# 

[**INTRODUCTION**](#_beli6smw0qs1) **1**

[**Deep Learning with MRI**](#_czmgnpq7ima4) **2**

[Alzeimer’s Disease Detection](#_ezohi2dwhtp7) 2

[Multiple Sclerosis Lesion Segmentation](#_6psk2oi81ltq) 4

[**Deep Learning with EEG**](#_1dprzf271rwx) **9**

[Epilepsy Detection and Prediction](#_6wrb01ebyknn) 10

[Cognitive Task Detection](#_ds2hmezifblw) 11

[**CONCLUSION**](#_5edt5khisc2a) **13**

# INTRODUCTION

Deep Learning has been quite successful recently in the area of Computer Vision, achieving improvements in accuracies sometimes as high as 30 %. The main strength in deep learning, also differentiating it from other methods, is their automatic feature extraction capability. Normally, raw data has to processed automatically or manually to extract meaningful and useful features. This process requires time and careful analysis, and includes subjectivity on the part of the expert, which might bias the results or produce erroneous results. However, in deep learning, the feature extraction is driven by data, an appropriate loss function and a learning algorithm, which removes the subjectivity, randomness and expert knowledge to a certain degree. Moreover, the features obtained are hierarchical, each layer producing more abstract features using the less abstract features obtained in the previous layer. This way feature extraction is carried out step-by-step, which is more likely to produce more complex and useful features. Another strength of deep learning is its ability to represent very complex functions, which might be also its weakness since it is prone to easily overfitting, but with the correct guidance and regularization methods the overfitting can be prevented. Deep learning methods are also robust to outliers, which is very common in neuroimaging data.

There are also disadvantages to deep learning methods such as hard-to-explain nature of it and computational complexity. Visualisation of the features in methods like Convolutional Neural Networks(CNN) helps for better explainability of deep learning models. Novel methods such as CNNs/Convolutional Encoder Networks(CENs) provide weight sharing methods which eases the computational costs. Moreover, improving hardware (GPUs) and the advent of cloud technologies(for example Amazon Web Services)removes the computational efficiency need to a great extent.

With the growing success of Deep Learning in various fields including Computer Vision, neuroscience has also taken an interest in deep learning methods for its own problems so far unsolved or partially solved. Similar to the success obtained in other fields, substantial improvements have been achieved using these techniques compared to other machine learning methods or conventional methods. Deep learning has been applied to a wide range of problems such as classification or detection of diseases affecting the brain(such as Multiple Sclerosis, Alzheimer’s, Schizophrenia, Epilepsy), segmentation of brain lesions caused by such diseases, classification or detection of cognitive states/functions etc. In this essay, I will analyze some of the studies carried out with deep learning applied on some of these problems using either MRI or EEG data.

# Deep Learning with MRI

Magnetic Resonance Imaging (MRI) is a spectroscopic medical imaging technique to form pictures of the anatomy and medical processes of the body, which make use of magnetic fields, radio waves or field gradients. MRI is a very powerful technology in neuroscience allowing to investigate the details of human brain in various perspectives. It is so flexible in the sense that one can measure the macrostructural features of the brain (such as cortex thickness or volumes of different sub-structures), microstructural features (such as diffusion within white matter tracts) or functional activations in the brain (task-driven activations or resting state properties). The scale of the MRI image data may vary depending on the modality and the resolution used. MRI can be considered under three categories, structural MRI, functional MRI and diffusion MRI.

**sMRI:** sMRI provides the static anatomical information, such as the size and shape of the white and gray matter. In structural MRI, different structures appear in different levels of contrast. It is used to observe the structural abnormalities in the brain, such as tumors, lesions, etc. It has several modalities, such as T1, T2 and FLAIR (See Figure 1.).

**fMRI:** Functional MRI (fMRI) measures indirectly the brain activity by directly measuring the oxygenation level of blood throughout the brain. The main aim of this technique is to find a relationship between structure and function. The scanning is typically accompanied with a behavioural task and it is done taking many scans through time within a short time period. Thus, the resulting data is a 4D image, which is converted to a 3D contrast image.

**Diffusion MRI:** In this scanning technique the relative motion of water through each voxel in different directions is captured. A tensor analysis is carried out to calculate the direction and the intensity of the motion through each voxel. Since it is known that water moves more easily along axons than across them, this technique mainly gives information about the connectivity throughout the brain. The abnormality in the connections may indicate a structural or functional abnormality in the brain, such as tumors.

## Alzeimer’s Disease Detection

Alzheimer’s disease is a progressive, degenerative brain disorder characterized by the death of neurons throughout the brain which results in the loss of cognitive functions mainly related with the memory. At the advanced stages, symptoms become so severe that living becomes a great challenge for the affected person. It is a common form of dementia among the elderly. Since the damage in the brain begins long before the onset of the symptoms, it is possible to detect the disorder in its early stages. Early detection is critical to delay and even to prevent the onset of the life-debilitating symptoms, to devise a health care plan, to take safety measures for the patient and to develop new medicines. Historically sMRI, fMRI, PET and CSF images are used to obtain features to detect the Alzheimer’s disease. sMRI stands out particularly to better show the progression of the disease.

In the study of [1] features are being extracted from sMRI data using a stack of unsupervised 3D convolutional autoencoder(3D-CAE) followed by a supervised 3D convolutional neural network(3D-CNN). In the 3D-CAE stage, the unlabeled images are pre-trained to obtain generic features while in the supervised 3D-CNN stage, the classification is obtained by fine-tuning the network with labeled images. The generic features obtained are general AD biomarkers such as ventricular size, hippocampus shape, cortical thickness etc. The generic features obtained in the 3D-CAE stage can be seen in the figure 1. The accuracy rates obtained by this technique were superior to other conventional methods surpassing 90%.

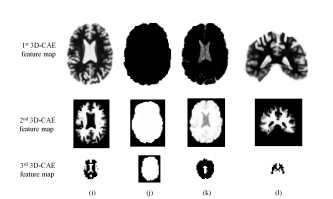


Figure 1- Generic features obtained by 3D-CAE

In another study on Alzheimer’s disease detection using deep learning, two types of brain imaging is used as data. sMRI, which provides information about structural integrity of the brain; and PET scans, providing information glucose metabolism and amyloid deposits in the brain because Alzheimer’s disease is known to produce abnormal changes in both tissue loss and a change in the glucose and amyloid metabolism. The data used is obtained from Alzheimer’s Disease Neuroimaging Initiative(ADTI) database, which is an initiative to bring together and offer publicly brain data to encourage research on the Alzheimer’s disease. Since the number of data available was was small, instead of training a neural network, which might end up overfitting, the researchers decided to use pre-trained neural network(Alex-Net trained on Imagenet) and fine-tune it with the available data. AlexNet is a Convolutional Neural Network consisting of 5 convolutional layers and 3 fully connected layers. For the training they obtained 2D slices with varying view angles from the brain image and augmented the data by taking mirror images. To merge information from different modalities(MRI and PET), they fine-tuned separate AlexNets up to the final classifier for each angle or modality and combined them at the classifier layer(see Figure 2). They experimented with changing the view angles or modalities used in training to see which ones are more useful in obtaining good accuracy. According to their findings, PET images are more likely to provide valuable clues as to the onset of the Alzheimer’s disease than MRI images.

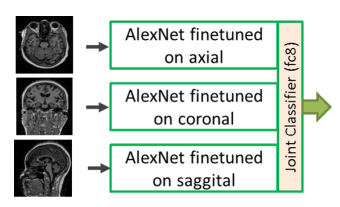


Figure 2 - AlexNet on Alzheimer’s detection

## Multiple Sclerosis Lesion Segmentation

Multiple Sclerosis (MS) is a common neurological disease that afflicts especially the young population and it remains a very challenging disease to diagnose and treat, due to its variability in its clinical expression. MS is characterized by lesions throughout the brain that is caused by the loss of myelin sheath around neurons in the brain, which is also known as demyelination. The lesions are generally ovoid in shape, are of varying sizes, are scattered throughout the brain and are seen mainly in the white matter, and sometimes in the gray matter of the brain. The diagnosis and prognosis of MS is currently guided by conventional structural MRI of brain and spinal cord, seeking for evidence of both dissemination in time and dissemination in space of the MS lesions.

The number of MS lesions and the total volume of the MS lesions are indicative of the disease stage and are used to track disease progression. Therefore, the accurate segmentation of lesions is quite important for MS disease in the medical world with the purposes of correct diagnosis, adequate treatment development and prognosis follow-up. Manual segmentation by experts is the most commonly used technique of MS lesions and is still considered to produce the most accurate results although it suffers from many complications. First of all, it is subject to intra- and inter- expert variability, which means there are significant differences between two segmentations performed by two different experts(due to slightly varying definitions) or by the same expert at different times (due to fatigue). Secondly there is a shortage of adequately trained experts given the huge amount of segmentation need. Thirdly, the segmentation task requires valuable expert time and concentration, which could be dedicated to other tasks. These drawbacks with manual segmentation makes it necessary and desirable to develop a semi-automatic segmentation method that would assist experts in the task with a reduced amount of time and intra- inter- expert variability or, in the ideal case, a fully automatic segmentation method which would obviate the need for experts and produce accurate/reproducible results.

Lately, deep learning methods have become rather popular in MS lesion segmentation due to the segmentation performances obtained very close to the human expert level. One example to deep learning methods applied applied to MS segmentation problem is a voxel-wise classifier with a 3 layer convolutional neural network(CNN) used in [Vaidya2015]. According to this architecture, multiple channels of 3D patches of are extracted from MRI images of modality T1, T2 and FLAIR and fed to a CNN. The network has two convolutional layers and a fully-connected network with average pooling layers in-between CNN layers. For efficient training, sub-sampling methods and sparse convolutions are used. This implementation won the 2015 MS Challenge. According to the authors, dice scores comparable to inter-rater results were obtained. The advantage of this implementation is its relative simplicity to other CNN architectures. The disadvantage is the segmentation is done voxel-by-voxel which might mean long training and testing times.

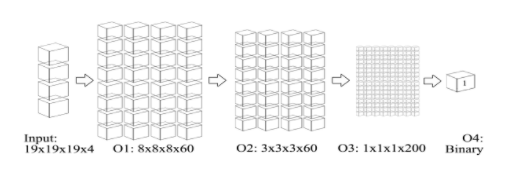


Figure 3 -

Another example to a CNN solution with 3D patches is a 3 layer convolutional neural network(CNN) explained in [Brosch2015]. The network consists of 3 layers, the input layer of voxels, one convolutional layer that extracts features for each voxel, one deconvolutional layer which predicts the class of the voxel using previously extracted features. This special type of CNN, consisting of convolutional layers followed by deconvolutional layers is called Convolutional Encoder Network(CEN) (see Figure 4) During training entire MRI volumes as samples are used instead of patch-based training. The entire network is trained all at once, therefore, the learning process is guided by segmentation performance. The proposed architecture is similar to a convolutional auto-encoder. A special objective function with weighted sensitivity and specificity measures is used to deal with the unbalanced nature of the training dataset, i.e. the error measure is a weighted average of squared differences for lesion and non-lesion voxels. 3 MRI modalities(T1, T2, FLAIR) are exploited. The use of convolutional and deconvolutional layers allow the segmentation result to be of the same resolutions as the input, which is the whole brain. The advantage of this implementation is although one training step might be more computationally intensive, the whole training time is reduced since the whole MRI image is used as one sample in training. Testing time is also significantly lower than a patch-based testing which evaluates one voxel at a time.

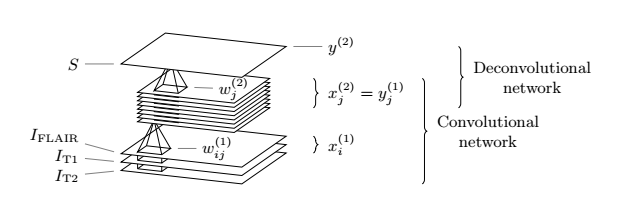


Figure 4 - A simple CEN Architecture

In a study one year later, the same authors proposed a similar architecture [Brosch2016] with convolutional and deconvolutional pathways using the whole brain as input. This is a special type of CNN, which is called Convolutİonal Encoder Network (CEN). Initially, a convolutional Restricted Boltzmann Machine(cRBM) is trained, with convolutional layers and pooling layers in between, to pre-train the initial weights and biases. With this initialization, another network is trained consisting of a convolutional pathway followed by a deconvolutional pathway. In the convolutional pathway, there are alternating convolutional layers and pooling layers. The low-level features obtained from the first convolutional layer of this pathway are kept to be used later. The features resulting from the second convolutional layer are fed to the deconvolutional pathway, which consists of alternating deconvolutional and unpooling layers. To the high-level features obtained after the first deconvolutional+unpooling layer are added the low-level features obtained in the previous stage, to obtain a probabilistic lesion mask. The reason to use the deconvolutional layers is to obtain a result the same size as the input, so that the segmentatıon can be done as a whole. As can be seen from Figure 5 the number of convolutonal layers has been varied to see the effect of increasıng the depth. The advantage to this method is again the segmentation of the whole MRI image in a single-step. The disadvantage is the relative complexity of the model which might lead to overfitting.

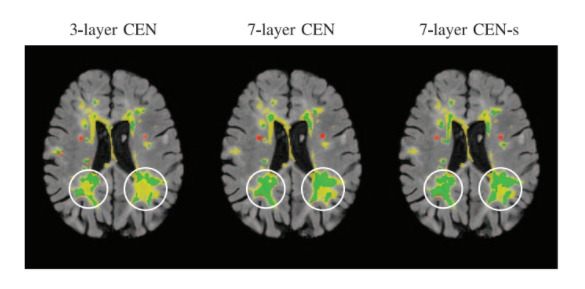


Figure 4 - As the depth of the CEN is increased the segmentation of the bigger lesions gets better due to the increased size of the receptive field.

In [Ghafoorian2016] they propose to integrate the anatomical location in CNN architectures to increase the accuracy, since the probability of a lesion occurring in certain regions are higher. Different from the previous architectures, 2D patches are used instead of 3D patches. Two different ways of feeding the location information the network is implemented, the first by considering multi-scale patches and the second by adding the location features explicitly. 2D patches are sampled from 500 patients, selecting the positive and negative samples in a random fashion. The patches are selected in 3 different sizes, 32\*32, 64\*64 and 128\*128. The first architecture they propose is a CNN with 4 convolutional layers followed by 3 fully connected layers. This architecture uses only patches of size 32\*32. In a second architecture, they consider all the different scales as the input and down-sampling the 64\*64 and 128\*128 patches into 32\*32 patches, they obtain 3-channels for each input. This architecture takes into both the local information with the smaller patches and the global information with the bigger patches. Another architecture proposed in the same study trains 3 different convolutional branches for each input type and then fuse them with fully-connected layers. Alternatively to this architecture where global information is taken into account, they propose to use only the smallest patches and add explicit location features at the fully-connected layer. They find that incorporating the spatial location information improves the segmentation results. The advantage to this approach is the explicit or implicit location information added to the network, which might lead to improved accuracies. Another advantage is the reduced training and testing times due to 2D patches. However, usage of 2D patches also create a disadvantage, because some valuable information is lost due to the usage of 2D patches instead of 3D patches.

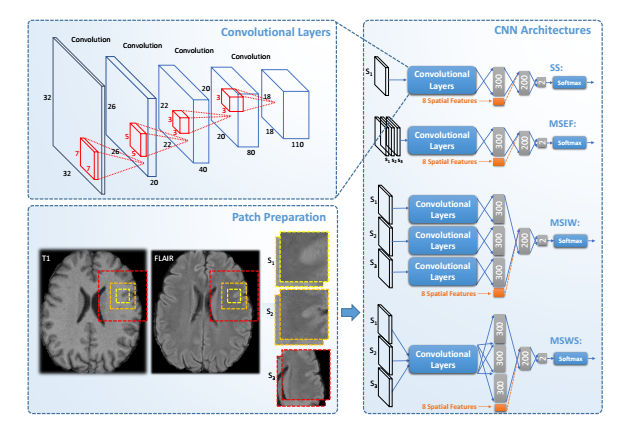


Figure 5 - Alternative proposed architectures in the study by [Ghafoorian2016]

An improvement to the implementations suggested above could be the cascade based approach in [Valverde2017]. This implementation is based on a cascade based training with two stages, using 3D patches of size 11\*11\*11. For the first training step a CNN with two convolutional layers each followed by max-pooling layer and a fully-connected layer followed by a soft-max layer at the end is used. The novel idea in this paper is that first a relatively "coarser" CNN is trained, with the explained architecture, which is supposed to find the "candidate" lesions and, later, using the wrongly classified negative examples from the application of the first CNN (together with all the positive ones), a "finer" CNN with the same architecture is trained, which is more sensitive, because trained with harder cases. This two-stage technique is applied during testing as well, during which the first network is used to find the “candidate” lesions, discarding the low-probability results. These candidate lesions are later fed to the second network to make finer predictions. This idea might also be used to select the training samples in a more smart fashion but train only one network in the end to be used during testing, which might decrease the testing time. The advantage of this approach is the two stage evaluation, which might increase the accuracy results. The disadvantage is high training and testing times.

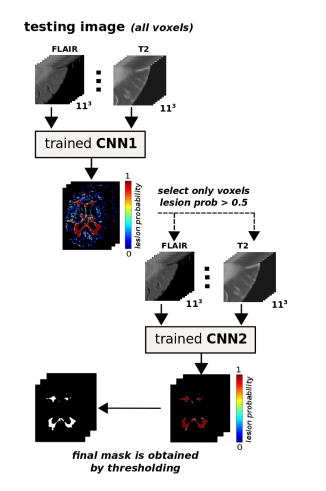
****

Figure 6 - Testing in the proposed architecture by [Valverde]

# Deep Learning with EEG

EEG is a non-invasive invasive technology to measure the electrical activity in the brain. Unlike MRI, it does not require a patient to remain stationary in an uncomfortable position in complex and expensive machines but it is simply based on electrodes placed carefully on the correct spots of the scalp of a patient. EEG can produce high time resolution data, which makes it also suitable for real time systems. However, its spatial resolution is quite low compared to MRI since it is the collection of the noisy outputs collected from at most a hundred or so electrodes. Moreover, it can hardly capture the electrical activity deep within the brain since the electrodes are placed superficially on the scalp. It is also subject to noise and artifacts caused by other internal or external factors such as heart beating, moving of a limb or the rhythmic pulse of the country’s electrical system.

EEG classification has been quite popular in recent days for detecting brain conditions such as seizures or strokes, or simply to detect basic brain activities such as intent classification. Conventional methods for EEG classification rely on frequency information, which is extracted from the EEG signals by Fourier transform or other transforms. Deep learning, like in many other classification problems with other types of data, is becoming very popular and promising with EEG data. Recurrent Neural Networks are specifically the favourite deep learning method with EEG due to its spatial/temporal nature.

## **Epilepsy** **Detection and Prediction**

Nearly one percent of the world population suffers from epilepsy and in many cases even medication and surgery do not completely remove the seizures from the patient’s daily life. The uncertain nature of the seizures is a stress factor for a patient and sometimes a threat to his/her safety. Therefore, it would substantially improve the patient’s quality of life to be able to predict beforehand if a seizure is imminent or detect seizure occurrence for more accurate diagnosis and prognosis.

In a study with deep learning using EEG data, the epilepsy seizure prediction, which is also known as online prediction, is taken as an objective. The first step applied is feature reduction to the EEG signal using Principal Component Analysis and Independent Component Analysis in the aim to increase classification accuracy(reducing the noise) and decrease computational cost. This decreased dimensional input is later fed to a two layer stacked-autoencoder and each layer is trained individually in an unsupervised fashion. Thus, high numbers of unlabeled data can be exploited to train the stacked autoencoder. These two layers are followed by a softmax layer to produce probabilities for each class and the whole network is trained in a supervised fashion after the initial unsupervised trainings of autoencoder layers. The power of cloud computing is leveraged to handle the storage and computational cost requirements. The accuracy obtained is reported to be 94% which showed a big improvement over other methods such as random forest, SVM or MLP Neural Networks.

In one study[12] with EEG data and deep learning, epilepsy seizure detection (also called offline detection) is targeted. One of the challenges with this problem is the inter- and intra-patient variability in the manifestation of seizures in EEG signals. Detection of seizures and pre-ictal spikes, which are the characteristic features of epilepsy, are important in the diagnosis and prognosis of the condition. This task is normally performed by trained professional requiring several hours of careful inspection for one day of recording. The scarcity of such professionals and the prevalence of this condition makes automated methods necessary. Most automatic methods so far developed depended on hand-crafted features obtained from EEG signals. However, these features need to be robust to inter- and intra- patient variability, which is a difficult task. The feature extraction ability of deep learning techniques together with the robustness of these features, previously proven in other computer vision and speech recognition tasks, makes these techniques particularly suitable for the task of seizure detection. The deep learning technique proposed in this study is a recurrent convolutional network, specifically chosen to make use of the spatial and temporal and spectral information from the EEG signals. The data used consists of a diverse set of patients with several observations and contains different types of seizures. The 30-second EEG signals are converted into image based representations before being fed to the recurrent convolutional neural network. Long short term memory(LSTM) units are used for the temporal nature of the data. The results are compared with the commercialised algorithm called REVEAL and increased average sensitivity rate (85% as opposed to 67%) and decreased false positive rate (1.7/hour as opposed to 0.8/hour) is reported with this deep learning technique.

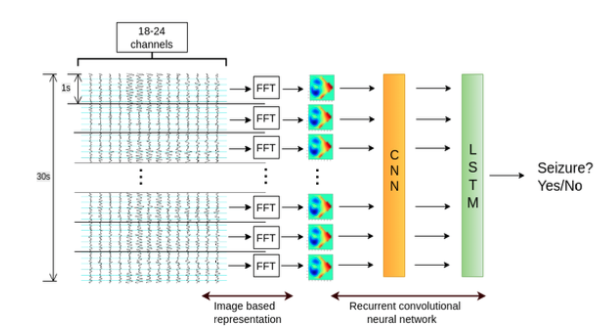


Figure 7 - Recurrent Neural Network using image based representation of EEG data

## Cognitive Task Detection

In another study researchers try to convert electrophysical signals from the brain obtained with EEG technology into the classification of user intent using a deep learning method. Normally experts are needed to convert these signals to useful features which can be implicated in the involvement of an action. The idea of using deep learning is based on the ability of deep learning to extract useful features for the task at hand. For preprocessing the data, the artifacts are removed from the signal using z-scores and to further remove the remaining noise signal processing methods such as Fourier transform and Morlet Wavelet is applied to the signal(though it is argued that this may not be necessary due to the ability of CNNs to automatically learn and perform these tasks). The preprocessed output from the electrodes are then fed to a CNN with one convolutional layer and two fully connected layers followed by a logistic regression to classify the output. In the convolutional layer, the filters are arranged in a fashion that they are applied to 3 electrodes for 5 observations(spatial and temporal dimensions). According to the author, this study suffered from very few training data and, therefore, overfitting but promising results were obtained encouraging further experiments on the problem using deep learning.

Another interesting study[9] with EEG deals with the classification of 3D/2D perception of the brain. The aim of this study is to understand if EEG data can be used to detect if a person is seeing a 2D or a 3D image and if deep learning is effective in this task. The subjects are presented with 2D and 3D images and their EEG signals from 88 electrodes are recorded during the task. Later the EEG signals are processed to extract raw signal features(instead of Fourier transform). 5 methods are compared in this study, elastic net regression which is not a deep learning method, 1-layer and 2 layer multi layer perceptron(MLP), Elman Recurrent Neural Network and Time-dependent Elman Recurrent Neural Network. Surprisingly the two RNNs were the worst performing methods, possibly due to overfitting. However, the MLPs outperformed all the other methods, including the conventional non-deep learning method, with the best test accuracy value of 0.72.

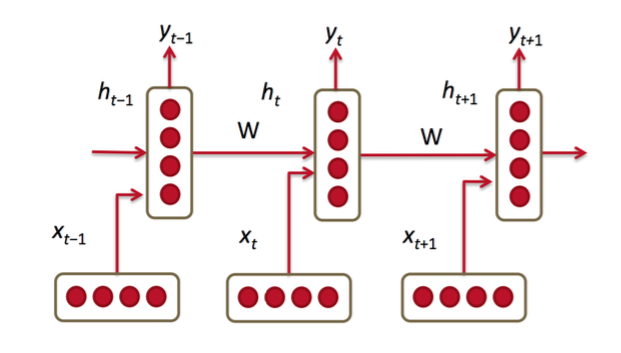


Figure 8 - Elman Recurrent Neural Network Structure

# CONCLUSION

The success obtained by applying deep learning methods on a wide range of problems in neuroscience should be more than mere chance. Although it is difficult to explain the obtained model, there are many clues as to why successful models are obtained with deep learning. There are a huge number of practical and theoretical studies concerning deep learning that we are informed about a new deep learning technique or approach, or an improved fine-tuning technique, activation function, loss function etc. very often. Moreover, the data in neuroscience is accumulating and getting better in quality, and there are more and more efforts to centralize and the standardize the data. All these improvements will work to the advantage of deep learning methods, which require a high amount of data and appropriate tuning to perform better. The results already obtained in neuroscience problems are already very promising and considering the efforts put in improving the existing techniques and data, the results in the future are set to be even better.

8. Bibliography

[1] Essan Hosseini Asl, Robert Keynton, Ayman El Baz, Alzheimer’s Disease Diagnostics by Adaptation of 3D Convolutional Neural Network, ICIP, 2016

[2] Tanya Glozman, Orly Liba, Hidden Cues: Deep Learning for Alzheimer’s Disease Classification, Stanford University, 2016

[3]

[4] Bhatkoti Pushkar, Manoranjan Paul, Early Diagnosis of Alzheimer's Disease: A Multi- class Deep Learning Framework with Modified k- sparse Autoencoder Classification, Charles Sturt University, 2016

[5]

[6]

[7]

[8] Sergey M. Plis, Devon R. Hjelm, Ruslan Salakhutdinov, Elena A. Allen, Henry J. Bockholt, Jeffrey D. Long, Hans J. Johnson, Jane S. Paulsen, Jessica A. Turner and

Vince D. Calhoun, Deep learning for neuroimaging: a validation study

[9] Alex S. Greaves, Classification of EEG with Recurrent Neural Networks, Stanford University

[10] Ian Walker, Deep Convolutional Neural Networks for brain Computer Interface using Motor Imagery, Imperial College London, 2015

[11] M.P. Hosseini, H. Soltanian-Zadeh, K. Elisevich, D. Pompili Cloud-based Deep Learning of Big EEG Data for Epileptic Seizure Prediction, IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, 2016.

[12] Pierre Thodoroff, Joelle Pineau, Andrew Lim, Learning Robust Features using Deep Learning for Automatic Seizure Detection, 2016