. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV); 2015; Santiago, Chile.
61. Saxe AM, McClelland JL, Ganguli S. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. arXiv preprint arXiv:1312.6120. 2013.

62. Bottou L. Large-scale machine learning with stochastic gradient descent. In: *Proceedings of COMPSTAT'2010*. Berlin, Germany: Springer-Verlag Berlin Heidelberg ; 2010:177-186.
63. Bottou L, Curtis FE, Nocedal J. Optimization methods for large-scale machine learning. *SIAM Rev*. 2018;60(2):223-311.
64. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014.
65. Yadava M, Kumar P, Saini R, Roy PP, Dogra DP. Analysis of EEG signals and its application to neuromarketing. *Multimed Tools Appl*. 2017;76(18):19087-19111.
66. Murugappan M, Murugappan S, Gerard C. Wireless EEG signals based neuromarketing system using fast fourier transform (FFT). Paper presented at: 2014 IEEE 10th International Colloquium on Signal Processing and its Applications; 2014; Kuala Lumpur, Malaysia.
67. Gevins A, Smith ME, Leong H, et al. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Hum Factors*. 1998;40(1):79-91.
68. Smith ME, Gevins A, Brown H, Karnik A, Du R. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Hum Factors*. 2001;43(3):366-380.
69. Perry A, Stein L, Bentin S. Motor and attentional mechanisms involved in social interaction—evidence from mu and alpha EEG suppression. *Neuroimage*. 2011;58(3):895-904.
70. Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. *J Neural Eng*. 2007;4(2):R1.
71. Pfurtscheller G, Neuper C, Flotzinger D, Pregenzer M. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr Clin Neurophysiol*. 1997;103(6):642-651.
72. Chiappa S, Bengio S. *HMM and IOHMM Modeling of EEG Rhythms for Asynchronous BCI Systems*. Research Report. Martigny, Switzerland: IDIAP; 2003.
73. Millan JR, Mouriño J. Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Trans Neural Syst Rehabil Eng*. 2003;11(2):159-161.
74. Lehmann C, Koenig T, Jelic V, et al. Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *J Neurosci Methods*. 2007;161(2):342-350.
75. Simpraga S, Alvarez-Jimenez R, Mansvelder HD, et al. EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease. *Sci Rep*. 2017;7(1):5775.
76. Goli P, Rad EM, Ghandehari K, Azarnoosh M. Early assessment of mild Alzheimer's disease using Elman neural network, LDA and SVM methods. *Mach Learn Res*. 2017;2(4):148-151.
77. Trambaiolli LR, Lorena AC, Fraga FJ, Kanda PAM, Anghinah R, Nitrini R. Improving Alzheimer's disease diagnosis with machine learning techniques. *Clin EEG Neurosci*. 2011;42(3):160-165.
78. Song Y, Crowcroft J, Zhang J. Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. *J Neurosci Methods*. 2012;210(2):132-146.
79. Tzimourta KD, Tzallas AT, Giannakeas N, Astrakas LG, Tsalikakis DG, Tsipouras MG. Epileptic seizures classification based on long-term EEG signal wavelet analysis. In: Maglaveras N, Chouvarda I, de Carvalho P, eds. *Precision Medicine Powered by pHealth and Connected Health*. Singapore: Springer Nature Singapore Pte Ltd; 2018:165-169. *IFMBE Proceedings*; vol. 66.
80. Golmohammadi M, Torbati AHHN, Lopez de Diego S, Obeid I, Picone J. Automatic analysis of EEGs using big data and hybrid deep learning architectures. arXiv preprint arXiv:1712.09771. 2017.
81. Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI. A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. *Comput Biol Med*. 2018;99:24-37.

How to cite this article: Li G, Lee CH, Jung JJ, Youn YC, Camacho D. Deep learning for EEG data analytics: A survey. *Concurrency Computat Pract Exper*. 2019;e5199. https://doi.org/10.1002/cpe.5199