# Overview of Deep Learning Architectures for Classifying Brain Signals

Lachezar Bozhkov[1]

*Abstract* **One of the challenges in modeling cognitive events from electroencephalogram (EEG) data is finding representations that are invariant to inter- and intra-subject differences, as well as the inherent noise associated with EEG data collection. Herein, we explore the capabilities of the recent deep neural architectures for modeling cognitive events from EEG data. In this paper, we present recent achievements applying deep learning for EEG signal classification. We investigate the use of feed forward, convolutional, recurrent neural nets, as well as deep belief networks, echo-state networks, reservoir computing, and denoising auto encoder models. We present the application of these architectures for classifying user intent generated through different motor imagery; BCI to control wheelchair and robotic arm; mental load classification; discriminating emotional state; feature dimensionality reduction for EEG data. Many of the models prove to be more accurate and more efficient than current state-of-the-art models.**

*Index Terms:* **EEG data, Deep Learning, Brain-Computer Interface, Convolutional Neural Networks, Auto-Encoders, Recurrent Neural Networks, Echo State Networks, Deep Belief Networks, Reservoir Computing, Affective Computing**

## I. Introduction

Over the last decade, deep learning techniques have become very popular in various application domains such as computer vision, automatic speech recognition, natural language processing, and bioinformatics where they produce state-of-the-art results on various tasks. At the same time, there has been very little progress investigating the application of deep learning in cognitive neuroscience research, where these techniques could be used to analyze signals recorded with electroencephalography (EEG) – a non-invasive brain imaging technique that relies on electrodes placed on the scalp to measure the electrical activity of the brain. EEG is especially popular for the development of brain-computer interfaces (BCIs), which work by identifying different brain states from the EEG signal.

Working with EEG data poses several challenges. Brain waves recorded in the EEG have a very low signal-to-noise ratio and the noise can come from a variety of sources. For instance, the sensitive recording equipment can easily pick up electrical line noise from the surroundings. Other unwanted electrical noise can come from muscle activity, eye movements, or blinks.

Usually, only certain brain activity is of interest, and this signal needs to be separated from background processes. EEG lacks spatial resolution on the scalp with additional spatial smearing caused by the skull but it has a good (millisecond) time resolution to record both, slowly and rapidly changing dynamics of brain activity.

Hence, in order to identify the relevant portion of the signal, sophisticated analysis techniques are required that should also take into account temporal information.

## II. Brain Computer Interface and Electroencephalography

Electroencephalography (EEG) is a method to analyse brain activity by measuring the electrical activity across a subject's scalp. EEG offers many advantages for construction of a BCI system, but also several disadvantages. Firstly, and most importantly, EEG is a non-invasive method for measuring brain activity. This removes the need for costly and risky surgical procedures, such as electrophysiology, in which intracortical devices such as needles or tubes may be inserted directly into the brain material, or electrocorticography, in which an array of electrodes is implanted under the skull. Both systems risk permanent and life threatening damage to a patient's brain and require costly surgical expertise to carry them out safely. Also useful for designing a BCI, EEG does not require the patient to be stationary like other noninvasive imaging systems such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG), both of which can only be carried out by large scale and expensive equipment. In contrast, EEG simply requires the placement of a set of electrodes along the scalp, which, although the exact placement is important for valid results, can be carried out in a straightforward manner. EEG has the ability to produce high time resolution data, which is a necessity for near real time systems. However, EEG does possess some major drawbacks.

A Brain Computer Interface (BCI), or Brain Machine Interface (BMI), is a system, which allows users to communicate or manipulate external devices using only their brain signals as opposed to the standard methods for carrying out such tasks (Mak & Wolpaw, 2009).

BCIs allowing patients to accomplish tasks simply by thinking about a certain action has many diverse applications that provide real benefits to their lives. Recent research has shown that much is possible with this technology. BCIs have allowed individuals to control a robotic arm to manipulate the surrounding environment (McFarland & Wolpaw, 2008). They have allowed patients to move around in a wheelchair. They have even been shown to increase the neuroplasticity of recovering tetraplegic patients and improve their quality of life. BCIs

[1] Lachezar Bozhkov is with Technical University of Sofia, Bulgaria, Computer Systems Department, 8 St.Kliment Ohridski Boulevard,Sofia 1756

which allow for three dimensional cursor motion have also been constructed. This suggests there are many more potential applications of the technology to entertainment and real-world manipulation more broadly. (McFarland & Wolpaw, 2008) provide an excellent survey of recent work in the area of BCIs as well as the basic definitions used throughout the field. They describe a BCI as carrying out the following four procedures. First, it must have a method for signal acquisition, a way of measuring the neurophysiological state of the brain. This can be accomplished by recording electrophysiological signals, via devices like the EEG or intracortical implants discussed below. Second, it must have a process for feature extraction, whereby the useful information gleaned from the signals, removing artifacts or noise from the analysis. Third, it then uses a translation algorithm to convert the extracted features into a signal of user intent. Finally, the system executes the user's desired output. Most research into BCIs has focused on the first three steps with the relevant question of interest being how to cheaply, safely, and accurately translate a user's brain activity into real-time actions.

## III. Deep Learning

Deep neural networks have recently achieved great success in recognition tasks within a wide range of applications including images, videos, speech, and text (Krizhevsky, et al., 2012) (Graves, et al., 2013) (Karpathy & Toderici, 2014) (Zhang & LeCun, 2015) (Hermann, et al., 2015). Convolutional neural networks (ConvNets) (see Fig.1) lie at the core of best current architectures working with images and video data, primarily due to their ability to extract representations that are robust to partial translation and deformation of input patterns. On the other hand, recurrent neural networks have delivered state-of-the-art performance in many applications involving dynamics in temporal sequences, such as, for example, handwriting and speech recognition (Graves, et al., 2013). In addition, combination of these two network types have recently been used for video classification (Ng, et al., 2015).

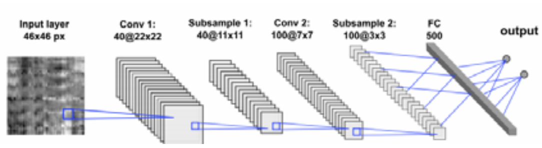


Fig.1 Convolutional Neural Network

Despite numerous successful applications of deep neural networks to large-scale image, video and text data, they remain relatively unexplored in neuroimaging domain. Perhaps one of the main reasons here is that the number of samples in most neuroimaging datasets is limited, thus making such data less adequate for training large-scale networks with millions of parameters. As it is often demonstrated, the advantages of deep neural networks over traditional machine-learning techniques become more apparent when the dataset size becomes very large. Nevertheless, deep belief network and ConvNets have been used to learn representations from functional Magnetic Resonance Imaging (fMRI) and Electroencephalogram (EEG) in some previous work with moderate dataset sizes (Plis, et al., 2014). They showed that adding several Restricted Boltzman Machine layers to a deep belief network and using supervised pretraining results in networks that can learn increasingly complex representations of the data and achieve considerable accuracy increase as compared to other classifiers. In other works, convolutional and recurrent neural networks have been used to extract representations from EEG time series (Mirowski, et al., 2009). These studies demonstrated potential benefits of adopting (down-scaled) deep neural networks in neuroimaging, even in the absence of extremely large, million-sample datasets, such as those available for images, video, and text modalities. However, none of these studies attempted to jointly preserve the structure of EEG data within space, time, and frequency.

Herein, we explore the capabilities of deep neural nets for modeling cognitive events from EEG data. In the following chapters we will present recent achievements applying deep learning for EEG signal classification.

## IV. Related Work

In (Stober, et al., 2016) the authors introduce and compare several strategies for learning discriminative features from electroencephalography (EEG) recordings using deep learning techniques. EEG data are high dimensional with a poor signal-to-noise ratio, and there is considerable variability between individual subjects and recording sessions. The proposed techniques specifically address these challenges for feature learning. Cross trial encoding forces auto-encoders to focus on features that are stable across trials. Similarity constraint encoders learn features that allow distinguishing between classes by demanding that two trials from the same class are more similar to each other than to trials from other classes. This tuple-based training approach is especially suitable for small datasets. Hydra-nets allow for separate processing pathways adapting to subsets of a dataset and thus combine the advantages of individual feature learning (better adaptation of early, low-level processing) with group model training (better generalization of higher-level processing in deeper layers). This way, models can, for instance, adapt to each subject individually to compensate for differences in spatial patterns due to anatomical differences or variance in electrode positions. The different techniques are evaluated using the publicly available OpenMIIR dataset of EEG recordings taken while participants listened to and imagined music. In (Stober, et al., 2016) the authors propose several novel techniques for deep feature learning from EEG recordings that address specific challenges of this application domain.

(Walker, 2015) presents a novel application of convolutional neural networks, classifying user intent generated through motor imagery and signaled using EEG data, with the intent of using it as input to a real-time brain-computer interface (BCI). The motivation is to design a system using which a player can control a video game character. The paper proposes a novel method for defining convolutional filters along the scalp to disjoint groups of electrodes that measure activity in similar regions of the brain. The preliminary results from a revised experimental setup still show significant learning and opportunities to improve the results in future research.

Deep learning technology is uniquely suited to analyze neurophysiological signals such as the electroencephalogram (EEG) and local field potentials (LFP) and promises to outperform traditional machine-learning based classification and feature extraction algorithms. Furthermore, novel cognitive computing platforms such as IBM's recently introduced neuromorphic TrueNorth (*Fig.* ) chip allow for deploying deep learning techniques in an ultra-low power environment with a minimum device footprint. In (Nurse, et al., 2016) merge deep learning and TrueNorth technologies for real-time analysis of brain-activity data at the point of sensing will create the next generation of wearables at the intersection of neurobionics and artificial intelligence.

Fig. 2 TrueNorth board deployed to decode neural signals during movement to control robotic arm.

Effectively extracting EEG data features is the key point in Brain Computer Interface technology. (An, et al., 2014) aiming at classifying EEG data based on Motor Imagery task, Deep Learning (DL) algorithm was applied. For the classification of left and right hand motor imagery, firstly, based on certain single channel, a weak classifier was trained by deep belief net (DBN); then borrow the idea of Ada-boost algorithm to combine the trained weak classifiers as a more powerful one. During the process of constructing DBN structure, many RBMs (Restricted Boltzmann Machine) are stacked on top of each other by setting the hidden layer of the bottom layer RBM as the visible layer of the next RBM, and Contrastive Divergence (CD) algorithm was exploited to train multilayered DBN effectively. The performance of the proposed DBN was tested with different combinations of hidden units and hidden layers on multiple subjects, the experimental results showed that the proposed method performs better with 8 hidden layers. The recognition accuracy results are compared with Support vector machine (SVM) and DBN classifier demonstrated better performance in all tested cases. There was an improvement of 4 – 6% for certain cases.

3-D perception is a task that is growing in popularity in television and entertainment. Algorithms and innovations that mimic 3-D perception are of great importance to those in this industry, and as such they need a metric for how well a particular innovation is working. Electroencephalogram (EEG) recordings are an accurate and objective method of evaluating brain activity, and so (Greaves, 2014) primary task is to use EEG recordings score different methods of mimicking 3-D perception. As a first step in doing this the author finds the best features and methods to classify EEG recorded when participants are viewing regular 2D stimuli, and actual 3D stimuli. Hence, in this paper, they explore methods to address the following goal: Can we use EEG signals to accurately classify whether someone is viewing a 2D or 3D image? Using over 5,000 training examples, they investigated the effectiveness of multiple models in achieving this task, with an emphasis on neural networks, and in particular, Recurrent Neural Networks.

In (Jingwei, et al., 2015), the multi-scale deep convolutional neural networks are introduced to deal with the representation for imagined motor Electroencephalography (EEG) signals. We propose to learn a set of high-level feature representations through deep learning algorithm, referred to as Deep Motor Features (DeepMF), for brain computer interface (BCI) with imagined motor tasks. As the extracted DeepMF are dissimilar for different tasks and alike for the same tasks, it is convenient to separate the diverse EEG signals for imagined motor tasks apart. Our approach achieves 100% accuracy for 4 classes imagined motor EEG signals classification on Project BCI - EEG motor activity dataset. Moreover, thanks to the highly abstract features DeepMF learned, only 4.125 seconds trials of training data are needed, compared with the conventional BLDA algorithm for 8.75 seconds trials demand to achieve the same accuracy, accordingly the BCI response time and the required trials for training are almost declined by half. Experiments are provided to illustrate the effectiveness of the proposed design approach.

Signal classification is an important issue in brain computer interface (BCI) systems. Deep learning approaches have been used successfully in many recent studies to learn features and classify different types of data. However, the number of studies that employ these approaches on BCI applications is very limited. In (Tabar & Halici, 2016) the authors aim to use deep learning methods to improve classification performance of EEG motor imagery signals. They investigate convolutional neural networks (CNN) and stacked autoencoders (SAE) to classify EEG Motor Imagery signals. A new form of input is introduced to combine time, frequency and location information extracted from EEG signal and it is used in CNN having one 1D convolutional and one max-pooling layers. The authors also proposed a new deep network by combining CNN and SAE. In this network, the features that are extracted in CNN are classified through the deep network SAE.

In recent years, there are many great successes in using deep architectures for unsupervised feature learning from data, especially for images and speech. In (Zheng, et al., 2014), the authors introduce recent advanced deep learning models to classify two emotional categories (positive and negative) from EEG data. They train a deep belief network (DBN) with differential entropy features extracted from multichannel EEG as input. A Hidden Markov Model (HMM) is integrated to accurately capture a more reliable emotional stage switching. They also compare the performance of the deep models to KNN, SVM and Graph regularized Extreme Learning Machine (GELM). The average accuracies of DBN-HMM, DBN, GELM, SVM, and KNN in the experiments are 87.62%, 86.91%, 85.67%, 84.08%, and 69.66%, respectively. The experimental results show that the DBN and DBN-HMM models improve the accuracy of EEG-based emotion classification in comparison with the state-of-the-art methods.

Deep learning has the advantage of approximating the complicated function and alleviating the optimization difficulty associated with deep models. Multilayer extreme learning machine (MLELM) is a learning algorithm of an artificial neural network, which takes advantages of deep learning and extreme learning machine. Not only does MLELM approximate the complicated function but it also does not need to iterate during the training process. The authors of (Ding, et al., 2015) combine with MLELM and extreme learning machine with kernel (KELM) put forward deep extreme learning machine (DELM) and apply it to EEG classification in this paper. This paper focuses on the application of DELM in the classification of the visual feedback experiment, using MATLAB and the second brain-computer interface (BCI) competition datasets. By simulating and analyzing the results of the experiments, effectiveness of the application of DELM in EEG classification is confirmed.

An alternative pathway for the human brain to communicate with the outside world is by means of a brain computer interface (BCI). A BCI can decode electroencephalogram (EEG) signals of brain activities, and then send a command or an intent to an external interactive device, such as a wheelchair. The effectiveness of the BCI depends on the performance in decoding the EEG. Usually, the EEG is contaminated by different kinds of artefacts (e.g., electromyogram (EMG), background activity), which leads to a low decoding performance. A number of filtering methods can be utilized to remove or weaken the effects of artefacts, but they generally fail when the EEG contains extreme artefacts. In such cases, the most common approach is to discard the whole data segment containing extreme artefacts. This causes the fatal drawback that the BCI cannot output decoding results during that time. In order to solve this problem, (Lia, et al., 2015) employ the Lomb–Scargle periodogram to estimate the spectral power from incomplete EEG (after removing only parts contaminated by artefacts), and Denoising Autoencoder (DAE) for learning. The proposed method is evaluated with motor imagery EEG data. The results show that their method can successfully decode incomplete EEG to good effect.

With the ultimate intent of improving the quality of life, identification of human's affective states on the collected electroencephalogram (EEG) has attracted lots of attention recently. In this domain, the existing methods usually use only a few labeled samples to classify affective states consisting of over thousands of features. Therefore, important information may not be well utilized and performance is lowered due to the randomness caused by the small sample problem. However, this issue has rarely been discussed in the previous studies. Besides, many EEG channels are irrelevant to the specific learning tasks, which introduce lots of noise to the systems and further lower the performance in the recognition of affective states. To address these two challenges, in (Li, et al., 2013), the authors propose a novel Deep Belief Networks (DBN) based model for affective state recognition from EEG signals. Specifically, signals from each EEG channel are firstly processed with a DBN for effectively extracting critical information from the over thousands of features. The extracted low dimensional characteristics are then utilized in the learning to avoid the small sample problem. For the noisy channel problem, a novel stimulus-response model is proposed. The optimal channel set is obtained according to the response rate of each channel. Finally, a supervised Restricted Boltzmann Machine (RBM) is applied on the combined low dimensional characteristics from the optimal EEG channels. To evaluate the performance of the proposed Supervised DBN based Affective State Recognition (SDA) model, the authors implement it on the Deap Dataset and compare it with five baselines.

Time series (particularly multivariate) classification has drawn a lot of attention in the literature because of its broad applications for different domains, such as health informatics and bioinformatics. Thus, many algorithms have been developed for this task. Among them, nearest neighbor classification (particularly 1-NN) combined with Dynamic Time Warping (DTW) achieves the state of the art performance. However, when data set grows larger, the time consumption of 1-NN with DTW grows linearly. Compared to 1-NN with DTW, the traditional feature-based classification methods are usually more efficient but less effective since their performance is usually dependent on the quality of hand-crafted features. To that end, in (Zheng, et al., 2014), the authors explore the feature learning techniques to improve the performance of traditional feature-based approaches. Specifically, they propose a novel deep learning framework for multivariate time series classification. They conduct two groups of experiments on real-world data sets from different application domains.

One of the challenges in modeling cognitive events from electroencephalogram (EEG) data is finding representations that are invariant to inter- and intra-subject differences, as well as to inherent noise associated with such data. In (Bashivan, et al., 2016), the authors propose a novel approach for learning such representations from multichannel EEG time-series, and demonstrate its advantages in the context of mental load classification task. First, they transform EEG activities into a sequence of topology-preserving multi-spectral images, as opposed to standard EEG analysis techniques that ignore such spatial information. Next, the authors train a deep recurrent-convolutional network inspired by state-of-the-art video classification to learn robust representations from the sequence of images. The proposed approach is designed to preserve the spatial, spectral, and temporal structure of EEG which leads to finding features that are less sensitive to variations and distortions within each dimension. Empirical evaluation on the cognitive load classification task demonstrated significant improvements in classification accuracy over current state-of-the-art approaches in this field.

In (Bozhkov, et al., 2017) we propose a new approach for feature dimensionality reduction based on Reservoir Computing (Echo State Networks). The method is validated with EEG data to identify the common neural signatures based on which the positive and negative valence of human emotions across multiple subjects can be reliably discriminated. The key step in the proposed approach is the Intrinsic Plasticity (IP) adaptation of the reservoir states. Learning Echo State Networks (ESN) with IP maximizes the entropy of the distribution of reservoir vectors given static data as a fixed input, which is supposed to follow Gaussian distribution. The equilibrium reservoir vector is extracted for each static input vector by iterating updates of the reservoir vector until it converges. Standard classification and clustering models provided with selected combinations of reservoir neurons are ranked based on their discriminate performance. The IP tuned ESNs is more powerful technique to map the high dimensional input feature vector into a low dimensional representation and improve the emotion valence discrimination compared to classical ESNs and Deep Neural Encoders.

The aim of this study (Bozhkov, et al., 2016) is to identify the common neural signatures based on which the positive and negative valence of human emotions across multiple subjects can be reliably discriminated. The brain activity is observed via Event Related Potentials (ERPs). ERPs are transient components in the Electroencephalography (EEG) generated in response to a stimulus. ERPs were collected while subjects were viewing images with positive or negative emotional content. Building inter-subject discrimination models is a challenging problem due to the high ERPs variability between individuals. We propose to solve this problem with the aid of the Echo State Networks (ESN) as a general framework for extracting the most relevant discriminative features between multiple subjects. The original feature vector is mapped into the reservoir feature space defined by the number of the reservoir equilibrium states. The dominant features are extracted iteratively from low dimensional combinations of reservoir states. The relevance of the new feature space was validated by experiments with standard supervised and unsupervised machine learning techniques.

## V. Results

In (Walker, 2015) the highest accuracy achieved in the revised results is 80.08% for subject KP in the left and right hand imagery experiments, a significant improvement over a random model, which would be expected to have a 50% accuracy using CNN.

To determine which music piece somebody listened to based on the EEG is a challenging problem. In (Stober, et al., 2016) trained CNN classifiers for a 12-class problem and achieved accuracies significantly above chance in the ranges of 20 – 27.8%.

In (Nurse, et al., 2016) see *Fig.* , the authors achieve 81% accuracy on binary classification using CNN of a single participant who performed a self-paced hand squeeze task. They were also able to transfer the model to a TrueNorth chip thus decreasing the gap between machine learning and deep learning software and wearable hardware.

(Greaves, 2014) concludes that it is not straightforward to apply RNNs to EEG data. While it is possible that a more complex RNN would have done better at classification, it seems that a simple feed-forward network outperforms a simple RNN for binary discrimination using EEG data of subject watching 3D or 2D stimuli. Still, both RNN and FF NN models managed to get fairly above chance results, and so future work should involve researching more complex recurrent models for EEG data.

In (Jingwei, et al., 2015) the authors achieve 100% accuracy for 4 classes imagined motor EEG signals classification using CNNs.

The classification performance obtained by the proposed method in (Tabar & Halici, 2016) on BCI competition IV dataset 2b in terms of kappa value is 0.547. Their approach yields 9% improvement over the winner algorithm of the competition. Their results show that deep learning methods provide better classification performance compared to other state of art approaches. These methods can be applied successfully to BCI systems where the amount of data is large due to daily recording.

The experimental results in (Zheng, et al., 2014) show that the Deep Belief Networks and Deep Belief Networks with Hidden Markov Models improve the accuracy of EEG-based emotion classification in comparison with the state-of-the-art methods.

The results in (Lia, et al., 2015) show that their method can successfully decode incomplete EEG to good effect using Denoising Autoencoder (DAE) for learning.

Extensive experimental results from (Li, et al., 2013) show that the proposed algorithm (DBN and RBM) can successfully handle the aforementioned two challenges and significantly outperform the baselines by 11.5% to 24.4%, which validates the effectiveness of the proposed algorithm in the task of affective state recognition.

The final results in (Zheng, et al., 2014) show that the novel deep learning framework for multivariate time series classification is not only more efficient than the state of the art but also competitive in accuracy. It also demonstrates that feature learning is worth to investigate for time series classification (Zheng, et al., 2014).

In the results of (Bashivan, et al., 2016) the authors demonstrated significant improvements in classification accuracy over current state-of-the-art of the trained deep recurrent-convolutional network inspired by state-of-the-art video classification to learn robust representations from the sequence of images.

In (Bozhkov, et al., 2017) we shown that using reservoir computing pre-training is beneficial for selecting the most relevant discriminative features and reaching state-of-the-art performance for subject independent emotion valence recognition. This study confirmed the observation that the reservoir computing approach can be used not only for time series processing but also for high dimensional static data representation.

The relevance of the ESN extracted feature vectors in (Bozhkov, et al., 2016) was validated by experiments with typical clustering and classification techniques. 80% of correct clustering of the positive and negative emotion valence across multiple subjects obtained with feature vectors having the surprising low dimension of 2 elements is a very promising result in the context of this challenging application. Moreover, 95% correct classification obtained with 4D feature vectors outperform all previously published results of affective computing based on brain neural data (ERPs).

## VI. Conclusion

In the studies reviewed above, we showed that using recent deep learning architecture, we can achieve very high accuracy and efficiency on decoding the human brain signal recorded using EEG. In recent years deep learning proved to be excellent technique in solving vastly different and hard problems. Deep learning techniques proved to have better accuracy and efficiency then humans in various tasks such as recognition in images, sound and text. DL also helped in beating the human champion in the game of Go. The reviewed studies show that DL will also help for solving problems for other domains such as EEG data. The necessity of huge amount of EEG recorded and labeled signal and data, however stays and is even severer for DL research to achieve exceptional results.

## References


An, X. et al., 2014. A Deep Learning Method for Classification of EEG Data Based on Motor Imagery. *Intelligent Computing in Bioinformatics,* pp. 203-210.

Bashivan, P., Rish, I., Yeasin, M. & Codella, N., 2016. Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks. *ICLR 2016.*

Bozhkov, L., Koprinkova-Hristova, P. & Georgieva, P., 2016. Learning to decode human emotions with Echo State Networks. *Neural Networks, Volume 78, June 2016, Pages 112–119.*

Bozhkov, L., Koprinkova-Hristova, P. & Georgieva, P., 2017. Reservoir computing for emotion valence discrimination from EEG signals. *Neurocomputing, Volume 231, 29 March 2017, Pages 28–40.*

Ding, S. et al., 2015. Deep Extreme Learning Machine and Its Application in EEG Classification. *Mathematical Problems in Engineering, vol. 2015, Article ID 129021.*

Graves, A., Mohamed, A.-R. & Hinton, G. E., 2013. Speech recognition with deep recurrent neural networks.. *s. Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on.*

Greaves, A. S., 2014. Classification of EEG with Recurrent Neural Networks.

Hermann, K. M. et al., 2015. Teaching Machines to Read and Comprehend.. *arXiv.*

Jingwei, L., Yin, C. & Weidong, Z., 2015. Deep learning EEG response representation for brain computer interface. *Control Conference (CCC), 2015 34th Chinese.*

Karpathy, A. & Toderici, G., 2014. Large-scale video classification with convolutional neural networks.. *Computer Vision and Pattern Recognition (CVPR).*

Krizhevsky, et al., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems.*

Lia, J., Struzik, Z., Zhang, L. & Cichocki, A., 2015. Feature learning from incomplete EEG with denoising autoencoder. *Neurocomputing, Volume 165.*

Li, K., Li, X., Zhang, Y. & Zhang, A., 2013. Affective state recognition from EEG with deep belief networks. *2013 IEEE International Conference on Bioinformatics and Biomedicine.*

Mak, J. N. & Wolpaw, J. R., 2009. Clinical applications of brain-computer interfaces. *IEEE reviews in biomedical engineering, vol. 2.*

McFarland, D. & Wolpaw, J., 2008. Brain-computer interface operation of robotic and prosthetic devices. *Computer, vol. 41.*

Mirowski, P., Madhavan, D., LeCun, Y. & Kuzniecky, R., 2009. Classification of patterns of EEG synchronization for seizure prediction.. *Clinical Neurophysiology, 120(11):1927–1940, 2009..*

Ng, J. et al., 2015. Beyond Short Snippets: Deep Networks for Video Classification. *CVPR.*

Nurse, E. et al., 2016. *Decoding EEG and LFP Signals using Deep Learning: Heading TrueNorth.* s.l.:s.n.

Plis, S. M. et al., 2014. Deep learning for neuroimaging: a validation study.. *Frontiers in Neuroscience, 8(August):1–11, 2014.*

Stober, S., Sternin, A., Owen, A. M. & Grahn, J. A., 2016. DEEP FEATURE LEARNING FOR EEG RECORDINGS. *International Conference on Learning Representations.*

Tabar, Y. R. & Halici, U., 2016. A novel deep learning approach for classification of EEG motor imagery signals. *Journal of Neural Engineering, Volume 14, Number 1.*

Walker, I., 2015. *Deep Convolutional Neural Networks for Brain Computer Interface using Motor Imagery.* s.l.:s.n.

Zhang, X. & LeCun, Y., 2015. Text Understanding from Scratch.. *arXiv preprint arXiv:1502.01710, 2015..*

Zheng, W.-L., Zhu, J.-Y., Peng, Y. & Lu, B.-L., 2014. EEG BASED EMOTION CLASSIFICATION USING DEEP BELIEF NETWORKS. *IEEE International Conference on Multimedia & Expo.*

Zheng, Y. et al., 2014. *Lecture Notes in Computer Science, Volume 8485, pp 298-310.*

Bozhkov L., P. Georgieva, R. Trifonov, Brain Neural Data Analysis Using ML Feature Selection and Classification Methods. *EANN 2014, pp. 123–132, Springer 2014*