

Using Deep Learning algorithms to detect the success or failure of the Electroconvulsive Therapy (ECT) sessions

¹Cyrus Kalantarpour, ¹Usef Faghihi, ²François-Xavier Roucaut,

¹Université du Québec à Trois-Rivière

²Université du Québec à Montréal

cyrus.kalantarpour@uqtr.ca, usef.faghih@uqtr.ca, francois-xavier_roucaut@ssss.gouv.qc.ca

Abstract

Major Depression Disorder (MDD) is a big problem in our society. MDD can cause suicide and take families apart. When treatment with medications fail, mental healthcare professionals, use Electroconvulsive Therapy (ECT) to treat patients with MDD. During an ECT session, electroencephalogram (EEG) signals let the mental healthcare professionals record patients' brain activities which are helpful to decide whether the treatment was successful. However, there is no standard way to know how and with what intensity a healthcare professional needs to apply electroshock to treat patients with MDD. So far, to our knowledge, researchers have used multi-parametric magnetic resonance imaging (MRI) techniques combined with statistical methods and/or linear machine learning algorithms to predict patients' responses to ECT. However, the aforementioned methods are very expensive and time-consuming. In this study, we will be using Deep learning algorithms to detect the effectiveness of ECT sessions based on the EEG.

1 Introduction

Major Depression Disorder (MDD) is the cause of more than one million suicide per year (Sun et al. 2020). Electroconvulsive (ECT) therapy has been used by mental healthcare professionals since 1930 to treat patients with MDD (Tsuchiyama et al. 2005). Yet, there is no methodological technique to individualize ECT in order to obtain successful results. So far, most of the researchers used multi-parametric magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and/or resting state-fMRI techniques combined with statistical methods and/or support vector regression model (Gong et al. 2020, Gärtner et al. 2020, Van Waarde et al. 2015), to categorize and predict the success or failure of the ECT method with 67-70% accuracy (Sun et al. 2020). Furthermore, most of the aforementioned studies suggest frontal and temporal networks of the brain as good predictors for the success or failure of ECT

treatments. In another study, (Min et al. 2020), limited their experience to MDD patients suffering from schizophrenia. They used a random-forest algorithm with fMRI for their predictions. Their results suggest higher connectivity in the patient's frontal area with 82% accuracy of predicting successful and unsuccessful ECTs. Another limitation with this study is that the authors only used patients with schizophrenia.

Furthermore, using MRI/ fMRI/rs-fMRI with ECT is very expensive and time consuming. In this study, we suggest the use of Deep learning algorithms to predict the result of the ECT from patient pre-during-post shock EEGs. Comparing to MRI family techniques mentioned above our technique uses patients' brain EEGs pre-shock to predict the success or failure of the ECT techniques. This makes our approach less expensive and timely efficient comparing to the above techniques. Deep learning algorithms are capable of learning unseen patterns (Faghihi et al. 2020, Robert et al. 2020). We believe that DLs can help us to individualize ECT techniques for patients suffering from MDDs.

Among others, they are used for detecting different types of cancers (Cruz-Roa et al. 2013), sentiment analysis (Baziotis, Pelekis and Doulkeridis 2017), detecting cataract (Yu et al. 2019).

However, to our knowledge, so far, there is no study used deep learning algorithms for predicting the success or failure of the ECT technique only based on EEG data. One reason is that EEGs are very complex and DLs cannot use EEGs directly (Hu and Zhang 2019). Furthermore, while some of the mental health professionals may use lateral regions of the brain for the ECT shocks, others may use lateral and frontal regions. This will result in DLs behave differently when the nature of data changes (Chen et al. 2020).

Thus, before applying DLs to data, one needs to do preprocessing and adapting the data in a way that DLs can process them.

In what follows, we will first very briefly explain EEGs preprocessing techniques. We will also explain how one can distinguish between a good and bad ECT using EEG traces. We then, very briefly explain DL architecture we used in this study. Finally, we will compare different DLs performance for ECT outcome prediction.

2 Pre-Processing EEGs

This section is divided into two subsections 1) using correlation technique to find whether there are correlations between Hyperpolarization, Depolarization, and Repolarization phases and post-shock phases; 2) Using noise reduction techniques to prepare our data for deep learning algorithms.

2.1 Correlation

In the first phase of this study, we wanted to test whether there are correlations between EEGs segments using cross-correlation technique (explained below). Another technique that is widely used in the field of channel processing is Fourier transform (Hu and Zhang 2019). Roughly speaking, Fourier transform breaks a channel into an alternative representation that is characterized by sinus and cosines. However, using Fourier transform may result in losing an important portion of data.

Before explaining cross-correlation technique, we will explain very briefly EEG records of an ECT experiment. An EEG signal can be divided into pre-during-post shock phases.

Figure 1, shows the during shocking phase of an EEG channel after applying electroshock to an MDD patient’s scalp. The during shocking phase starts with a hyperpolarization phase, followed by repolarization and,

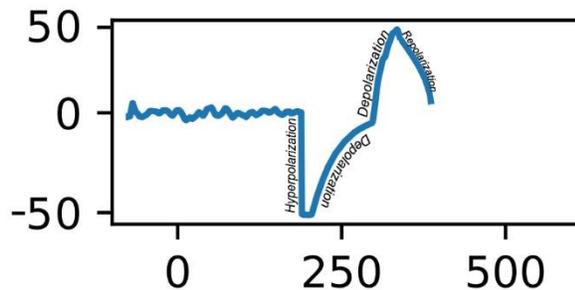


Figure 1. Hyperpolarization, Depolarization, and Repolarization phases.

then depolarization phases. In Figure 1, the x-axis corresponds to time and the y-axis corresponds to the amplitude of the EEG channel. The depolarization phase finishes by oscillation around zero on the x-axis. It is worth mentioning that the trace in Figure 1 is the average output of about one hundred million neurons¹ activities after applying an ECT shock to the patient’s brain.

Figure 2, demonstrates the last part of two complete EEG channels gathered from two patients' scalps during the electroshock procedure. The EEGs contain a patient’s pre-during-post shock phases. To distinguish the good and bad ECTs, healthcare professionals use different criterion such as the quality of the pattern of the crisis, and/or the length of the crisis and/or the smoothness of the end of the EEG channel. If the end of the channels becomes smooth as demonstrated in Figure 2.A, the ECT session is considered successful.

However, if the end of the channel is not smooth (Figure 2.B), the electroshock procedure is considered unsuccessful. In this article, for convenience, the EEG associated with a good ECT test is called a good EEG and vice versa.

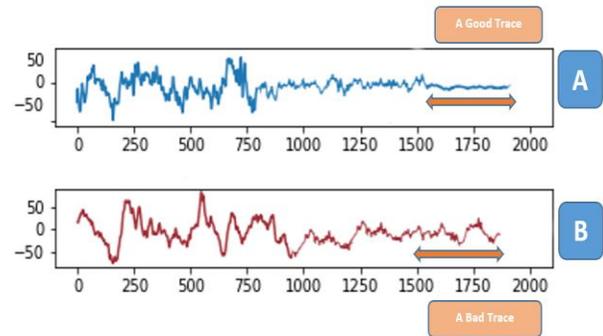


Figure2: The horizontal axis is time, and its length equals 2 seconds (2000ms). The vertical axis is amplitude. The last segment (which is underlined by an orange arrow) determines to what extent the trace is good or bad.

Postulated by one of our mental healthcare professional colleagues, the first hypothesis in our work was that there should be some type of correlation between EEGs’ Hyperpolarization, Depolarization, Repolarization phases (Figure 1), and final phase of the EEG (Figure 2, the interval between 1500-2000). In our EEG database, each EEG file contains up to 4 channels. That is, to apply the ECT, our healthcare professional used two electrodes on the frontal lobe, and two on the temporal lobe, symmetrically.

¹ The human brain has about 100 billion neurons and an area of 1,200 square centimeters. Given that the area of an electrode is about one

square centimeter. Therefore each electrode records the electrical activity of approximately 83 million neurons.

To verify our above hypothesis, we used the cross-correlation technique. The cross-correlation between two EEG channels measures the level of dependency between hyperpolarizations (HP), depolarizations (DP), and the repolarizations (RP) phases of the channels.

More specifically, If the cross-correlation of two EEG channels is 1 at a time t_0 , the EEGs are either hyperpolarized or depolarized, or repolarized. That is, the amplitudes for the two channels are equal.

Similarly, if the cross-correlation between two EEG channels is 0.8 at time t_1 , then the EEGs are hyperpolarized, depolarized, or repolarized at the t_1 and the amplitudes of their HP or RP or DP are equal to 80%.

Any negative correlation means the channels are correlated but their behavior is the opposite. That is if the HP value is increasing the RP or DP values are decreasing. Furthermore, the values around zero means EEG channels are acting independently (Dowdy, Wearden, & Chilko, 2011).

The EEGs we used in this study are gathered on a daily basis at the St-Marie hospital at Trois-Rivières (QC) by mental health professionals and then anonymized for research usage. So far, we have 290 EEGs traces. The gathering of EEGs by mental health professionals at the St-Marie hospital continues as the more EEG we have, the better results we obtain from our neural network (explained below). We applied cross-correlation technique to the above 290 EEG traces.

Every EEG has an average length of 75,000 milliseconds (75sec) in total. However, according to our hypothesis, we need to extract the HP, RP and DP parts of the EEGs. Once extracted, we apply cross-correlation technique to the aforementioned parts of the EEGs and the last part of the EEGs which contain the successful or failed ECT (Figure 2).

Because the total average of EEG files' length is 75 seconds, we need to divide them into smaller segments. In this article, we split the EEG files into two-seconds segments.

Therefore, we have $(75 \div 2 \approx 37) \times (75 \div 2 \approx 37) \times 202$ cross-correlation plots. That is, after applying cross-correlation technique to the segmented EEGs, we will obtain $202 \times 37 \times 37$ plots. As an example, Figure 3, shows the plot of two EEG segments with the length of two seconds (Figure 3. A and Figure 3. B).

Instead of 2 seconds, one can divide the EEGs into larger or smaller segments. Choosing smaller numbers than the number 2, makes cross-correlations comparison

meaningless. That is, the segment has very little information that decreases the cross-correlation technique's performance. On the other hand, choosing larger numbers make comparisons difficult as every segment contains too much information. After dividing EEG channels into segments of two seconds, we: 1) compared all segments of the good and bad EEGs; 2) extracted the HPs, RPs, DPs, and the last segment of the good and bad EEGs.

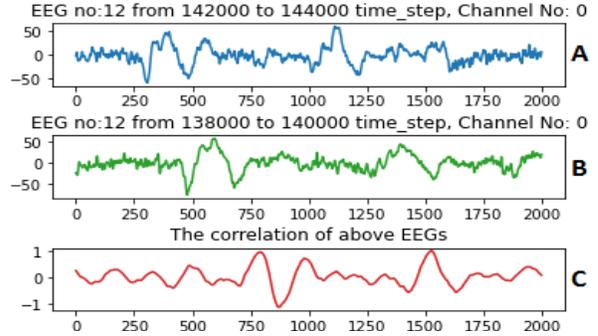


Figure 3: The red subplot (Figure 3.C) demonstrates the cross-correlation of two segments (Figure 3.A, Figure 3.B) from the twelfth EEG signal of our dataset. The horizontal axis is time (ms). In Figure 3.A and Figure 3.B, the vertical axis is EEG amplitude and in Figure 3.C is correlation value.

2.1.1 Comparing two seconds segments

In this subsection, using cross-correlation technique, we compared all the 2-seconds segments we extracted from the entire data set. One problem with the Cross-correlation technique is it compares every segment with themselves which causes redundancy in comparison.

So, in order to avoid calculating duplicate cross-correlations, we filtered the extracted segments so it only considered the unique combination of 2-second segments.

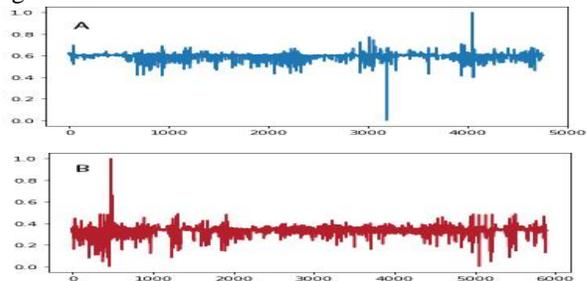


Figure 4: In Figure 4.A and B, the Y-axis shows computed averages of the cross-correlations of different segments for the first channel of the good EEGs (Figure 4.A) and bad EEGs (Figure 4.B). The X-axis shows the EEG segment pair numbers.

We then, averaged the cross-correlations of every 2-second segments for the entire data set. In our case every comparisons produced one single point which is demonstrated in Figure 4.A,B. Figure 4.A, B shows the result of the filtering process and averaging of the whole data set which resulted in more than 5000 points (each considered a point in the Figure 4).

It must be noted that according to the differences between channel size (Figure 4.A,B good EEGs with 5000 and bad EEG with 6000 length), we obtain different data points on the x-axis.

In Figure 4, the average correlation for the good EEGs is equal to 0.6 and 0.25 for the bad EEGs (maximum should be equal to 1).

In Figure 5, we see the average cross-correlation for the second channel of the good EEGs and bad EEGs which are equal to 0.6 and 0.4 respectively. We obtained similar results for the third and fourth channels. Our results show that there are some correlations in general between EEGs segments. However, we would like to remind you that so far, we have not extracted HPs, RPs and DPs from two seconds segments (see below).

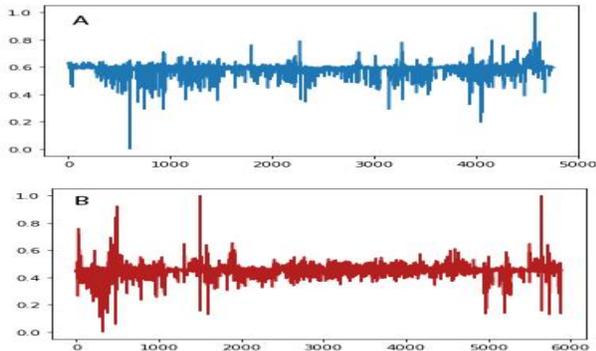


Figure 5: In Figure 5.A and B, The Y-axis shows computed averages of the cross-correlations of different segments for the second channel of the good EEGs (Figure 5.A) and bad EEGs (Figure 5.B). The X-axis shows the EEG segment pair numbers.

It is also worth mentioning that in Figures 4 and 5, at some points the spikes' values are almost one which are considered as outliers in our case. This is because the number of spikes is less than 10, while the total number of calculated cross- correlations is more than 5,000. In the next subsection, we will analyze the possible correlations of HPs, RPs, DPs with ECT outcome.

2.1.2 HPs, RPs, DPs cross-correlation with ECT results

In this subsection, we will test our mental healthcare professional colleagues' hypothesis which postulates

that there should be a logical connection between the HPs, DPs, RPs, and the ECT outcomes.

To do so, we extracted, concatenated, and averaged the HPs, RPs, and DPs segments from the entire EEGs data set.

Since, we obtained very similar results for the cross-correlation of HPs, DPs, RPs, here we will briefly explain RPs (Figure 6).

It must be noted that every EEG contains many HPs, DPs, RPs phases. We extracted and concatenated all RPs from the entire data set and obtained more than 100000 points (Figure 6). We then, calculated the cross-correlation of all RPs and the end of the good and bad EEGs.

Most of the average cross-correlation values in Figure 6 varies between -0.25 and 0.25, which is very low. However, there are some specific points that demonstrate

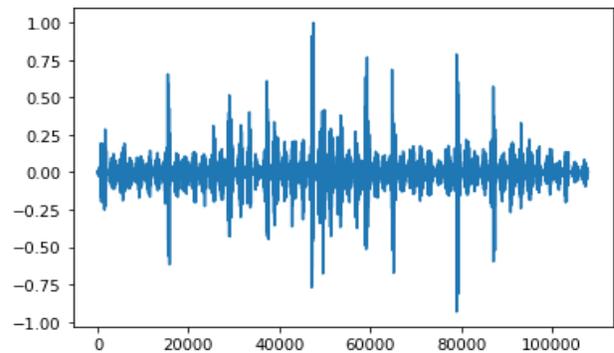


Figure 6: The Y-axis shows computed averages of the cross-correlations of RPs. The X-axis shows the EEG segment pair numbers.

good correlations. The number of these specific points are small comparing to the whole data set. Consequently, they cannot be considered as a solid indicator for the predictions of successful versus unsuccessful ECTs.

We obtained similar results for HPs and DPs. That is, HPs, RPs, DPs cannot be considered as the good predictors of the successful and unsuccessful of ECTs (see the link to the code).

Our next hypothesis was the use of Deep Learning algorithms (DLs) that uses patients' EEGs pre-shock and during the shock phase in order to predict the successful and unsuccessful ECTs.

However, EEGs contain noises which degrade substantially DLs performance. Thus, we must ideally delete or reduce the noises. For instance, patients' fast or low winking results in different EEGs (Hu and Zhang 2019). The noises directly affect DLs performance. In a preliminary experiment, we directly applied Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber 1997) to *anonymized* EEGs, without obtaining good results.

Consequently, one crucial problem when processing the EEGs gathered during ECT sessions is how to separate the useful data and noises. Therefore, before we use the DLs to predict the ECTs results, we must reduce/suppress the noises. In the next section, we explain the noise reduction method we used in this study.

2.2 Noise Reduction (NR)

In order to do Noise Reduction (NR) in our data, we used Moving Averages (MAs) or Moving Mean (MM) technique (Booth, Mount and Viers 2006). MA takes as

$$MA_n = \frac{X_{(k-1)*n} + X_{((k-1)*n+1)} + X_{((k-1)*n+2)} + \dots + X_{k*n}}{k}$$

Formula 1: The Simple Moving average formula

input a dataset and creates many subsets of it. It then, returns the average of the subsets by smoothing subsets' variations. This technique can be seen as noise reduction. Researchers use different versions of MA for noise reduction (Booth et al. 2006). We used a simple version of it (Figure 7):

Where:

- X_k = The average of input signal amplitudes in i th-period.
- n = The n th point of output signal.
- k = Length of the periods.

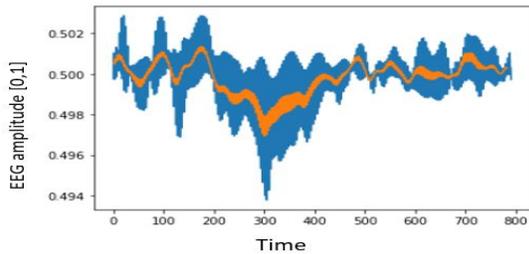


Figure 7: The blue signal is the EEG traces from our dataset and the orange signal is moving average of the blue signal. The x axis is time, and the y axis is EEG amplitude which is between (0, 1)

After applying the noise reduction (SMA), the data is ready to be fed into the Deep Learning algorithms (DLs). In the next section, we will create our DL.

3 Deep Learning algorithms

3.1 Predictor Neural Network Architecture

In this section, we will examine another hypothesis. That is, there is a strong logical relationship between the pre-shock segment of the EEGs and the success or failure of the ECT results.

To do so, we need to design and implement a hybrid architecture that is capable of detecting the temporal features of the EEGs and the relationships between different phases of the EEGs.

Our hybride DL architecture (Figure 8) uses Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber 1997), Convolutiona Neural Networks (CNN) (Sainath et al. 2020), and Multi-layer Perceptron (Gardner and Dorling 1998).

In this step, we used the same 2-second segments we used in subsection 2.1.1. Out of 290 EEG traces divided into 2-second segments, we considered 70% for training and 30% for test.

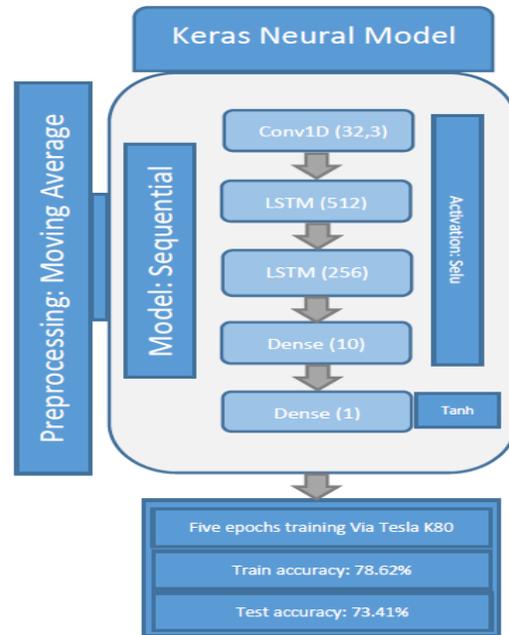


Figure 8: The Predictor CNN- Neural Architecture

In Figure 8, the denoised and averaged data (see previous section) is first fed into a one dimensional CNN (32 neurons and a kernel with size 3). It then fed into two LSTMs-*first having 512 neurons and the second 256 neurons*.

Table 1 shows different configurations for our hybrid DL architecture. We obtained the best performance (82%) using the configuration in the first line in Table 1.

| No | Model | Optimizer | Dense layers | LSTM layers | Conv Layers | Train Accuracy | Test Accuracy | Number of epochs |
|----|------------|-----------|---------------------------------------|--|--|----------------|---------------|------------------|
| 1 | Sequential | ADAM | 2 layers • Dense(10) • Dense(1) | 2 layers • LSTM(512) • LSTM(256) | 1 layer • Conv1D(32,3) | 81.71% | 82.26% | 5 |
| 2 | Sequential | ADAM | 2 layers • Dense(10) • Dense(1) | 2 layers • LSTM(1024) • LSTM(512) | 0 layer | 68.5% | 66.2% | 6 |
| 3 | Sequential | ADAM | 2 layers • Dense(10) • Dense(1) | 3 layers • LSTM(1024) • LSTM(512) • LSTM(256) | 2 layers • Conv1D(32,3) • Conv1D(64,3) | 75.25% | 72.12% | 7 |

Table 1: Different configurations of our DL architecture

4 Conclusion

Currently, mental healthcare professionals (MHP) use trial and error method for the Electroconvulsive Therapy (ECTs) sessions. Consequently, a patient may experience many ECTs before noticing some results. This is a waste of time and resources. In this study, using only EEG traces gathered from MDD patients’ scalps, we used hybrid Deep learning algorithms to predict successful and unsuccessful ECTs. This is the first attempt toward creating methodological technique for individualized ECTs.

It must be noted that we did not used expensive techniques such as MRI/ fMRI/rs-fMRI that are used in previous studies. Deep learning algorithms are capable of learning unseen patterns. Although, our data set was small, we obtained 82% precision for detecting good and bad ECTs. Thus, we believe DLs can help us to create a methodological approach to individualize ECT techniques for patients suffering from MDDs.

At this point, we demonstrated our results to other mental healthcare professionals who offered their support to this project by giving us more EEG traces gathered from MDD patients’ scalps. Having more EEG traces will improve our DLs precision and prediction capability.

One problem with our current EEG files is that they do not have the degree to which ECTs are applied to the MDD patients’ scalps. Our future work will be to change our current DL’s architecture so by merely having the MDD patients pre-shock data it can assist MHPs to what degree they need to apply ECTs in order to get successful results.

5 References

Baziotis, C., N. Pelekis & C. Doukeridis. 2017. Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis. In *Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017)*, 747-754.

Booth, E., J. Mount & J. H. Viers (2006) Hydrologic variability of the Cosumnes River floodplain. *San Francisco Estuary and Watershed Science*, 4.

Chen, H., J. D. Janizek, S. Lundberg & S.-I. Lee (2020) True to the Model or True to the Data? *arXiv preprint arXiv:2006.16234*.

Cruz-Roa, A. A., J. E. A. Ovalle, A. Madabhushi & F. A. G. Osorio. 2013. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 403-410. Springer.

Faghihi, U., S. Robert, P. Poirier & Y. Barkaoui (2020) From Association to Reasoning, an Alternative to Pearl’s Causal Reasoning. In *Proceedings of AAAI-FLAIRS 2020. North-Miami-Beach (Florida)*.

Gardner, M. W. & S. Dorling (1998) Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32, 2627-2636.

Gärtner, M., E. Ghisu, A. L. Herrera-Melendez, M. Koslowski, S. Aust, P. Asbach, C. Otte, F. Regen, I. Heuser & K. Borgwardt (2020) Using routine MRI data of depressed patients to predict individual responses to electroconvulsive therapy. *Experimental Neurology*, 113505.

Gong, J., L.-B. Cui, Y.-B. Xi, Y.-S. Zhao, X.-J. Yang, Z.-I. Xu, J.-B. Sun, P. Liu, J. Jia & P. Li (2020) Predicting response to electroconvulsive therapy combined with antipsychotics in schizophrenia using multi-parametric magnetic resonance imaging. *Schizophrenia research*, 216, 262-271.

Hochreiter, S. & J. Schmidhuber (1997) Long short-term memory. *Neural computation*, 9, 1735-1780.

Hu, L. & Z. Zhang. 2019. *EEG Signal Processing and Feature Extraction*. Springer.

Min, B., M. Kim, J. Lee, J.-I. Byun, K. Chu, K.-Y. Jung, S. K. Lee & J. S. Kwon (2020) Prediction of individual responses to electroconvulsive therapy in patients with schizophrenia: Machine learning analysis of resting-state electroencephalography. *Schizophrenia research*, 216, 147-153.

Robert, S., U. Faghihi, Y. Barkaoui & N. Ghazzali (2020) Causality in Probabilistic Fuzzy Logic and Alternative Causes as Fuzzy Duals. *ICCCI 2020*, Ngoc-Thanh Nguyen, et al.

Sainath, T. N., A. W. Senior, O. Vinyals & H. Sak. 2020. Convolutional, long short-term memory, fully connected deep neural networks. Google Patents.

Sun, H., R. Jiang, S. Qi, K. L. Narr, B. S. Wade, J. Upston, R. Espinoza, T. Jones, V. D. Calhoun & C. C. Abbott (2020) Preliminary prediction of individual response to electroconvulsive therapy using whole-brain functional magnetic resonance imaging data. *NeuroImage: Clinical*, 26, 102080.

Tsuchiyama, K., H. Nagayama, K. Yamada, K. Isogawa, S. Katsuragi & A. Kiyota (2005) Predicting efficacy of electroconvulsive therapy in major depressive disorder. *Psychiatry and clinical neurosciences*, 59, 546-550.

Van Waarde, J., H. Scholte, L. Van Oudheusden, B. Verwey, D. Denys & G. Van Wingen (2015) A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. *Molecular psychiatry*, 20, 609-614.

Yu, F., G. S. Croso, T. S. Kim, Z. Song, F. Parker, G. D. Hager, A. Reiter, S. S. Vedula, H. Ali & S. Sikder (2019) Assessment of automated identification of phases in videos of cataract surgery using machine learning and deep learning techniques. *JAMA network open*, 2, e191860-e191860.