Classifying Brain Waves of Left and Right Hand Movement Imagery with a Portable Electroencephalograph

Ruben Teunisse

5896649

Utrecht University

Bachelor Artificial Intelligence

15 EC

28-6-2019

First Supervisor

Prof. Dr. Ir. M.J.C. Eijkemans

Second Supervisor

Dr. M. van Ommen

Utrecht University

Abstract

Objective: This is a follow-up to research done by Li, Xu, Zhu (2015). This research attempts to show that a portable electroencephalograph (EEG) with its electrodes placed on the forehead is capable of classifying eye movements, but is not capable of classifying movement imagery. **Background**: Brain-Computer Interfaces (BCIs) can already enable severely disabled patients to interact with the environment (Chaudhary, Birbaumer, & Ramos-Murguialday, 2016). However, the necessary traditional EEG systems are difficult and inconvenient to use for portable BCIs. In practice, a portable EEG system is desired, but using them to classify movement imagery lacks research. **Method**: The portable EEG system Muse was used to gather data while participants followed a stimulus on a screen. The stimulus alternated between appearing left and appearing right. Participants followed the stimulus either with their eyes, or by imagining closing their hand at the same side as the stimulus. For each of these and two other tasks a Support Vector Machine and a Neural Network were trained. **Results**: Both algorithms were able to classify looking to the left versus looking to the right with an accuracy above 80%. On the contrary, neither of them was capable of classifying imagining closing the left versus the right hand. **Conclusion**: These results show that a portable EEG system is capable of classifying the direction of eye movements. In addition, and unlike the statement by Li et al. (2015), this research suggests that Muse is not capable of classifying left and right hand movement imagery. **Application**: Patients with classical locked-in syndrome can still make vertical eye movements (Bauer, Gerstenbrand, & Rumpl, 1979). In case this research' results generalize to all eye movements, there is potential for these patients for easier interaction with the external world.

Keywords: Brain-Computer Interface (BCI), electroencephalography (EEG), Muse, movement imagery, classification, Support Vector Machine (SVM), Neural Networks.

Contents

1. Introduction

1.1 Brain-Computer Interfaces

A Brain-Computer Interface (BCI) is a system one can use to communicate with the external world by solely and voluntarily changing one's mental state, e.g. imagining moving a body part or thinking about an emotion. In the upcoming years BCIs will become an important part of everybody's lives. It may even be the step towards direct brain-to-brain communication (Rao et al., 2014). Currently, BCIs can for example help in enhancing creativity (Gruzelier, 2018), or help people with locked-in syndrome to communicate (Chaudhary, Birbaumer, & Ramos-Murguialday, 2016). Until recently, a big obstacle in spreading these systems was that they were inaccessible, hard to use and non-portable.

1.2 Classifying Imaginations at Home

This thesis is a follow-up to research done by Li, Xu & Zhu (2015) in which they investigated whether they could discriminate brain waves of left and right hand movement imagery with Muse in order to play a simple game. Muse is an affordable, accessible and portable electroencephalograph. Despite these benefits, the cost cannot be ignored; less accuracy due to more noise and fewer and less precisely placed electrodes. Li et al. (2015) collected data from participants while they imagined moving their left and right hand while rotating their eyeballs in the corresponding direction. Using a machine learning algorithm called support vector machines, they reached an accuracy of 95%, making the subjects able to flexibly control a plane in a one dimensional game.

To improve the model's applicability and intuitive control, this experiment will have three major differences. Firstly, the data will be gathered while the participants are looking at the game. The object will move automatically and perfectly. When the object moves to the left the data being collected will be labeled "left" and likewise for the other direction. This will bring the training environment closer to the test environment, suggesting that the model will perform better when applied. Secondly, rather than explicitly asking the participants to rotate their eyeballs without having anything to look at, they will be asked to look at the place they are headed for in the game. In addition, the subject will not be asked to imagine dribbling or waving a badminton racket as in the research by Li et al. (2015), but rather to imagine closing the specified hand. Both should contribute to a more intuitive way of controlling the game. Thirdly, Li et al. (2015) did not at all show whether imagining the mentioned hand motions had any influence on the data. This research will include a control group to test whether or not the imagination part had any positive influence on top of the eyeball rotation.

To classify the data, a Support Vector Machine and a Artificial Neural Network will be trained. Both have successfully been used in the study mentioned above or in similar studies (Mutasim, Tipu, Raihanul Bashar & Amin, 2017; Bird et al, 2018; Bird, Ekart, Buckingham & Faria, 2019).

To conclude, this experiment will show whether the chosen machine learning algorithms are useful for classifying 'left' and 'right' data from a portable EEG. Despite the extra restrictions on how the data is gathered, the classifier will be considered useful with a minimum of 80% accuracy, just as in the research of Li et al. (2015). Since the experiments only slightly differ in what the participants are looking at and thinking about, it is expected that this research will show similar results. Nevertheless, it is also hypothesized that the imagination of hand movement cannot be extracted from the data for reasons explained in the section 'Electrode Placement'.

1.3 Place in AI

This study connects to Artificial Intelligence in two ways. Firstly, in order to create non-human intelligence it is useful to study the human brain. By measuring the brain's electrical activity under certain conditions, knowledge is acquired about how the brain functions. Secondly, already existing AI models will be applied to analyze the data. Comparing the different machine learning techniques will show which algorithm with which parameter settings works best with this kind of data.

1.4 Thesis Structure

In the next chapter, the process from brain activity to useful data will be explained. This will bring to light some of the possibilities and limitations of this study. Next, a traditional experimental research structure will be followed, containing sections for methods, results, the models, discussion and conclusions.

2. Theoretical Background

2.1 From Neurons to Game Control

It is quite a process to get from brain activity inside a skull to a desired movement on a computer screen. This chapter will give a overview of the entire process.
In the next section a technique called electroencephalography will be discussed, showing how to obtain data out of the skull. Unfortunately, this data is not yet useful. In the section thereafter different techniques to preprocess the data will be explained. Preprocessing will result in data from which for example the two classes “left” and “right” can be learned through machine learning techniques discussed in the last section.

2.2 Electroencephalography

In order to get information directly from the brain, a variety of techniques have been developed in the last few decades. All of them rely on differences in activation of the relevant brain regions. These activation differences can for example be measured by measuring the electrical activity (EEG), the magnetic field (MEG), or the blood flow (MRI).
To put it simplistically, an electroencephalograph, or EEG for short, is a voltmeter used to measure the brain’s electrical potentials, in this case with the electrodes placed along the scalp.

To understand the possibilities and limitations of EEG, the brain’s electrical activity has to be examined more closely. The brain internally communicates through neurons. These nerve cells transmit information through a combination of chemical and electrical signals. A neuron receives a signal when neurotransmitters bind its neurotransmitter receptor proteins. This signal is electrically conducted to the other end of the neuron where it releases neurotransmitters as the presynaptic neuron. These neurotransmitters flow through the synapse, the gap in the range of nanometers between neurons, and bind the neurotransmitter receptors of the next postsynaptic neuron. Important to EEG is that neurons do not necessarily ‘fire’, i.e. transfer the chemical signal into an electrical one that travels to the other end of the neuron. Whether a neuron fires depends on whether its threshold is reached. This threshold, or action potential, is the potential difference between the inside and the outside of the neuron from which the neuron accepts the signal and fires. These potential differences of individual neurons can not be detected on the scalp, but they can when thousands to neurons line up and fire together.

2.2.1 Electrode Placement

Traditionally, hand movement is measured with electrodes placed on the locations C3 and C4 (e.g. LaFleur et al., 2013) according to the International 10-20 System (*figure 2.1* shows them in the red circles). This seems logical, since these locations are located right on top of the motor cortex; part of this brain area becomes more active when a certain body part movement is being planned or visualized (Morash, Bai, Furlani, Lin, & Hallett, 2008). Nevertheless, Li et al. (2015) claim to pick up this signal at F7 and F8. The Muse headband used in this research has electrodes placed on TP9, TP10, AF7 and AF8 (blue in *figure 2.1*) for average sized heads. For smaller sized heads, the frontal electrodes tend to be placed more in the direction of F7 and F8. The reference electrode is placed at Fpz (shown in green). Its function will be explained in the next section.

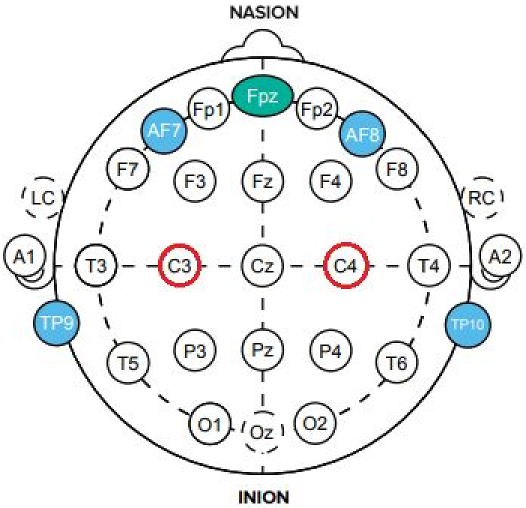

Figure 2.1; Electrode placements according to the International 10-20 system. Original image retrieved from http://developer.choosemuse.com/hardware-firmware/hardware-specifications

2.3 Preprocessing

Although the image above suggests that electrical voltage is measured at the electrodes' locations, in reality this is not possible. Voltage equals potential *difference* and is thus always measured relative to somewhere else. These 'channels' can for example be between two active electrodes (the electrodes already mentioned), forming chains of channels. In the case of Muse, the potential differences are measured between the active electrodes and the reference electrode.

2.3.1 Eye movement

How neurons create potential differences is already mentioned, but eye movements work differently.

Firstly, the eye acts as a dipole, since the retina is slightly negatively charged in comparison to the cochlea (Berg & Scherg, 1991). Whenever one looks to the right, frontal electrodes to the right will measure a higher voltage, both compared to before the eye movement as well as compared to the left electrode which will measure an even lower voltage. Secondly, the extraocular muscles are electrically controlled as well. Their contractions generally result in a smaller spike, just before the actual eye movement (Björk & Kugelberg, 1953).

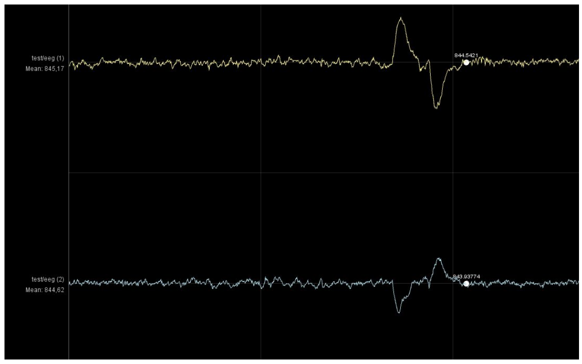

Figure 2.2; Image is a screenshot of Muse Lab after a participant looked first to the left and then to the right. The upper line shows the raw voltage at AF7 and the lower line shows the raw voltage at AF8. The contractions of the extraocular muscles are not visible.

2.3.2 Analog to Digital

The electrodes read an analog signal which is in the range of microvolts (Teplan, 2012). To be able to digitize this signal accurately, it first has to be amplified. After the amplification, an A/D converter brings it into digital form.

2.3.3 Noise reduction

The measured electrical voltages over time create a wave. These waves already contain more information than static potential differences. However, they are quite rough and contain a lot of irrelevant oscillations that seem to be nothing but noise. To remove this noise, three filters will be discussed.

Firstly, the low-pass filter, also known as the high-frequency filter, filters out all frequencies which are uninterestingly high, while leaving the relevant frequencies unaffected. In the case of EEG, the cut-off frequency is commonly chosen around 100 Hz, with a transition band starting around 70 Hz. Of course, these values depend on the pursued signal. Filtering out these very high frequencies result in decreased muscle artefacts (Muthukumaraswamy, 2013).

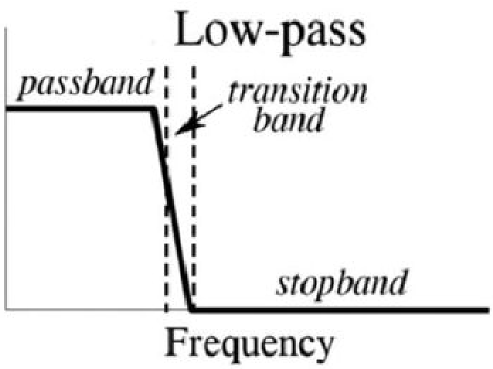

Figure 2.3; Obtained from research done by Obeid et al. (2017). On the vertical axis the amplification of the amplitude is shown.

Secondly, a high-pass filter, also known as a low-frequency filter, is applied. This filter works analogously to the high-frequency filter. It allows high frequencies to pass, while filtering out very low frequencies, e.g. below 1 Hz. The filter reduces the effect of for example eye blinks.

Lastly, the notch filter eliminates electronic interference coming from any electronic device that is plugged into a wall. In Europe, alternating current is produced at 50 Hz. This frequency is clearly visible in the EEG signal and needs to be filtered out. The notch filter functions similarly to the filters mentioned before, but with a very narrow stopband, immediately around 50 Hz. All other frequencies, except for a small transition band, stay unaffected.

2.3.4 Frequency Bands

Even with less noise, the electric voltage waves seem to not contain much information about a mental state. This raw voltage wave can be split into a discrete number of different frequencies, using Fast Fourier Transform (FFT). How FFT works is too complex to dive into here. From the output of the algorithm, the amplitude and phase of each frequency can be derived. This results in many waves. To clarify what FFT did, adding up all the resulting waves (taking the sum of the amplitude), would result in the original raw voltage wave.

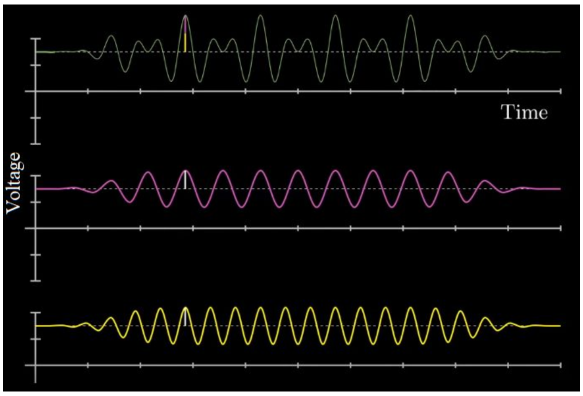

Figure 2.4; The 440 Hz yellow wave and the 294 Hz purple wave add up to the complex wave at the top. In the case of EEG, the process starts with the complex wave and splits it into many more than just two waves. For example, one for every 0.5 Hz in the range of 0 Hz to 100 Hz, resulting in 200 different waves, all with their own static frequency.

By convention, all of these waves fall into one of five frequency *bands.*

Gamma	30-44 Hz
Beta	13-30 Hz
Alpha	7.5-13 Hz
Theta	4-8 Hz
Delta	1-4 Hz

Table 2.1; There definitely exists some variation in these numbers between different EEG systems as well as between different researchers, but these are the ones used by Muse. Overlap between consecutive bands is not uncommon, neither does it result in complications.

The slow waves (roughly below 7 Hz) correspond with low arousal and usually indicate that one is sleeping or in deep meditation. These are not relevant for this study. Fast waves correspond with high arousal and are divided in alpha, beta and gamma.

Alpha waves show a high amplitude when one's eyes are closed or one is very relaxed (Valipour, Shaligram, & Kulkarni, 2013). Alpha activity is also strongly related to attention (Klimesch, 2010).

Beta waves are dominant during normal state of wakefulness with open eyes (Teplan, 2012). They tend to be even more dominant when one is processing information or is thinking analytically. Beta activity is also related to planning and executing movements of body parts (Takahashi, Saleh, Penn, & Hatsopoulos, 2011).

Correlations with the activity of gamma waves are more controversial. Some researchers argue that gamma is simply a byproduct of brain activity, while others propose it directly contributes to brain functioning (Jia & Kohn, 2011). Gamma would be modulated by working memory, attention and sensory input and actively facilitate communication between different brain regions. Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell (2008), associate gamma with rapid eye movements.

It stands out that whatever frequency band is being look at, the eyes have a major influence on the signal.

2.3.5 Absolute Band Powers

In practice, considering tens of waves in order to find relations between a frequency band and brain behaviour is very inconvenient. Rather, a single measure of frequency band activity is desired. For this purpose, the power spectral density (PSD) is calculated. The PSD describes the power of each wave at a given moment. This power is not to be confused with the electric power measured in watt, but refers to the square of the amplitude.

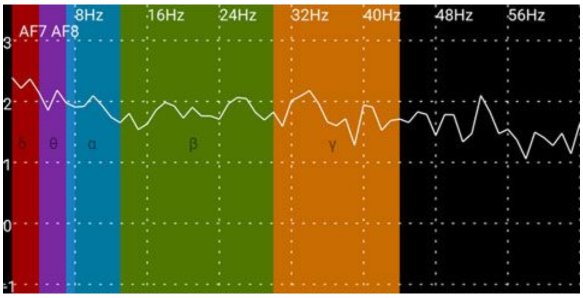

Figure 2.5; a screenshot of Muse Monitor displaying the PSD, also known as the power spectrum or periodogram. The different frequency bands are coloured.

Still, there is no single number per frequency band, but now it can easily be calculated. One option for such a number is called the absolute band power which theoretically equals to the integral of the PSD between its frequency boundaries. In *figure 2.6*, the absolute band power of gamma would be the orange area under the graph. Since there is no closed-form formula to integrate this area, it has to be approximated. This approximation represents the total power of the gamma frequency band, also called the absolute gamma value. This value can of course be calculated for every frequency band.

Muse calculates the absolute band power differently, which is one of the reasons why it is hard to compare EEG data between different studies. Muse takes the logarithm of the sum of all PSD values available within the frequency range.

2.3.6 Band Relations

Often when trying to find relations between more complex mental states and brain activity, it turns out useful to consider relative brain waves.

One option is to compare one absolute band power to the sum of absolute powers of all frequency bands. This is called the relative power of the frequency band. Relative powers are commonly used within EEG research. For example, relative beta may be correlated with symptoms of ADHD (Kropotov et al., 2005) and all relative frequency bands differ in Parkinson dementia patients compared to controls (Soikkeli, Partanen, Soininen, Pääkkönen, & Riekkinen Sr., 1991).

The second option is to calculate the ratio between one frequency band and one other, rather than all others. Ratios are commonly used in EEG research as well. For example, the alpha/theta ratio is correlated with cognitive creativity (Gruzelier, Thompson, Redding, Brandt, & Steffert, 2014) and beta/theta ratio with response inhibition and attentional control (Putman, van Peer, Maimari, & van der Werff, 2010).

2.3.7 Epochs

Even after noise reduction and frequency band division, the data is still coarse and needs improvement. Since EEG has a very high temporal resolution, many data points per second can be gathered (512 in the case of Muse). When the pursued signal spreads over multiple data points, there is no need to consider each datapoint individually. Moreover, considering each datapoint individually adds noise due to very short, but irrelevant fluctuations in the signal. Rather, it is useful to combine many data points into what are called 'epochs'. These epochs can vary from 100ms to a few seconds, depending on the duration of the signal that is aimed for.

2.4 Machine Learning

In order to find relations between EEG data and mental states, machine learning can be very useful. Even though a lot of knowledge about brain regions and their functions is available, it is difficult to find which parts of the signal come from which brain region. Machine learning can help search through multiple features of the signal and relate them to the reported mental state. What these features can be will first be discussed, followed by explanations of two machine learning algorithms: Support Vector Machines and Artificial Neural Networks.

2.4.1 Feature Selection

As stated before, the raw EEG signal does not reveal much about what is going on inside, but the signal does contain a lot of characteristics which seem to be easier linked to mental states. These include the absolute band powers, the relative band powers and the ratios between absolute band powers. Including all possible ratios, this already adds up to 35 features. Fortunately, a selection can be made to use in the model based on previous research. Sometimes, adding features will keep on improving the model, but that is only guaranteed if the features do not overlap and do contain information about the mental state that is looked for. For example, if brain region A is responsible for cognitive function F but only sends out alpha waves, the absolute beta feature does not contain any valuable information and will therefore at best not influence the model, but more realistically worsen it.

2.4.2 Support Vector Machine

For a long time the Support Vector Machine (SVM) was the most popular classification method. It was introduced by Boser, Guyon and Vapnik in 1992. The SVM is effective in high dimensional spaces and memory efficient. Currently, it is in competition with deep neural networks. The SVM has the advantage of performing better with smaller data sets and can be interpreted more easily. To understand the mechanics of the SVM, it is easiest to start with its hard-margin version.

A hard-margin SVM can be applied to classify data that is linearly separable in its feature space. When d features are used to train the model, the feature space has d dimensions. Each datapoint is a vector with a value for each of the d features and is located in the feature space according to these values. Each datapoint belongs to one of two classes. The SVM tries to partition the space in two groups in such a way that the decision boundary has a maximum margin without any data points and that the two groups perfectly match the two classes. In order to find this decision boundary, the data points closest to the opposing class, called the support vectors, are determined. The decision boundary is drawn as a hyperplane, a *d*-dimensional generalization of a line, right in the middle of these support vectors.

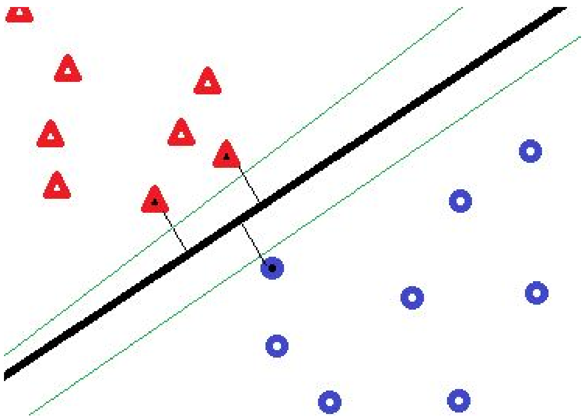

Figure 2.6; The green lines split the data perfectly as well, but only the black one does so with the maximum margin. The three data points filled with black are the support vectors. Their distances to the decision boundary are all of the same length.

In practice, the data is almost never linearly separable. This does not mean that the SVM is useless. To overcome this problem, overlap with the margin can be allowed, by using the soft-margin version of the SVM. Another option is to use a kernel function to transform the feature space to higher dimensions.

Since the soft-margin SVM allows for misclassification, simply choosing the extreme data points as the support vectors will not work. An in sample error can be calculated and minimized to find the optimal support vectors. This in sample error function increases with the number of misclassified data points, as well as with the distance of the misclassified data points to the decision boundary.

One of the reasons why SVMs are so popular is because of their resistance to overfitting. This is one of the benefits of linear models in general. However, if the nature of the studied phenomenon is not linear, than neither should the classification algorithm. To seemingly improve the SVM's flexibility, kernels can be used. These kernel functions can transform the priorly poorly separable data to data that can be separated by the linear SVM. When the data is transformed back to the original feature space, it seems like the SVM has separated the data non-linearly.

2.4.3 Artificial Neural Network

The Artificial Neural Network (ANN) may currently be the best known machine learning algorithm, thanks to its many successes within Artificial Intelligence. Especially its version with multiple hidden layers, called deep learning, has enormously improved the state-of-the-art speech recognition, visual object recognition and many other domains such as drug discovery and genomics (LeCun, Bengio, & Hinton, 2015).

A neural network consists of an input layer, an output layer and zero or more hidden layers. Each layer is a set of nodes, called neurons, which are connected to all nodes of the previous and the next layer. These connections are called weights and determine the influence of the node preceding the connection on the node posterior to the connection. Each input neuron represents one of the model's features. In the case of binary classification, the output layer consists of precisely one neuron. Each neuron has a value, which is calculated as the sum of all preceding neurons' values multiplied by the corresponding weight. As regards the input neurons, their values are determined by the feature values. Training the network means establishing the weights, which can for example be done using the backpropagation algorithm. When the trained network is fed with a new data point, the value returned by the output neuron can be rounded and used as a classification.

2.4.4 Cross Validation

Generating a model that perfectly predicts the classes of the data points it is trained on is easy and contains no value. What matters is the model's performance on unseen data points. Therefore, a data set is split into a training and a test set. If the model performs great on the training set, but poorly on the test set (overfitting), the model does not generalize well and should be given less flexibility. In the case of the SVM this can for example be done by not using the kernel trick, or in the case of an ANN by removing hidden neurons.

Especially with small data sets, it may be the case that the test set, which is usually between 10 and 30 percent of the entire data set, is not representative for the entire data set. This may result in unreliably high or low performance scores. To minimize this effect, k-fold cross validation can be applied. This time, the entire data set is not split into a fixed train and test set, but rather in k subsets of the same size. The model is trained and tested k times, each time on a different subset as test set, with the remaining sets as training data.

3. Methods

3.1 Participants

Data has been collected from 6 adults (four male and two female). 5The subjects' ages ranged from 20 to 53. None of them had any significant eye or brain abnormalities. Two of them had experience with EEG neurofeedback training, while four of the subjects had not been exposed to BCIs before.

3.2 Materials

3.2.1 Hardware

Muse refers to the EEG system created by the company Interaxon 2014 used in this study. Muse has been chosen for many reasons. If there would not have been a constraint on the EEG system being used, this study would not have been interesting. Numerous researchers have shown they could classify visualizations of left and right hand movement with traditional EEG systems (e.g. Morash et al., 2008). Muse adds the benefits of portability, accessibility, affordability and the ease with which it can be used without the help of someone that is not wearing the headband. To summarize, the results of research done with muse can be applied 'at home', or even outdoors. Of course, these benefits are not unique to Muse. Muse is simply chosen as one of them.
Inside the muse headband are a gyroscope and an accelerometer, capturing the head orientation and movement.

As a laptop, a Dell XPS 15 was used. It could easily run all programs without noticeable delay.

3.2.2 Software

Muse sends out each package of information individually as soon as a certain value is determined for every electrode. Such a package contains for example one raw voltage each electrode (about 256 times per second). Already inside the headband the absolute and relative values of every frequency band is calculated. These ten different packages are transmitted as well (about 10 times per second). Likewise, muse sends out information about connection quality, gyroscope and accelerometer values, battery status and blink and jaw clench detection.

To use this data in a C# program, three pieces of software are considered. The smartphone applications Muse Monitor and Muse Direct work fine, but slow down the connection and increase the risk of losing connection, as the data first has to be sent to a phone before it can be transmitted to the computer. On the other hand, the windows application Muse Direct allows Muse to directly connect with the laptop via bluetooth. For that reason, the windows application has proven to be the most suitable.

Muse Direct forwards the data via Open Sound Control (OSC). A C# program has been written to pick and process this signal. Subsequently, datapoints are formed by combining all information packages within the duration of one epoch. Multiple values of the same measurement (e.g. the raw voltage measured by electrode 1) are averaged to reduce noise and get one value per measurement per epoch. Lastly, the program to train and test the Support Vector Machines were written in C# as well.

All C# programs were written in Microsoft Visual Studio 2015. The data has been stored in SQL databases as well as in CSV files.
The neural network was generated in R, using the library neuralnet and Rstudio as the editor.

Muse Lab is the software used to visualize the data gathered by Muse. It has been used to make sure all electrodes had a good connection and to show and explain participants what was being measured.

3.3 Procedure

Figure 3.1; The one-dimensional snake game.

The subjects were shown a screen similar to the one in *figure 3.1*. The participants did not see all the lines at the same time, but each new line was printed on top of the previous one, creating the sense of a moving 'Y'. The 'Y' represents the snake in a one-dimensional snake game. The 'O' represents the food. Whenever the snake catches the food, the score increases with one. The 'tail' of the snake does not lengthen, since that would inevitably create an unplayable game. Note that during the experiment the participant had no influence on the game at all, since there is no model to determine the participants' desired input yet. Rather, the snake moves automatically and perfectly; it always runs in the direction of the food. New food always emerges at the opposite side, near the end of the level. This ensures consistent and large space intervals between the old and the new food, resulting in consistently large eye movements.

In a test round (with one of the participants) before the actual experiment, different combinations of level width, snake speed and epoch duration were tried in order to find reasonable values.

The actual experiment consisted of five rounds, with the first being a test round. The other four rounds were randomly ordered. Before the first round, Muse was properly fitted to the participant's head and the connection quality was checked in Muse Lab. Briefly, the data being gathered was explained. If needed, the electrodes were wetted to improve signal quality.

Next, the program was set up and the laptop was placed either on the participant's lab, or on the table in front of him/her. In both situations the screen was located at a comfortable distance.

Each round took 60 seconds; long enough to gather sufficient data, while not too long to bore the participant. Between each round there was a break in which the instructions for the next round were given. During each round, the participants were allowed to blink whenever they wanted to.

During the first round the elements on the screen were explained. The subject could ask questions and look anywhere on the screen. Data was gathered to function as a baseline.

In the 'eyes' round, subjects were instructed to look at the 'O' on the screen. They were explicitly asked to not follow the 'Y', but to look at the new 'O' on the other side of the screen immediately after the old 'O' had disappeared.

Before the 'visualization' round, the participants were asked to place their hands either on their legs or on the table, with their palms up. Whenever the 'Y' would start moving to the right they should imagine closing their right hand (making a fist) and imagine opening it up before the 'Y' would reach the 'O'. Likewise for when the 'Y' moves to the left, but then with their left hand. The participants were explicitly asked to try not to move their hands, but solely imagine closing them. To avoid overlap with the eyes round, the subjects were asked to look at the centre of the screen, such that they could still notice the direction of the ‘Y’, but without making eye movements.

The ‘real movement’ round was very similar to the visualization round, except for asking the participants to *imagine* closing the corresponding hand, they were asked to *actually* close it. Detecting real movement with a portable EEG is not the main aim of this study, but it may provide extra useful knowledge.

The ‘both’ round combined the eyes round with the visualization round. The participants were asked to both follow the ‘O’ with their eyes, as well as imagine closing and opening the corresponding hand at the same time.

After the last round, the participants were shown their own data in scatterplots.

3.4 Measurements

As soon as the ‘Y’ on the screen changed direction, the following data would be stored for the duration of one epoch: the raw voltages and all absolute and relative values for all frequency bands for all channels. In addition, the connection quality of each electrode was saved, in order to afterwards be able to detect noise due to bad connection.

4. Experimental Results

4.1 Test phase

During the test phase multiple level widths, snake speeds and epoch duration were tested. These variables turned out to have major influence on the results. The level had to be wide enough to make sure the eyeball rotations were big enough to observe in the data. As regards the snake speed, a tradeoff exists between data points per minute on the one hand and enough time to clearly visualize closing a hand on the other. A well chosen epoch duration turned out to be crucial for gathering the right data. Since only eye movement, rather than statically staring in one direction, impacts the eeg signal, only one epoch of data per change of direction was gathered. The epoch started as soon as the snake caught the food. An epoch duration chosen too short would label data already 'LEFT', while the participant did not have had enough time to react, since the average reaction time for a visual stimulus is 150ms (Fischer & Ramsperger, 1984). If, on the other hand, the epoch duration is chosen too long, the stimulus would be smoothed out.

4.1.1 Level width

The plot in *figure 4.1* clearly shows the need for a wide enough window. When the width increases, the two classes move away from each other. The sweet spot was found around 161 character spots of the console window, equal to 23 cm on the used computer screen.

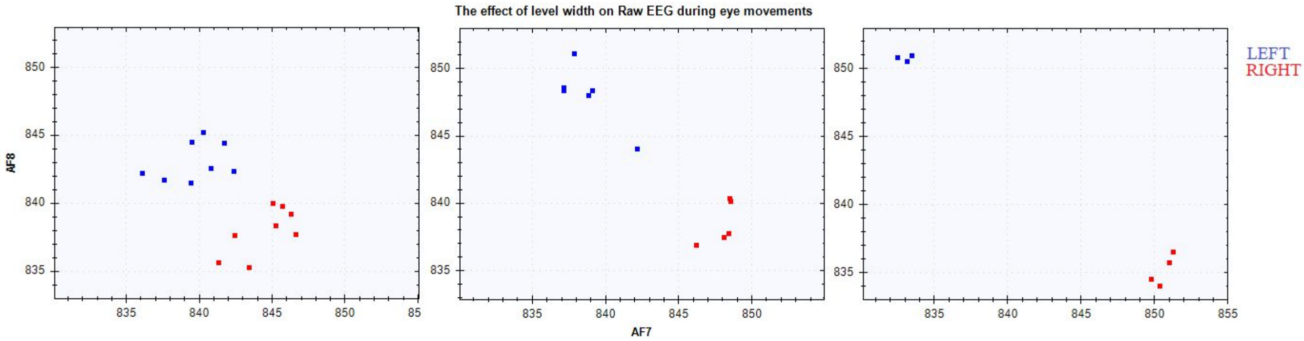

Figure 4.1. The plots show the improvement of data quality from widening the screen. From left to right the plots show a level width of 81, 121 and 161. The game speed (1 step per 10 milliseconds) and the epoch duration (500 milliseconds) remained unchanged. In all three rounds, data was gathered for 30 seconds. Given the constant game speed, the fact that 1 data point was gathered per change of direction and the widening of the screen, inevitably the total number of gathered data points decreased. This is an argument against a screen as wide as possible and in favour of a high game speed.

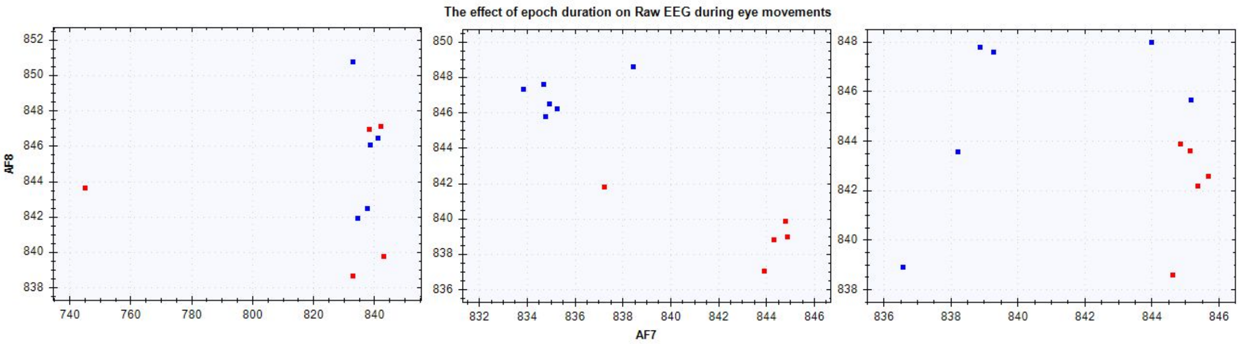

Figure 4.2; The plots show the effect of different epoch durations on the raw EEG data during eye movements. From left to right the chosen epoch durations are 100 ms, 500 ms and 900 ms, all with a level width of 121 and a game speed of one step per 10 ms. Noise increases as the chance for eye blinks within the epoch interval increases.

4.1.2 Epoch duration

Figure 4.2 shows the tradeoff between the probability of including the participant's reaction on the stimulus and the magnitude with which the effect has been smoothed out.

4.1.3 Game speed

In case of the eyes round, the speed could be quite high, since it is easy to quickly alternate between looking left and looking right. On the contrary, after the visualization round participants reported that it took time to clearly imagine feeling and seeing their hands close. This was anticipated; since it was reported in previous research as well (Li et al., 2015). The two rounds could not differ in game speed, then the two could not fairly be compared. This resulted in a balance of one step per 10 milliseconds, i.e. 1.61 seconds from one end to the other. The raw eeg plot did not reveal any easily visible significant patterns.

4.2 Experiment rounds

Due to bad electrode connection the data of two participants had to be left out. Furthermore, the results of the both round from one participant was lost. A different participant did all rounds twice. Only the outliers which were clearly due to bad connection were removed (the connection quality was logged).

4.2.1 Test round

As expected, data from the test round did not show any interesting patterns.

4.2.2 Eyes round

The test phase showed that eye movements are clearly visible in the raw eeg data. Even though the data of all participants showed some overlap between left and right in the raw eeg data, there exists a clear segregation between the majority of the two classes.

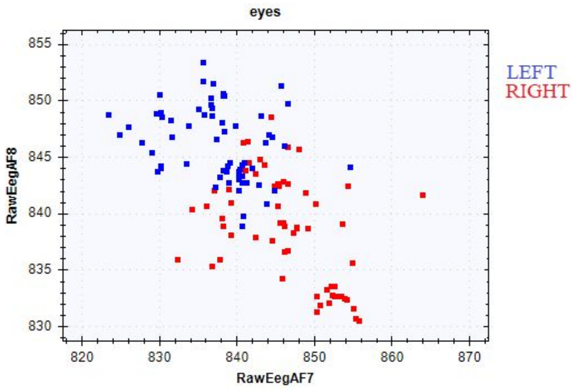

Figure 4.3; a scatterplot of the results from the eyes round. The values on both axes are in microvolts.

The data quality seems to be significantly better for individual participants than for all participants combined. *Figure 4.4* shows the data from the interesting frequency bands for both an individual participant as well as for all participants combined. Although the plots show the clearest ones, this pattern stood out for all participants; if their raw eeg data was clearly separable, than some distinction between 'left' and 'right' was visible in alpha absolute, gamma absolute and in gamma relative.

This difference between all participants individually and all participants combined, could be explained by natural differences in participants, or very small differences in connection quality which both could result in voltage differences bigger than the observed differences between 'left' and 'right' within participants. This may be solved by normalizing the results of each subject before aggregating them.

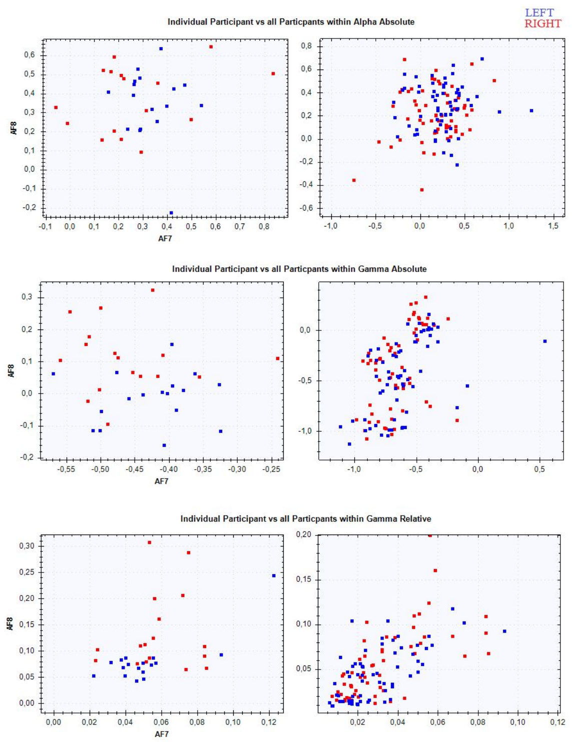

Figure 4.4; The eyes round data of an individual participant is shown on the left. The data of all participants combined, from the same round, and from the corresponding feature band, is shown on the right. The absolute and relative bands are measured in Bels and are thus on a log scale.

The other absolute and relative frequency bands did not show the same pattern consistently throughout all participants, or they showed no clear distinction between 'left' and 'right' at all. Which is interesting, since Li et al. (2015), used precisely absolute gamma and relative gamma as features for their model.

4.2.3 Visualization

Neither the absolute, nor the relative frequency bands, nor the raw EEG plots showed any clear distinction between 'left' and 'right' for the visualization round.

Even considering participants individually showed no support for any feature as a valuable separator.

4.2.4 Real movement

Visual inspection of the combined data from the real movement round showed no support for any feature as a valuable separator, nor did the data from individual participants.

4.2.5 Both

The both round combines the eyes round with the visualization round. Therefore, it is expected that the data will be separable with respect to its raw voltage, just as the data from the eyes round is. The plot in *figure 4.5* supports this, however, it is definitely not as clear.

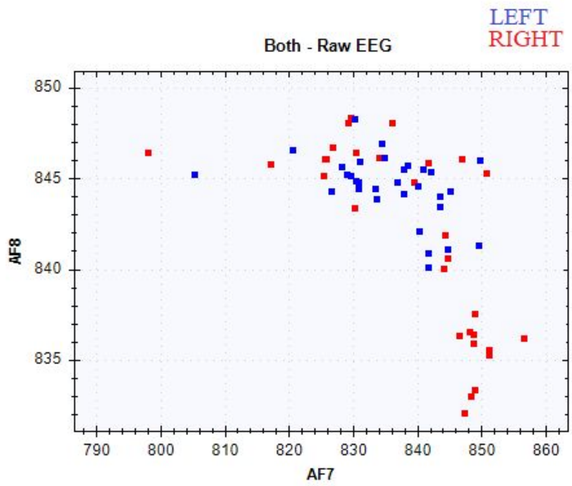

Figure 4.5 shows the raw EEG of all subjects obtained during the both round. The values are in microvolts.

5. Model

The previous section showed the results of the visual inspection of the collected data. Visual inspection is very limited, since viewing more than 3 dimensions at the same time is difficult. A machine learning model can be trained to consider multiple dimensions at once. Moreover, such a model can predict the class of new data points and quantify the separability of a data set.

Before any model was trained, the order of the data within each round was randomized. The values of all features were scaled to values between -1 and 1. Lastly, every reported accuracy is the result of 10-fold cross validation.

5.1 Feature sets

For each algorithm, a model was trained on each of the following feature sets.

- The raw voltages at AF7 and AF8
- The absolute and relative alpha values at AF7 and AF8
- The absolute and relative beta values at AF7 and AF8
- The absolute and relative gamma values at AF7 and AF8
- The raw voltages together with the relative gamma values at AF7 and AF8
- All features: The raw voltages, together with the absolute and relative delta, theta, alpha, beta and gamma values, all at TP9, AF7, AF8 and TP10, as well as the head orientation and head movement.

All individual fast frequency bands are included, since in the research of Li et al. (2015) all of them turned out to be correlated with what in this experiment is measured during the both round.

Ratios between frequency bands are excluded, since no previous research showed support for the value of a certain ratio for this application.

5.2 Support Vector Machine

A SVM was trained on five different values for the cost parameter C. This parameter determines the penalty for a misclassified datapoint and therefore the size of the margin. A high C value forces the model to fit very well on the training data, but risks low scores on test data due to overfitting. On the other hand, a low C value risks not training enough. A balance is required. 8 different C values were tried. Only the raw voltage feature set in combination with the data from the eyes and the both round is reported, since none of the other feature sets were significantly more effective than guessing at any round. The table below shows a clear parabolic pattern (especially in the both column). Although it is not entirely clear what the best C value is, 10^{-5} was chosen as the one to fit the final model with.

	eyes	**both**
10^{3}	71.2	49.6
10^{0}	74.1	50.9
10^{-3}	75.4	50.2
10^{-4}	82.0	67.3
10^{-5}	81.7	69.5
10^{-6}	82.9	67.4
10^{-7}	82.4	63.8
10^{-8}	66.2	53.9

Table 5.1; The accuracies in percentages for different C values for the eyes and the both round.

Next, a linear rather than a squared hinge-loss function was tried. The accuracies increased to 83.5% for the eyes round and 69.7% for the both round. Still, the SVM did not appear to be of any value on other rounds or feature sets.

5.3 Neural Network

A small neural network was trained as well. Adding more than one hidden layer worsened its performance. The optimal number of neurons depended on the number of features, but no attempted configuration was able to reasonably predict the right class for the visualization round, nor for the real movement round.
In contrast to the Support Vector Machine, adding features seemed beneficial. A neural network with 29 neurons in the hidden layer scored an accuracy of 80% on the both round with all (50) features. For the small feature sets, a neural network with only 5 hidden neurons turned out to be optimal, but still worse than the big neural network which could predict the data of the eyes round 87% accurate, combining the raw EEG data with the relative gamma data.

6. General Discussion & Conclusion

6.1 The Experiment

The experimental results already showed a clear difference between the eye movements and the movement imagery. The first showed a clear distinction between left and right in the raw EEG data and some separation potential in individual frequency bands. The latter did not show any sign of difference between left and right in any of the considered features. Both results were expected, given the placement of the electrodes. However, there are multiple other possible causes.

Parameter values were selected during the test phase based on data from eye movements and on assumptions with respect to movement imagery, since movement imagery was not visible in the test phase data. If these assumptions are wrong, it would not be a surprise that the final data did not show a difference between left and right hand movement imagery either. A few examples will be given.

Firstly, only one epoch per stimulus was stored. Which is logical as regards eye movements, since only the movement spikes the voltage, rather than the staring in a certain direction. However, the visualization of hand movement was not meant as a single action, but as one that continued until the next stimulus. Saving multiple epochs per stimulus may had improved the data quality.

Secondly, the time between two consecutive stimuli may be too short to fully catch the effect of the visualization. Actually, Li et al. (2015), found an optimal time frame from 3 to 7 seconds after the participant started visualizing.

Thirdly, reaction time was only partly taken into consideration. The epoch duration was made sure to be long enough to include the actual reaction, but that also meant that the epoch was partly filled with pre-reaction data. A better option would have been to determine the reaction time and only start the epoch (or epochs) just before the reaction initiated.

A possible cause of a different nature was the amount of data. Unfortunately, about one third of the data was lost due to bad electrode connection. On top of that, the 35 minutes of data that was collected in total (compared to 53 minutes in the research done by Li et al., 2015), was spread over 5 rounds, rather than 1 round.

6.2 Models

Combining the SVM and the ANN, many different parameter settings, on many different feature sets from multiple subsets of the data were tried. This decreases the significance of finding an interesting performance score. On the other hand, it increases the significance of not having found an acceptable model for the visualization data.

The testround proved valuable as a baseline. Both the SVM as well as the ANN could reach an accuracy of 60% on the testround data for certain feature sets. With that in mind, accuracies below 60% on the data of other rounds were considered insignificant.

Three more things about the model results stood out. Firstly, the ANN performed better than the SVM, especially with the big feature set. Secondly, neither of the two models was able to significantly predict real hand movement. Thirdly, both models performed better on the eyes round than on the both round. Which is surprising, since Li et al. (2015) argued that the eye movement would amplify the effect of the visualization, not that the visualization would weaken the effect of the eye movements. *Figure 6.1* summarizes the results.

	SVM	ANN
eyes	83.5	87.0
both	69.7	80.0
visualization	≤ 60	≤ 60
real movement	≤ 60	≤ 60
testround	60	60

Figure 6.1; The results of the SVM and the ANN on each of the five rounds. Values are in accuracy percentages.

6.3 Conclusions

Both algorithms were able to classify looking to the left versus looking to the right with an accuracy above 80%. On the contrary, neither of them was capable of classifying imagining closing the left versus the right hand. These results show that a portable EEG system is capable of classifying the direction of eye movements. In addition, and unlike the statement by Li et al. (2015), this research suggests that Muse is not capable of classifying left and right hand movement imagery.

Further research could reveal whether other portable EEG systems, for example with different electrode placement, are able to classify left and right hand movement imagery. If this turns out to be the case, a big leap towards portable BCIs is taken. Portable BCIs can have a big impact on everyday life. From fast brain to brain communication, to improved care for the disableds. For example, patients with classical locked-in syndrome cannot move at all, but they can imagine moving. Portable BCIs open up the potential for these patients to easily interact with the external world.

8. References

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of neurology, 221*(2), 77-91.

Berg, P., & Scherg, M. (1991). Dipole models of eye movements and blinks. *Electroencephalography and clinical Neurophysiology, 79*(1), 36-44.

Bird, J. J., Ekart, A., Buckingham, C. D., & Faria, D. R. (2019). Mental emotional sentiment classification with an eeg-based brain-machine interface. In *Proceedings of the International Conference on Digital Image and Signal Processing (DISP'19)*.

Bird, J. J., Manso, L. J., Ribiero, E. P., Ekart, A., & Faria, D. R. (2018). A study on mental state classification using eeg-based brain-machine interface. In *9th International Conference on Intelligent Systems, IEEE*.

Björk, A., & Kugelberg, E. (1953). The electrical activity of the muscles of the eye and eyelids in various positions and during movement. *Electroencephalography and clinical neurophysiology, 5*(4), 595-602.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152). ACM

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research, 57*(1), 191-195..

Chaudhary, U., Birbaumer, N., & Ramos-Murguialday, A. (2016). Brain–computer interfaces for communication and rehabilitation. Nature Reviews Neurology, 12(9), 513.

Gruzelier, J. H. (2018). Enhancing Creativity with Neurofeedback in the Performing Arts: Actors, Musicians, Dancers. In Creativity in Theatre (pp. 223-245). Springer, Cham.

Gruzelier, J. H., Thompson, T., Redding, E., Brandt, R., & Steffert, T. (2014). Application of alpha/theta neurofeedback and heart rate variability training to young contemporary dancers: State anxiety and creativity. *International Journal of Psychophysiology, 93*(1), 105-111.

Jia, X., & Kohn, A. (2011). Gamma rhythms in the brain. *PLoS biology, 9*(4), e1001045.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences, 16*(12), 606-617.

Kropotov, J. D., Grin-Yatsenko, V. A., Ponomarev, V. A., Chutko, L. S., Yakovenko, E. A., & Nikishena, I. S. (2005). ERPs correlates of EEG relative beta training in ADHD children. *International Journal of Psychophysiology, 55*(1), 23-34.

LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., & He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface. Journal of neural engineering, 10(4), 046003.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning. nature, 521*(7553), 436.

Li, Z., Xu, J., & Zhu, T. (2015). Recognition of brain waves of left and right hand movement imagery with portable electroencephalographs. *arXiv preprint arXiv:1509.08257*.

Morash, V., Bai, O., Furlani, S., Lin, P., & Hallett, M. (2008). Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries. Clinical neurophysiology, 119(11), 2570-2578.

Mutasim, A. K., Tipu, R. S., Bashar, M. R., & Amin, M. A. (2017, November). Video Category Classification Using Wireless EEG. In *International Conference on Brain Informatics* (pp. 39-48). Springer, Cham.

Muthukumaraswamy, S. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Frontiers in human neuroscience, 7*, 138.

Obeid, H., Khettab, H., Marais, L., Hallab, M., Laurent, S., & Boutouyrie, P. (2017). Evaluation of arterial stiffness by finger–toe pulse wave velocity: optimization of signal processing and clinical validation. *Journal of hypertension, 35*(8), 1618-1625.

Putman, P., van Peer, J., Maimari, I., & van der Werff, S. (2010). EEG theta/beta ratio in relation to fear-modulated response-inhibition, attentional control, and affective traits. *Biological psychology, 83*(2), 73-78.

Rao, R. P., Stocco, A., Bryan, M., Sarma, D., Youngquist, T. M., Wu, J., & Prat, C. S. (2014). A direct brain-to-brain interface in humans. *PloS one, 9*(11), e111332.

Soikkeli, R., Partanen, J., Soininen, H., Pääkkönen, A., & Riekkinen Sr, P. (1991). Slowing of EEG in Parkinson's disease. *Electroencephalography and clinical neurophysiology, 79*(3), 159-165.

Takahashi, K., Saleh, M., Penn, R. D., & Hatsopoulos, N. (2011). Propagating waves in human motor cortex. *Frontiers in human neuroscience, 5*, 40.

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review, 2*(2), 1-11.

Valipour, S., Shaligram, A. D., & Kulkarni, G. R. (2013). Spectral analysis of EEG signal for detection of alpha rhythm with open and closed eyes. *International Journal of Engineering and Innovative Technology (IJEIT), 3*(6), 1-4.

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron, 58*(3), 429-441.