Deep learning for EEG data analytics: A survey

Article *in* Concurrency and Computation Practice and Experience · February 2019

DOI: 10.1002/cpe.5199

Gen Li
Chung-Ang University

Young Chul Youn
Chung-Ang University Hospital

David Camacho
Universidad Politécnica de Madrid

Received 26 April 2016; | Revised 6 June 2016; | Accepted 6 June 2016

DOI: xxx/xxxx

ARTICLE TYPE

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

Correspondence
*Jason J. Jung, Computer Engineering, Chung-Ang University, 84 Heukseok, Seoul, Korea. Email: j2jung@gmail.com

Summary

Research on EEG signal processing have been got more focus while the EEG signal as a basis of prediction of brain behavior and diagnose disease. With the success of deep learning apply on the time series data, many studies have been start applying deep learning on the EEG signal processing. In order to summarize the technologies of EEG signal processing, we have conducted literature review about deep learning 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, we introduce 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, we discuss the challenges and limitations of EEG signal processing on the future.

KEYWORDS:
Deep learning, EEG signal processing, Clinical AI-based diagnosis

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], the EEG signals is 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 the machine learning algorithms have been demonstrated 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. [8] 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 Fig. 1 [7], it is not difficult to

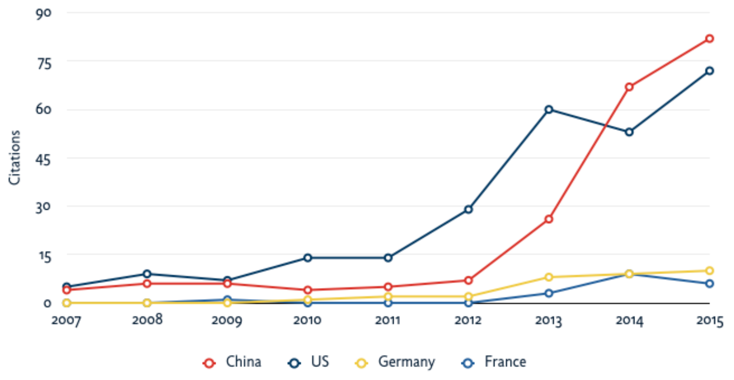

FIGURE 1 Number of Publications about deep learning for reference of countries: China, United States, Germany, France [7]

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 [10,11,12] 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.

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 use 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 Stober et al. [15] classified the rhythm and genre of music which performed to experimenter, and Cecotti et al. [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] discussed CNN for EEG to understand the behavior of human's brain. In this paper, we will analyze the basics on deep neural network 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:

(i) To understand the current research trends on deep learning for EEG signal processing and analytic by conducting literature review. The different available architectures and applications will revised.

(ii) How non-fixed EEG data can be analyzed by deep learning. It will be described a new model, build by combining CNN and LSTM models, developed by authors [58] and it will by discussed bow this model could be employed over other time-serial data.

(iii) Considering the application of EEG analysis and EEG signal processing by review some studies about the most popular applications like Neuromarketing, human factors, social interaction and BCI. We also survey 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 the 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 intial 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 rised their relevance when they were applied to specific application problems, such as the brain-computer

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, 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, meditation
Lambda	Evoked potential	When the eyes are stimulated by light, they are induced after 100ms (also known as P100)
P300	Evoked potential	When you see or hear what is imagined in your brain, 300ms induces it.

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.

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 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 1-30 times per second, it can be divided into four bands, δ (1-4Hz),θ (4-7Hz),α (8-15Hz),β (12.5-28Hz). 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 30-80Hz 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. The Table 1 shows the Band, Frecuencies and some other details related to the EEG signal and what it could means 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 frecuencies 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 better on data analytics in the 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 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. XW Wang et al. [22] 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 (i) that the spectrum feature is superior than other two features, (ii)using the feature smoothing method to improve a linear dynamic system could improve the accuracy of classification, (iii) the changes of emotion could be monitored by reducing the subject-independent features with manifold learning. M Kaper et al. [23] 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. AH shoeb et al. [24] used machine learning method for seizure detection.

TABLE 2 Summarize of deep learning for different EEG type

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

2.3 | Deep neural networks for EEG data

Since mixing with noise and artifacts in recorded EEG signals, as used features are more frequently than raw signals. In brain decoding, Jia et al.[25] and Jirayucharoensak et al.[26] used DBN and SAE to model a EEG signals classification for emotion detection. Zheng W-L et al.[27] presented the critical channels and frequency bands which related with emotion recognition by decoding EEG signals with DBN. An et al.[28] classify left- and right-hand motor imagery skills by analysis the frequency factor of EEG which applied DBN. In the other classification of deep neural networks for EEG[29,30,31,32]. Wulsin D et al.[29] used semi-supervised deep belief nets to make a fast classification and detection anomaly measurement, and Langkvist M et al.[32] used unsupervised feature learning to classify the human's sleep stage. M Hajinoroozi et al.[33] used DBN to extract the feature from EEG signals for cognitive the driver's states.

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

There we can see, most studies prefer doing preprocessing with transformation of raw data to frequency bands before using deep neural network training EEG data, that's because the frequency bands which be filtered has been more clean, it reduce the influence of noisy.

2.4 | Deep recurrent networks for EEG data

EEG signals also are 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 play a important role. For training EEG data, RNN could better to extract the feature informations from EEG data. In 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.

A Petrosian et al.[34] predicted seizures by applied RNN to raw EEG data and corresponding wavelet features.Davidson et al.[35] transformed the EEG data to spectra features and used LSTM to detect lapses. GR Minasyan et al.[36] developed a method for detection of seizures prior to or immediately after clinical onset using features derived from scalp EEG data. MA Naderi et al.[37] proposed a three stages technique for seizure detection in EEG signals, first used Welch method power spectrum density estimation to extracted 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. M Soleymani et al.[38] EM Fomey et al.[39] proposed the Elman Recurrent neural networks a new architecture of RNN to classifier EEG signals during imagined mental tasks. M Li et al. recognize MI-EEG combined with LSTM based network employing. S Patnaik et al.[40] used wavelet transform to deal with extraction of EEG features for predict the human's brain state. Z Ni et al.[41] 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.

Since the EEG data recorded the value of each temporary, the information of each times is very important for diagnose diseases, such as P300 signal will be recorder before the 300 ms of seizure prediction, and recurrent networks model has a better performance to training sequential data since RNN could save the information of previous time to help the next time to make prediction, therefore, RNN model usually used for diagnosing disease.

2.5 | Deep convolutional networks for EEG data

Convolutional neural networks has a good performance on image processing tasks recently, since it have a good capacity on feature extraction from images by convolution kernel, they can extract the information features by through passing every part of the images with number of kernels. For EEG signals, it is not only works well on raw EEG data, but also used on frequency spectrum diagram. There we categorized the researches in EEG signals processing into two groups: 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, that's can reduce the complexity of training, therefore most studies used CNN for EEG signals. For instance Cecotti et al.[16] classified characters that viewed by participants. Z Tang et al.[17] presented a new approach to extract the feature and classify the single-trial MI EEG. X Sun et al.[19] proposed a computational method to detect memory performance of remembered or forgotten by training EEG data while memory processing. Thodoroff et al.[42] combined of CNN and RNN to train robust features to automatically detect seizures. J Shamwell et al.[43] explored a new CNN architecture with 4 convolution layers and 3 full connect layers to generalized multi-class, single-trial EEG classification across subjects, aim to increase human-autonomy classification performance. Manor et al.[44] 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 compared the feature extraction by CNN and manually designing feature extraction algorithms. S Sakhavi et al.[45] presented a Parallel convulutional-linear network which an architecture that can make EEG data as a dynamic energy representation for input and utilizes CNN for imagery classification. Y Ren et al.[46] applied convolutional deep belief networks learning features from EEG data and evaluated it based on compared with the datasets from BCI. G Ruffini et al.[47] collected data from idiopathic RBD patients and healthy controls and proposed CNN for classifying rapid eye movement behavior disorder prognosis. M Hajinoroozi et al.[48] used covariance learning to train EEG data for driver's fatigue prediction. Z Jiao et al.[49] proposed improved CNNs methods for mental load classification task.

2.5.2 | Diagnose diseases

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

Mirowski et al.[50] extracted features as phase-locking synchrony and wavelet coherence and coded them as pixel colors to formulate two-dimensional patterns[2]. Liang et al.[51] adopted EEG datasets which are not directly related to seizure prediction and training by deep learning to detected seizure. Antoniades et al.[52] consider generating feature automatically from epileptic intracranial EEG data in time domain by deep leaning. Page A et al.[53] made a 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. UR Acharya et al.[54] presented a deep convolutional neural networks with 5 layers to detected normal, preictal and seizure classes.

Thus it can be seen, the development of deep learning for EEG signals processing are increasing, and the researches 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's the reason why many studies begin focus on deep learning for EEG signals processing. There we make a Table 4 to list some details of CNN architectures, decoding problems and input domain by survey from part of papers.

As reviewed as before, each architecture have this own characters for training EEG data. Since the EEG data is sequential data, thus the training by using deep neural network will spend more time, and the other reason that raw EEG data usually be recorded with noisy, deep neural work can not appear a good performance for EEG data processing. While RNN model especially LSTM model can leverage the information of previous time, that they can consider the influence of EEG data for prediction on time domain. More than this, while extract the EEG data on frequency domain, RNN model also has a good performance since the EEG data of frequency domain doesn't mix up many noisy. However, there are many weighted parameters in the RNN model, thus the training time of RNN model will be increased. Because of the better ability of local feature extraction, CNN model most used for sequential data processing such as speech recognition, translation and so on. Thus, more and more studies used CNN model to train EEG data, since local feature of EEG data advantageous to realize the statement of brain activity, and to determine the brain waveforms at the disease time.

TABLE 3 Survey papers of CNN for EEG signals

Study	Decoding problem	Input domain	Conv/dense layers
[13]	Type of music rhythm	Time 0.5-39Hz	3/3
[15]	Imaged movement classes	Time,8-30Hz	2/2
[17]	Memory performance	Time,0.05-15Hz	2/2
[26]	Dirver performance	Time, 1-50Hz	1/3
[35]	Start of epileptic seizure	Frequency, mean amplitude for 0-7Hz,7-14Hz,14-49Hz	3/1(+LSTM)
[36]	Oddball response	Time, 05-50Hz	4/3
[37]	Oddball response using RSVP and image	Time,0.3-20Hz	3/2
[38]	Imagined movement classes	Frequency 4-40Hz	2/2
[39]	Imagined movement classes	Frequency 8-30Hz	2/0
[40]	Eye movement classes	Stacked multi-channel spectrograms	4/1
[41]	Driver performance	Frequency 256Hz	2/2
[44]	Seizure prediction	Frequency, 0-200Hz	1/2
[45]	Epileptic discharge	Time, 1-50Hz	1/3
[46]	Seizure detection	Time, 0-128Hz	1-3/1-3
[47]	Seizure detection	Frequency, 4-7Hz,8-13Hz,13-30Hz	5/1

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 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 method like Self training, Generative model, S3VMs and so on. Y Li et al.[55]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. D F Wulsin et al.[56] applied DBNs in a semi-supervised learning to model EEG data for classification and prediction. LC shi et al.[57] used semi-supervised clustering method to analysis vigilance based on EEG data. D Wulsin[56] 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 P Bashiva et al.[58]

3.1 | Convolutional neural networks

On account of a greatly ability on feature extraction from multidimensional dataset, CNN has been achieved great success in many recognition issue like image recognition or behavior recognition through by extract information from photo. And 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 been become a 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, the CNN model for EEG data shown as Figure 2 . Be different with CNN, for RNN, one of the neuron output value not only pass to the next neuron, but also act on itself, that means, the output of the current moment has been combined the experience of this time and the history, that's the reason why RNN has a greatly success for sequential data. But for CNN, the essence of convolution operation is use 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 informations are useful for nature language processing. Thus, many researchers has been began to use CNN to analysis 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

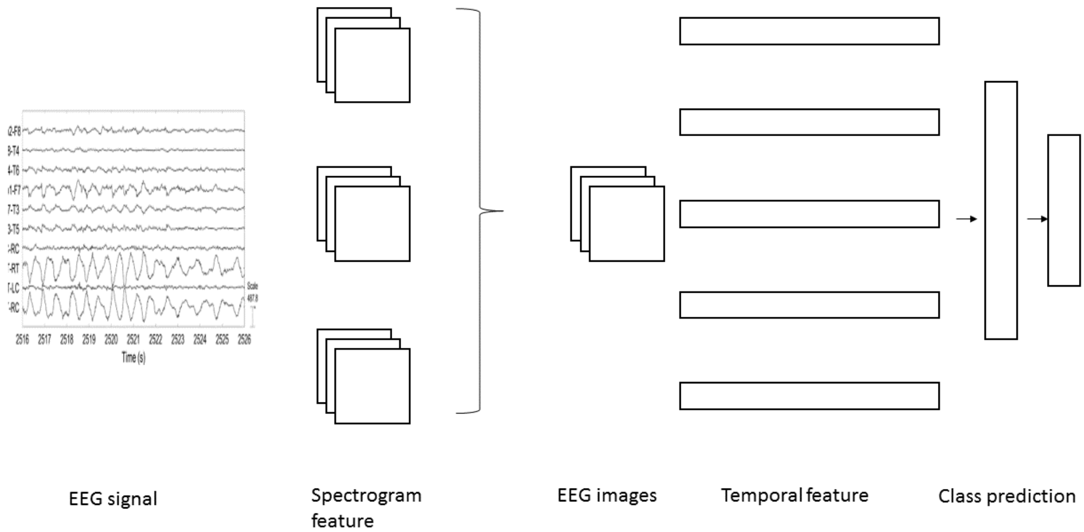

FIGURE 2 CNN for EEG signal processing

believe that we can decoding 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 field, but these always have some defects which consists none of the above models can analyze the overall logical sequence between the input information. These logical sequences are rich in content and has a complex time relationship with each other. In order to solve these sequence problems, recurrent neural networks arises 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 the other word, after pass the value to the next hidden layer, it always give this value to itself, it make this network has 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 equal a function which combined 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, JÃijrgen Schmidhuber et al.[59] proposed a new architecture of RNN named Long Short-Term Memory (LSTM), Be 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 a new data 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 to extract the useful information from varied length EEG data and classify to correct class. The input of this task is varied length signal $X = [x_1, x_2, ...x_k]$, and output 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.

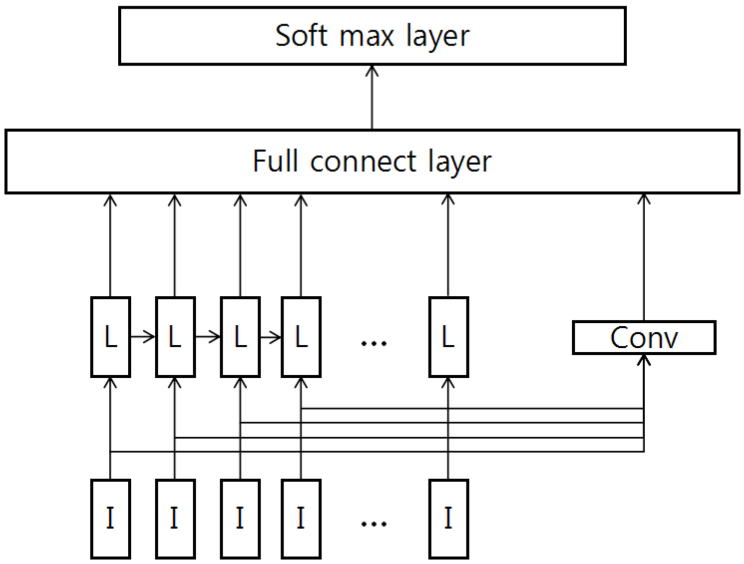

FIGURE 3 Connection between Conv layer and full connect layer. C: ConvNet, L: LSTM layer

Bashiva et al.[58] presented a model by combing LSTM and CNN for learning representations from EEG data, the model architecture shown as Figure 3 , they used 3 convolution layers to extracted the spectrogram features and got one EEG image sets, then pooling to got temporal features, then training these by LSTM. Connected with the last pooling layer is LSTM model proposed by Jrgen Schmidhuber et al.[59] The LSTM model include some cells, these like units of deep neural networks with some parameters, therefore, while a new data inputing, the LSTM model will forget no useful informations, then learning valuable informations from new input data and save on long memory, that's the reason why the LSTM model can learning series time data better. The cells of LSTM behavior is concluded 4 equations:

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

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

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

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

where $\mathrm{h_t}, \mathrm{x_t}, \mathrm{C} - \mathrm{t}$ are the output, input and private state at time t of cell. The $\mathrm{W_f}, \mathrm{W_c}, \mathrm{W_0}$ 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, 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 initializer[60] and for the LSTM cells the orthogonal initializer[61] could be used. Because they are shown a greatly 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.

(i) Intuitive motivation.

(ii) Practical motivation.

(iii) 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 list some advantages of Adam:

(i) A straightforward implementation.

(ii) The efficient calculation.

(iii) Less memory required.

(iv) Invariance of gradient diagonal scaling.

(v) Suitable for solving optimization problems with large data and parameters.

(vi) Applicable to non-stationary targets.

(vii) Suitable for solving problems involving very high noise or sparse gradient.

(viii) Super parameters can be interpreted intuitively and basically require only a very small number of parameters.

Therefore Adam sometimes will be a better choose. 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) is most popular research for analysis EEG data to build a direct connection between human brain and external devices, it allow 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 including the driving consumer decisions, brain areas that are active when consumer purchase a product or service and so on. M yadave et al. [65] 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. M Murugappan et al. [66] 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 always used for identifying the brain processes related to specific personality traits such as intro-/extroversion or social anxiety. A Gevins et al. [67] 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. ME Smith et al. [68] 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 are investigated. Thus EEG data analytics on social interaction is play a important role and many studies has research the social interaction based on EEG data analytics. In A Perry et al. [69]'s paper, Mu rhythms are EEG oscillations in 8-13 Hz recorded at sites located roughly over the sensory-motor cortex, this paper found that Mu suppression not only in response to actual activities but also when the participant observes actions executed by other peoples.

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, help paralyzed patients steer their wheelchairs or move a cursor on a screen. F Lotte et al.[70] surveyed some papers which 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]. And 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 EEG is the necessary basis for the diagnosis of epilepsy and AD. EEG also have 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 mostly studies focus on EEG processing by deep learning.

4.2.1 | Clinical application of Alzheimer's disease

Christoph L et al.[74] explored various machine learning algorithms with the ability of linear and non-linear such as principal component linear discriminant analysis(PC LDA), partial leaset 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. S Simpraga et al.[75] used complementary biomaker algorithms for EEG recording to get the signature of disease and pharmacological intervention and then in order to improve the performance of distinguish, they used machine learning algorithms making a classification. P Goli et al.[76] used Elman neural networks to drawn out the optimal features and used linear discriminant analysis (LDA) and SVM to generate 2 classifications for mild Alzheimer's disease. Y zhao et al.[31] 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. LR Trambaiolli et al.[77] 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

AH Shoeb et al.[24] used machine learning to construct patient classifier which predict the onset of epileptic seizure by EEG analyzing. This studies point at the problem which is brain's electrical activity is mix with numerous classes with characteristics, they presented a new algorithm which include make a machine learning framework and identifying the useful features from other types of brain activity. Y Song et al.[78] 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. But the proposed method in this paper did not achieved a high detection accuracy but also have a high speed of computation. KD Tzimourta et al.[79] presented a method for automated seizure which based on detection discrete wavelet transform. This paper extract five features from wavelet coefficients and training this features by SVM, for the result they get a high accuracy about the seizure detection. M 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, the third is generalized periodic epileptiform discharges. KM 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 expanding 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 play a center stage for international academic and lots of interests. In EEG signals, where having a good advance with deep learning to produce the expecting results. In this paper, we introduced the some architecture of popular deep learning, type of EEG data which as 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 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 research applications for EEG data analytics, that including the neuromarketing, human factors, social interaction and BCI. For the research of BCI, we discussed

TABLE 4 Survey papers of EEG analysis on Clinical application

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

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 popular issues of recent like image captioning, video summarization and image question answering are combined with CNN and RNN to applied. And in this paper we already seen related papers which combined these two architectures to learning representations. Reviewed these studies, we can consider multiple model combined to denosing and doing preprocessing for increase the performance of combined model will be the trend of future. And most of these papers which we saw were used supervised learning, in terms of EEG data processing, semi-supervised and reinforcement learning are 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 require the smallest number of personnel to do the work, and at the same time, it can bring higher accuracy. Reinforcement learning also play a important role in the artificial intelligence, it can adjust itself by reward function for completing the optimization if its own scheme. In the EEG signals, reinforcement can produces 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 a important on BCI research in future.

For the limitations of EEG data, since two reasons, labeled data is more rarely. First is the doctor label one sample will spend lots time, they need to combine the variety influence factors such as gender, ethnic group and constitutions. Second, because of the influence from the noisy, they may make a error prediction to tag a label. Therefore, the challenges of EEG signal processing is get a high accuracy with less label. Although the semi-supervised learning could reach this task, however the most component of semi-supervised leaning is unsupervised learning, thus the method of increasing the performance of semi-supervised learning will be a new challenge not only EEG signal processing but also most of fields of deep learning. The generative adversarial networks has been get pretty success on the image processing recently, applying GAN model to the EEG signal for learning loss signal or prejudging illness by generating the not appeared EEG signal will be the new challenge.

Although deep learning has been used in many fields and got lots achievements, it is not always provide great results in EEG signals processing. The main reasons 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 obviously, 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 results 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.

ACKNOWLEDGEMENT

This research was supported by the Chung-Ang University Research Grants in 2017.

References

1. Manyika James, Chui Michael, Brown Brad, et al. *Big data: The next frontier for innovation, competition, and productivity.* : McKinsey & Company; 2011.

2. Min Seonwoo, Lee Byunghan, Yoon Sungroh. Deep learning in bioinformatics. *Briefings in bioinformatics.* 2017;18(5):851–869.

3. Zhao Gang, Liu Shan, Wang Qi, et al. Deep convolutional neural network for drowsy student state detection. *Concurrency and Computation: Practice and Experience.* ;:e4457.

4. Liu Jung-Chun, Leu Fang-Yie, Lin Guan-Liang, Susanto Heru. An MFCC-based text-independent speaker identification system for access control. *Concurrency and Computation: Practice and Experience.* 2018;30(2):e4255.

5. Ferrucci David, Brown Eric, Chu-Carroll Jennifer, et al. Building Watson: An overview of the DeepQA project. *AI magazine.* 2010;31(3):59–79.

6. Oncology. IBM Watson. IBM https://console.bluemix.net/catalog/services/health-score.

7. Etten. Jaap. Datenna http://www.datenna.com/industry/china-marching-forward-with-artificial-intelligence.

8. Hinton Geoffrey E, Osindero Simon, Teh Yee-Whye. A fast learning algorithm for deep belief nets. *Neural computation.* 2006;18(7):1527–1554.

9. LeCun Yann, Bengio Yoshua, Hinton Geoffrey. Deep learning. *nature.* 2015;521(7553):436.

10. Farabet Clement, Couprie Camille, Najman Laurent, LeCun Yann. Learning hierarchical features for scene labeling. *IEEE transactions on pattern analysis and machine intelligence.* 2013;35(8):1915–1929.

11. Szegedy Christian, Liu Wei, Jia Yangqing, et al. Going deeper with convolutions. In: Cvpr; 2015.

12. Chorowski Jan K, Bahdanau Dzmitry, Serdyuk Dmitriy, Cho Kyunghyun, Bengio Yoshua. Attention-based models for speech recognition. In: :577–585; 2015.

13. Luong Minh-Thang, Pham Hieu, Manning Christopher D. Effective approaches to attention-based neural machine translation 2015.

14. Guan Ziyu, Chen Long, Zhao Wei, Zheng Yi, Tan Shulong, Cai Deng. Weakly-Supervised Deep Learning for Customer Review Sentiment Classification.. In: :3719–3725; 2016.

15. Stober Sebastian, Cameron Daniel J, Grahn Jessica A. Using Convolutional Neural Networks to Recognize Rhythm Stimuli from Electroencephalography Recordings. In: :1449–1457; 2014.

16. Cecotti Hubert, Graser Axel. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE transactions on pattern analysis and machine intelligence.* 2011;33(3):433–445.

17. Tang Zhichuan, Li Chao, Sun Shouqian. Single-trial EEG classification of motor imagery using deep convolutional neural networks. *Optik-International Journal for Light and Electron Optics.* 2017;130:11–18.

18. Lawhern Vernon J, Solon Amelia J, Waytowich Nicholas R, Gordon Stephen M, Hung Chou P, Lance Brent J. Eegnet: A compact convolutional network for eeg-based brain-computer interfaces. *arXiv preprint arXiv:1611.08024.* 2016;.

19. Sun Xuyun, Qian Cunle, Chen Zhongqin, Wu Zhaohui, Luo Benyan, Pan Gang. Remembered or Forgotten?âĂŤAn EEG-Based Computational Prediction Approach. *PloS one.* 2016;11(12):e0167497.

20. Chang Jie, Gu Yujun, Zhang Ya. Chinese Typeface Transformation with Hierarchical Adversarial Network. *arXiv preprint arXiv:1711.06448.* 2017;.

21. Goodfellow Ian, Bengio Yoshua, Courville Aaron, Bengio Yoshua. *Deep learning.* MIT press Cambridge; 2016.

22. Wang Xiao-Wei, Nie Dan, Lu Bao-Liang. Emotional state classification from EEG data using machine learning approach. *Neurocomputing.* 2014;129:94–106.

23. Müller Klaus-Robert, Tangermann Michael, Dornhege Guido, Krauledat Matthias, Curio Gabriel, Blankertz Benjamin. Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. *Journal of neuroscience methods.* 2008;167(1):82–90.

24. Shoeb Ali H, Guttag John V. Application of machine learning to epileptic seizure detection. In: :975–982; 2010.

25. Jia Xiaowei, Li Kang, Li Xiaoyi, Zhang Aidong. A novel semi-supervised deep learning framework for affective state recognition on eeg signals. In: :30–37IEEE; 2014.

26. Jirayucharoensak Suwicha, Pan-Ngum Setha, Israsena Pasin. EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World Journal.* 2014;2014.

27. Zheng Wei-Long, Guo Hao-Tian, Lu Bao-Liang. Revealing critical channels and frequency bands for emotion recognition from EEG with deep belief network. In: :154–157IEEE; 2015.

28. An Xiu, Kuang Deping, Guo Xiaojiao, Zhao Yilu, He Lianghua. A deep learning method for classification of EEG data based on motor imagery. In: :203–210Springer; 2014.

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. *Journal of neural engineering.* 2011;8(3):036015.

30. Turner JT, Page Adam, Mohsenin Tinoosh, Oates Tim. Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection. In: ; 2014.

31. Zhao Yilu, He Lianghua. Deep learning in the EEG diagnosis of AlzheimerâĂŹs disease. In: :340–353Springer; 2014.

32. Längkvist Martin, Karlsson Lars, Loutfi Amy. Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems.* 2012;2012.

33. Hajinoroozi Mehdi, Jung Tzyy-Ping, Lin Chin-Teng, Huang Yufei. Feature extraction with deep belief networks for driver's cognitive states prediction from EEG data. In: :812–815IEEE; 2015.

34. Petrosian Arthur, Prokhorov Danil, Homan Richard, Dasheiff Richard, Wunsch II Donald. Recurrent neural network based prediction of epileptic seizures in intra-and extracranial EEG. *Neurocomputing.* 2000;30(1-4):201–218.

35. Davidson Paul R, Jones Richard D, Peiris Malik TR. EEG-based lapse detection with high temporal resolution. *IEEE Transactions on Biomedical Engineering.* 2007;54(5):832–839.

36. Minasyan Georgiy R, Chatten John B, Chatten Martha Jane, Harner Richard N. Patient-specific early seizure detection from scalp EEG. *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society.* 2010;27(3):163.

37. Naderi Mohammad Ali, Mahdavi-Nasab Homayoun. Analysis and classification of EEG signals using spectral analysis and recurrent neural networks. In: :1–4IEEE; 2010.

38. Soleymani Mohammad, Asghari-Esfeden Sadjad, Pantic Maja, Fu Yun. Continuous emotion detection using EEG signals and facial expressions. In: :1–6IEEE; 2014.

39. Forney Elliott M, Anderson Charles W. Classification of EEG during imagined mental tasks by forecasting with Elman recurrent neural networks. In: :2749–2755IEEE; 2011.

40. Patnaik Suprava, Moharkar Lalita, Chaudhari Amogh. Deep RNN learning for EEG based functional brain state inference. In: :1–6IEEE; 2017.

41. Ni Zhaoheng, Yuksel Ahmet Cem, Ni Xiuyan, Mandel Michael I, Xie Lei. Confused or not Confused?: Disentangling Brain Activity from EEG Data Using Bidirectional LSTM Recurrent Neural Networks. In: :241–246ACM; 2017.

42. Thodoroff Pierre, Pineau Joelle, Lim Andrew. Learning robust features using deep learning for automatic seizure detection. In: :178–190; 2016.

43. Shamwell Jared, Lee Hyungtae, Kwon Heesung, Marathe Amar R, Lawhern Vernon, Nothwang William. Single-trial EEG RSVP classification using convolutional neural networks. In: :983622International Society for Optics and Photonics; 2016.

44. Manor Ran, Geva Amir B. Convolutional neural network for multi-category rapid serial visual presentation bci. *Frontiers in computational neuroscience.* 2015;9:146.

45. Sakhavi Siavash, Guan Cuntai, Yan Shuicheng. Parallel convolutional-linear neural network for motor imagery classification. In: :2736–2740IEEE; 2015.

46. Ren Yuanfang, Wu Yan. Convolutional deep belief networks for feature extraction of EEG signal. In: :2850–2853IEEE; 2014.

47. Ruffini Giulio, Soria David Ibanez, Dubreuil Laura, Gagnon Jean-Francois, Montplaisir Jacques, Soria-Frisch Aureli. Deep learning using EEG spectrograms for prognosis in idiopathic rapid eye movement behavior disorder (RBD). *bioRxiv.* 2018;:240267.

48. Hajinoroozi Mehdi, Zhang Jianqiu Michelle, Huang Yufei. Driver's fatigue prediction by deep covariance learning from EEG. In: :240–245IEEE; 2017.

49. Jiao Zhicheng, Gao Xinbo, Wang Ying, Li Jie, Xu Haojun. Deep Convolutional Neural Networks for mental load classification based on EEG data. *Pattern Recognition.* 2018;76:582–595.

50. Mirowski Piotr, Madhavan Deepak, LeCun Yann, Kuzniecky Ruben. Classification of patterns of EEG synchronization for seizure prediction. *Clinical neurophysiology.* 2009;120(11):1927–1940.

51. Liang Jian, Lu Rui, Zhang Changshui, Wang Fei. Predicting Seizures from Electroencephalography Recordings: A Knowledge Transfer Strategy. In: :184–191IEEE; 2016.

52. Antoniades Andreas, Spyrou Loukianos, Took Clive Cheong, Sanei Saeid. Deep learning for epileptic intracranial EEG data. In: :1–6IEEE; 2016.

53. Page Adam, Shea Colin, Mohsenin Tinoosh. Wearable seizure detection using convolutional neural networks with transfer learning. In: :1086–1089IEEE; 2016.

54. Acharya U Rajendra, Oh Shu Lih, Hagiwara Yuki, Tan Jen Hong, Adeli Hojjat. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in biology and medicine.* 2017;.

55. Li Yuanqing, Guan Cuntai, Li Huiqi, Chin Zhengyang. A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. *Pattern Recognition Letters.* 2008;29(9):1285–1294.

56. Wulsin Drausin, Blanco Justin, Mani Ram, Litt Brian. Semi-supervised anomaly detection for EEG waveforms using deep belief nets. In: :436–441IEEE; 2010.

57. Shi Li-Chen, Yu Hong, Lu Bao-Liang. Semi-supervised clustering for vigilance analysis based on EEG. In: :1518–1523IEEE; 2007.

58. Bashivan Pouya, Rish Irina, Yeasin Mohammed, Codella Noel. Learning representations from EEG with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448.* 2015;.

59. Hochreiter Sepp, Schmidhuber Jürgen. Long short-term memory. *Neural computation.* 1997;9(8):1735–1780.

60. He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: :1026–1034; 2015.

61. Saxe Andrew M, McClelland James L, Ganguli Surya. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv preprint arXiv:1312.6120.* 2013;.

62. Bottou Léon. Large-scale machine learning with stochastic gradient descent. In: Springer 2010 (pp. 177–186).

63. Bottou Léon, Curtis Frank E, Nocedal Jorge. Optimization methods for large-scale machine learning. *SIAM Review.* 2018;60(2):223–311.

64. Kingma Diederik P, Ba Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.* 2014;.

65. Yadava Mahendra, Kumar Pradeep, Saini Rajkumar, Roy Partha Pratim, Dogra Debi Prosad. Analysis of EEG signals and its application to neuromarketing. *Multimedia Tools and Applications.* 2017;76(18):19087–19111.

66. Murugappan M, Murugappan Subbulakshmi, Gerard Celestin, others . Wireless EEG signals based neuromarketing system using Fast Fourier Transform (FFT). In: :25–30IEEE; 2014.

67. Gevins Alan, Smith Michael E, Leong Harrison, et al. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human factors.* 1998;40(1):79–91.

68. Smith Michael E, Gevins Alan, Brown Halle, Karnik Arati, Du Robert. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors.* 2001;43(3):366–380.

69. Perry Anat, Stein Libi, Bentin Shlomo. Motor and attentional mechanisms involved in social interactionâĂŤEvidence from mu and alpha EEG suppression. *Neuroimage.* 2011;58(3):895–904.

70. Lotte Fabien, Congedo Marco, Lécuyer Anatole, Lamarche Fabrice, Arnaldi Bruno. A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of neural engineering.* 2007;4(2):R1.

71. Pfurtscheller Gert, Neuper Ch, Flotzinger Doris, Pregenzer Martin. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and clinical Neurophysiology.* 1997;103(6):642–651.

72. Chiappa Silvia, Bengio Samy. *HMM and IOHMM modeling of EEG rhythms for asynchronous BCI systems.* : IDIAP; 2003.

73. Millan Jdel R, Mouriño Josep. Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Transactions on Neural Systems and Rehabilitation Engineering.* 2003;11(2):159–161.

74. Lehmann Christoph, Koenig Thomas, Jelic Vesna, et al. Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *Journal of neuroscience methods.* 2007;161(2):342–350.

75. Simpraga Sonja, Alvarez-Jimenez Ricardo, Mansvelder Huibert D, et al. EEG machine learning for accurate detection of cholinergic intervention and AlzheimerâĂŹs disease. *Scientific reports.* 2017;7(1):5775.

76. Goli Peyman, Rad Elias Mazrooei, Ghandehari Kavian, Azarnoosh Mehdi. Early Assessment of Mild AlzheimerâĂŹs Disease Using Elman Neural Network, LDA and SVM Methods. *Machine Learning Research.* 2017;2(4):148.

77. Trambaiolli Lucas R, Lorena Ana C, Fraga Francisco J, Kanda Paulo AM, Anghinah Renato, Nitrini Ricardo. Improving Alzheimer's disease diagnosis with machine learning techniques. *Clinical EEG and neuroscience.* 2011;42(3):160–165.

78. Song Yuedong, Crowcroft Jon, Zhang Jiaxiang. Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. *Journal of neuroscience 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: Springer 2018 (pp. 165–169).

80. Golmohammadi Meysam, Torbati Amir Hossein Harati Nejad, Diego Silvia Lopez, Obeid Iyad, Picone Joseph. Automatic Analysis of EEGs Using Big Data and Hybrid Deep Learning Architectures. *arXiv preprint arXiv:1712.09771.* 2017;.

81. Tsiouris Kostas M, Pezoulas Vasileios C, Zervakis Michalis, Konitsiotis Spiros, Koutsouris Dimitrios D, Fotiadis Dimitrios I. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. *Computers in biology and medicine.* 2018;99:24–37.

AUTHOR BIOGRAPHY

empty.pdf

Author Name. This is sample author biography text this is sample author biography text.

How to cite this article: