UNIVERSITY OF OSLO

Master's thesis

Localization and identification of Neural Sources from simulated EEG Signals

Kamilla Ida Julie Sulebakk

Biological and Medical Physics
60 ECTS study points

Department of Physics
Faculty of Mathematics and Natural Sciences

Spring 2023

Kamilla Ida Julie Sulebakk

Localization and identification of Neural Sources from simulated EEG Signals

Acknowledgements

Massive thank-yous to my supervisor Gaute Einevoll and my co-supervisor Torbjørn Ness.

Contents

Chapter 1

Introduction

Electroencephalography (EEG) is a method for recording electric potentials stemming from neural activity at the surface of the human head, and it has important scientific and clinical applications. An important issue in EEG signal analysis is so-called source localization where the goal is to localize the source generators, that is, the neural populations that are generating specific EEG signal components. An important example is the localization of the seizure onset zone in EEG recordings from patients with epilepsy. A drawback of EEG signals is however that they tend to be difficult to link to the exact neural activity that is generating the signals.

Source localization from EEG signals has been extensively investigated during the last decades, and a large variety of different methods have been developed. Source localization is very technically challenging: because the number of EEG electrodes is far lower than the number of neural populations that can potentially be contributing to the EEG signal, the problem is mathematically under-constrained, and additional constraints on the number of neural populations and their locations must therefore be introduced to obtain a unique solution.

For the purpose of analyzing EEG signals, the neural sources are treated as equivalent current dipoles. This is because the electric potentials stemming from the neural activity of a population of neurons will tend to look like the potential from a current dipole when recorded at a sufficiently large distance, as in EEG recordings. Source localization is therefore typically considered completed when the location of the current dipoles has been obtained. However, an exciting possibility is to try to go one step further and identify the type of neural activity that caused a localized current dipole. For example, the type of synaptic input (excitatory or inhibitory) to a population of neurons, and the location of the synaptic input (apical or basal) will result in different current dipoles (Ness et al., 2022). It has also been speculated that dendritic calcium spikes can be detected from EEG signals, which could lead to exciting new possibilities for studying learning mechanisms in the human brain (Suzuki & Larkum, 2017). Identifying different types of neural activity from EEG signals would however require knowledge of how different types of neural activity are reflected in EEG signals. Tools for calculating EEG signals from biophysically detailed neural simulations have however recently been developed, and are available through the software LFPy 2.0 (Hagen et al., 2018; Næss et al., 2021). This allows for simulations of different types of neural activity and the resulting EEG signals, opening up for a more thorough investigation of the link between EEG signals and the underlying neural activity.

The past decade has seen a rapid increase in the availability and sophistication of machine learning techniques based on artificial neural networks, like Convolutional Neural Networks (CNNs). These methods have also been applied to EEG source localization with promising results. However, it has not been investigated if CNNs can also identify the neural origin of EEG signals, in addition to localizing neural sources. In this Master's thesis, the aim will be to investigate the possibility of using CNNs to not only localize current dipoles but also identify the neural origin of different types of neural activity, based on simulated data of different types of neural activity and the ensuing EEG signal.

1.1 Motivation

1.2 Goal and Objectives

1.3 Structure of the Thesis

Chapter 2

Background

Neurobiology is the study of the nervous system, including the structure, function, and development of neurons and neural circuits. The physics of the neuron is an important component of neurobiology, as it involves understanding the mechanisms by which neurons generate and transmit electrical signals. The basic unit of the nervous system is the neuron, which is capable of producing and transmitting electrical signals, or action potentials, across its membrane. These electrical signals are generated by the flow of charged ions into and out of the neuron, and are essential for communication between neurons and the transmission of information throughout the nervous system.

One technique for studying the electrical activity of the brain is electroencephalography (EEG), which measures the voltage fluctuations resulting from the electrical activity of neurons. EEG is a non-invasive technique that involves placing electrodes on the scalp, and has been used to study a wide range of cognitive and neural processes, including perception, attention, and memory. One of the challenges of interpreting EEG signals is the "inverse problem," which involves determining the location and nature of the underlying sources of electrical activity in the brain.

One approach to solving the inverse problem is source localization, which involves estimating the location and strength of the electrical sources in the brain that are responsible for the measured EEG signals. Source localization is a challenging problem due to the complexity of the brain and the fact that EEG signals are affected by a range of factors, including the conductivity of the scalp and the position and orientation of the electrodes. However, there are a number of techniques and algorithms that have been developed to address these challenges, including dipole modeling, distributed source modeling, and beamforming (Hämäläinen et al., 1993; Grech et al., 2008).

Overall, the physics of the neuron, EEG, and source localization are all important components of neurobiology that have contributed to our understanding of the nervous system and its functioning. By combining knowledge of the physical principles of neural signaling with advanced analytical techniques, researchers are able to gain valuable insights into the underlying neural processes that give rise to behavior and cognition.

2.1 Introduction to Neuroscience

2.2 Electroencephalograpy

2.3 The Inverse Problem

Chapter 3

Electroencephalograpy

Some words about what the background chapter will include.

3.1 Neuroscience and Electroencephalography

Electroencephalography (EEG) is a non-invasively technique for studying electrical potentials in the human brain. The technique was developed almost a century ago making it among the oldest methods for examining the brain's activity. Today, EEG remains one of the most important techniques being used to study electrical activity in the brain, with important applications in both neuroscientific and clinical research (Nunez Srinivasan 2006, Lopes da Silva 2013, Biasiucci et al. 2019, Ilmoniemi Sarvas 2019).

EEG signal is believed to originate from large numbers of synaptic inputs to populations of geometrically aligned pyramidal neurons (Nunez Srinivasan 2006, Pesaran et al. 2018).

Electroencephalography (EEG) is one of the most important techniques for studying cognition and disease in the brain non invasively [**?**]. EEG is a method used to measure brain waves. In practice, this involves electrodes consisting of small metal disks connected to the surface of the scalp. The electrodes detect electrical charges that result from activity of the brain cells. In cases where certain areas of the brain come out as more active than others, it might indicate abnormalities, in which can be signs of disease. In other words, the EEG technique can be used to evaluate multiple types of brain disorders, such as lesions of the brain, Alzheimer's disease, epilepsy or brain tumors [**?**].

An illustration of the typical EEG measurement setup is depicted in Figure 3.1.

EEG signals are generated from synaptic inputs to cells in the cortex. Synaptic inputs are electrical (or chemical) signals that are being transmitted from one neuron to another, causing changes in the membrane potential of the neurons. In other words, neurons are specialized to pass signals, and figures/EEG_electrodes.png

Figure 3.1: Illustration of the EEG method [?].

synapses are the structures that make this transmission possible [?].

Imagining a small part of the cortex, all of these cells will have dendrites pointing upwards in the same direction (lets say the z-direction). Due to rotational symmetry around the z-axis, the contributions in the x- and y-direction will cancel. This is illustrated in Figure 3.2. What we see is that the extracellular potential is configured in all sorts of weird ways (A-C) when there is only one synaptic input, while the extracellular potential reminds more of a dipole when we have multiple synaptic inputs (D-F). We can therefore argue that the total contribution to the extracellular potential can be modelled as a dipole in the z -direction for the case where we have multiple synaptic inputs. Hence, for each dipole moment in our simulations, we assume multiple synaptic inputs, and make sure to rotate the positions of the dipoles such that it is orientated along the depth of the cortex.

3.2 Currend Dipoles Approximation

EEG signals arise from cortical neural activity and are typically described in terms of current dipoles [?].

An optimal model of EEG signals would have consisted of multiple dipole moments. However, as such a model is complicated and computationally expensive, we will in this project only introduce one single dipole approximation $\mathbf{p}(t)$ for each multicompartmental neuron simulation. In this context, by multicompartmental modelling we refer to the widely used models within neuroscience, which acurately manufactures electical properties of single neurons. The sigle-dipole approximation might sound like a substantial simplification of the real biophysical properties, nevertheless it actually turns out to give a realistic modelling of EEG signals, when handling the single dipole moment, as an abnormality in the brain. We will be thinking of the abnormality as an epileptic seizures or a tumor in the brain, which among normal activity in the brain would have stuck out. The single-dipole approximation is implemented by summing up the multiple current dipole moments, $\mathbf{p}(\mathrm{t}) = \sum_{k=1}^{M} \mathbf{p}_k(t) = \sum_{k=1}^{M} I_k^{axial}(t)\mathbf{d}_k$, where I^{axial} is the current flowing along the neurite with the distance vector $\mathbf{d}_k$ and M denotes the number of axial currents [?]. The data set we will be using in this project will consist of measures of different EEG signals at a given time from 1000 patients. This means that we for each patient pick a random location (at $t = 0$) for the single current dipole.

3.3 The New York Head Model

The signals received from EEG are known to originate from cortical neural activity, which are often described by using current dipoles [?]. It is therefor reasonable to implement current dipoles in the brain for when generating a fig_chosen_dipoles.pdf

Figure 3.2: Extracellular potential from lonely synaptic input (A-C) and extracellular potential from multiple synaptic inputs (D-F).

biophysical modeling of EEG signals. The brain model used in this thesis is called the New York Head model [?], and is based on high-resolution anatomical MRI-data from 152 adult heads. The model utilizes the software tool LFPy, which is a Python module for calculation of extracellular potentials from multicompartment neuron models [?]. This model takes into account that electrical potentials are effected by the geometries and conductivities of various parts of the head [?].

The cortex matrix consists of 74382 points, which refer to the number of possible positions of the dipole moment in the cortex. When generating our data set, we will for each sample randomly pick the position of the dipole moment, such that one sample corresponds to one patient. In our head model we are considering 231 electrodes uniformly distributed across the cortex, meaning that each EEG sample will consist of this many signals for each time step. However, we are not interested in the time evolution of the signals as this does not affect nor say anything about the position of the dipole moment, and we therefor simply pick out the EEG signals for when $t = 0$ (note that the choice of time step could have been randomly picked). Our final design matrix will then consist of 1000 rows, corresponding to each patient, and 231 columns also refereed to as features, representing the signal of each electrode. The final output we are trying to predict is then the one dimensional vector with length 1000, where each element consists of the x-, y- and z- position of the dipole moment. An example of how the input EEG signals may look like is given in appendix A, where we also have marked the dipole moment with a yellow star.

3.4 The Inverse Problem and Source Localization

Chapter 4

Methods

4.1 New York Head Model

The New York Head Model is a computer model of the human head used to simulate the electrical activity of the brain. It was created by the Electrical Geodesics Incorporated (EGI) in 2004, and is based on the anatomical and electrical characteristics of the head of a typical adult human.

The model consists of a three-dimensional (3D) representation of the head and brain, with detailed information on the geometry and electrical properties of the different tissues and structures within the head. The model includes the scalp, skull, cerebrospinal fluid, gray matter, and white matter. The electrical properties of each of these tissues, such as conductivity and permittivity, are also included in the model.

The New York Head Model is used in research and clinical applications to understand the electrical activity of the brain and diagnose neurological disorders. It is particularly useful for studying the electrical activity of the brain during various cognitive and motor tasks, as well as during seizures and other abnormal brain activity.

The model is also used in the field of non-invasive brain stimulation, where it helps to guide the placement of electrodes or magnetic coils to target specific areas of the brain for therapeutic or diagnostic purposes. By using the model to predict the electrical activity of the brain, researchers and clinicians can optimize the placement of the electrodes or magnetic coils to achieve the desired effect.

Overall, the New York Head Model is an important tool for understanding the electrical activity of the brain and has numerous applications in research and clinical settings.

4.2 Machine Learning and Neural Networks

Computational neuroscience is a field that aims to understand the principles underlying information processing in the brain using mathematical and computational tools. The inverse problem in EEG, which involves estimating the location and strength of electrical sources in the brain based on measurements of electrical activity on the scalp, is a key challenge in computational neuroscience. Machine learning techniques, including feedforward neural networks, have been used to address this problem by learning to map the measured EEG signals to estimates of the underlying electrical sources in the brain.

Source localization using machine learning techniques has shown promise for improving the accuracy and efficiency of EEG analysis, and has been applied to a variety of cognitive and clinical applications. For example, machine learning-based source localization has been used to study the neural mechanisms underlying attention, memory, and perception (Wu et al., 2018; Lopes da Silva et al., 2019), as well as to diagnose and monitor neurological disorders such as epilepsy (Safieddine et al., 2019; Shah et al., 2020). These applications demonstrate the potential of machine learning and computational neuroscience to enhance our understanding of the brain and improve clinical outcomes.

Machine learning is a field of computer science that involves using algorithms and statistical models to enable computers to learn from data without being explicitly programmed. One popular type of machine learning algorithm is the feedforward neural network, which is a type of artificial neural network that is often used for tasks such as linear regression. In a feedforward neural network, data is passed through a series of layers of interconnected nodes, or "neurons," which perform mathematical operations to transform the data.

Linear regression is a common machine learning task that involves predicting a continuous quantity, such as the price of a house or the temperature of a city, based on a set of input features. In a feedforward neural network, linear regression can be accomplished by using a single neuron in the output layer of the network that computes a weighted sum of the input features and applies an activation function to produce the predicted output value. The weights on the input features are learned by the network during the training process, which involves adjusting the weights to minimize the difference between the predicted output values and the actual output values in the training data.

Overall, feedforward neural networks are a powerful machine learning tool that can be used to solve a wide range of problems, including linear regression. By adjusting the weights and biases of the neurons in the network during the training process, neural networks can learn to make accurate predictions based on input data, making them a valuable tool for a variety of applications.

Chapter 5

Results

As mentioned in chapter 1, an important topic in EEG signal analysis is the inverse problem of going from measured EEG signals to localized equivalent current dipoles, so-called source localization. In this chapter we will be training and presenting two different(?) neural networks used to localize single dipole sources in the human cortex. Section ... and ... deal with training a simple feed forward neural network and presenting its results.

5.1 The dataset

The cortex matrix of the New York Head Model (NYHM) consists of 74382 points, which refer to the number of possible positions for localization of a dipole source. When training the FFNN we will be using a data set consisting of simulated EEG signals corresponding to dipol sources with randomly selected positions within the cortex matrix. The final data set consists of 10 000 rows, where each row corresponds to one sample, or let us say - one patient. Within the data set we have 231 columns, also referred to as features, representing the dipole measure at every EEG electrode. Thus, we are left with a design matrix with size 10 000 x 231.

In order to train the network faster, one commonly split the data set into mini-batches, which is also done here. When splitting the data such a way, the weights of connection between neurons are updated after each propagation, making the network converge considerable faster. There might be Through trial and error we landed on a batch size equal to 30.

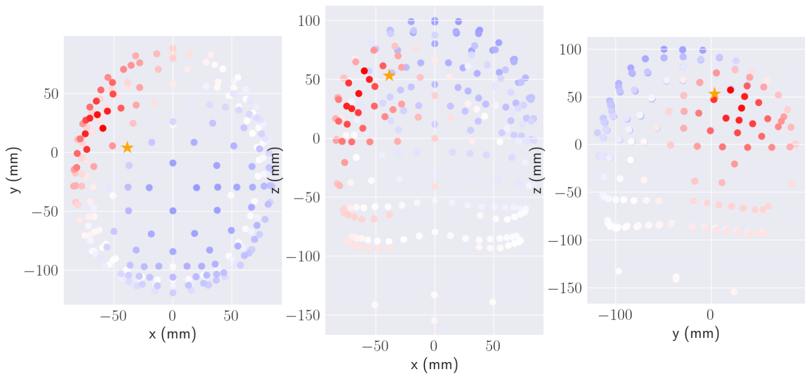

Figure 5.1: Example 1 dipole.

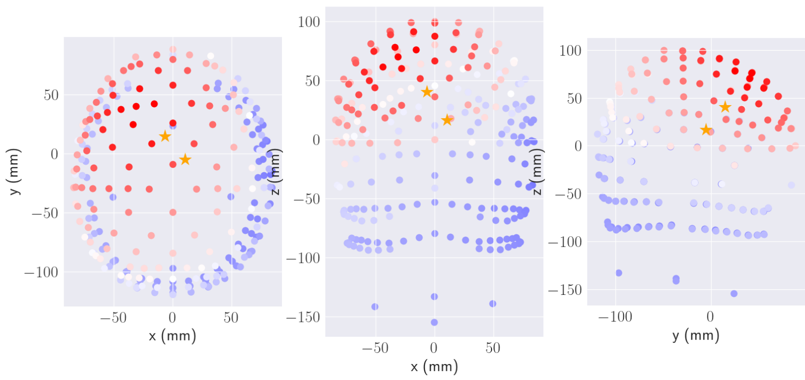

Figure 5.2: Example 2 dipoles.

5.2 Easy example

5.3 Feed-Forward Neural Network Approach for localizing single dipole sources

The feedforward neural network (FFNN) was one of the first artificial neural network to be adopted and is yet today an important algorithm used in machine learning. The feed forward neural network is the simplest form of neural network, as information is only processed forward, from the input nodes, through the hidden nodes and to the output nodes.

The FFNN that are trained to solve the inverse problem of ours has an input layer of 231 neurons, corresponding to the M = 231 electrode measures of the potentials. The input layer is followed by three hidden layers with 120, 84 and 16 hidden neurons, respectively. The final output layer holds the predicted x-, y- and z- position of the desired dipole source. For the neurons of the input layers we use the linear activation function ReLu, while for the neurons of the hidden and output layers, we chose the much used hyperbolic tangent activation function.

Cost function

5.3.1 Training, testing and evaluation

In order to make an ANN that generalizes well to new data we split our data into training and testing sets. Randomly selecting 80 percent of the rows in the full dataset, we put this into a separate one and call it our training set. The remaining 20 percent is put into the test set. In practice, the training data set consists of pairs of an input vector with EEG signals and the corresponding output vector, where the answer key is the x-, y- and z coordinate of the dipole source. The neural network is then feed with the training data and produces an estimation of the localization of the dipole. The estimation is found by the network through optimizing the parameters β minimizing the cost function, or said in other words, through finding parameters for the function that produces the smallest outcomes, meaning the smallest errors. The result provided by the network is then compared with the target, for each input vector in the training data. Adjustment of parameters...

When the network is fully trained, we have a final model fit on the training data set. Feeding the network with the test data set, we can assess the performance of the network. The predictions of the fully trained network can now be compared to the holdout data's true values to determine the model's accuracy.

In figure **??** we have provided the bias-variance trade-off for when using Tanh as activation function. We notice that error of the model is approaching 0 and that the variance between the two curves decreases for an increasing number of epochs.

5.4 Convolution Neural Network Approach for localizing single dipole sources

Some results for the prediction of location for single current dipoles.

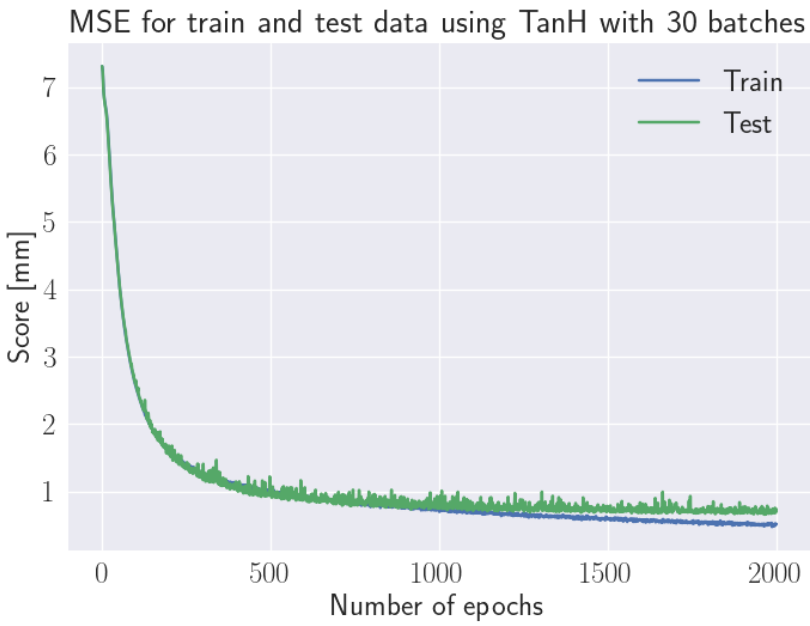

Figure 5.3: The validation accuracy for simple Feed Forward Neural Network with 10 000 samples with tanh activation function.

5.5 Region of Active Correlated Current Dipoles

Some results for the prediction of the size and location of current dipole populations.

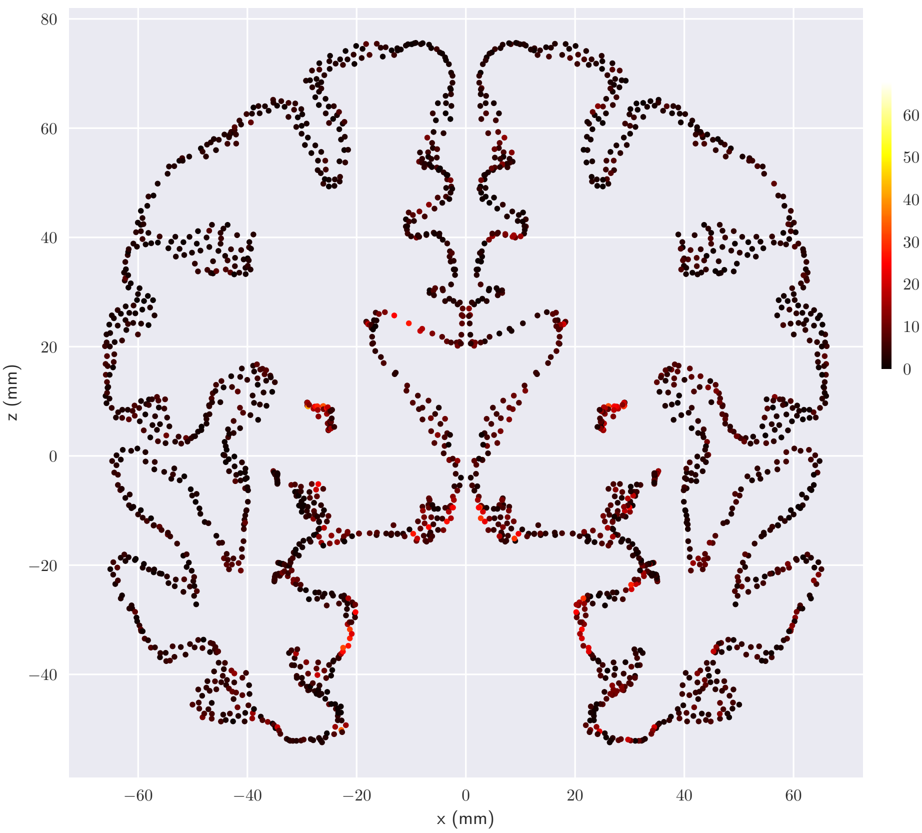

Figure 5.4: The mean squared error of the location accuracy (NN) at different dipole locations in the y cross section.

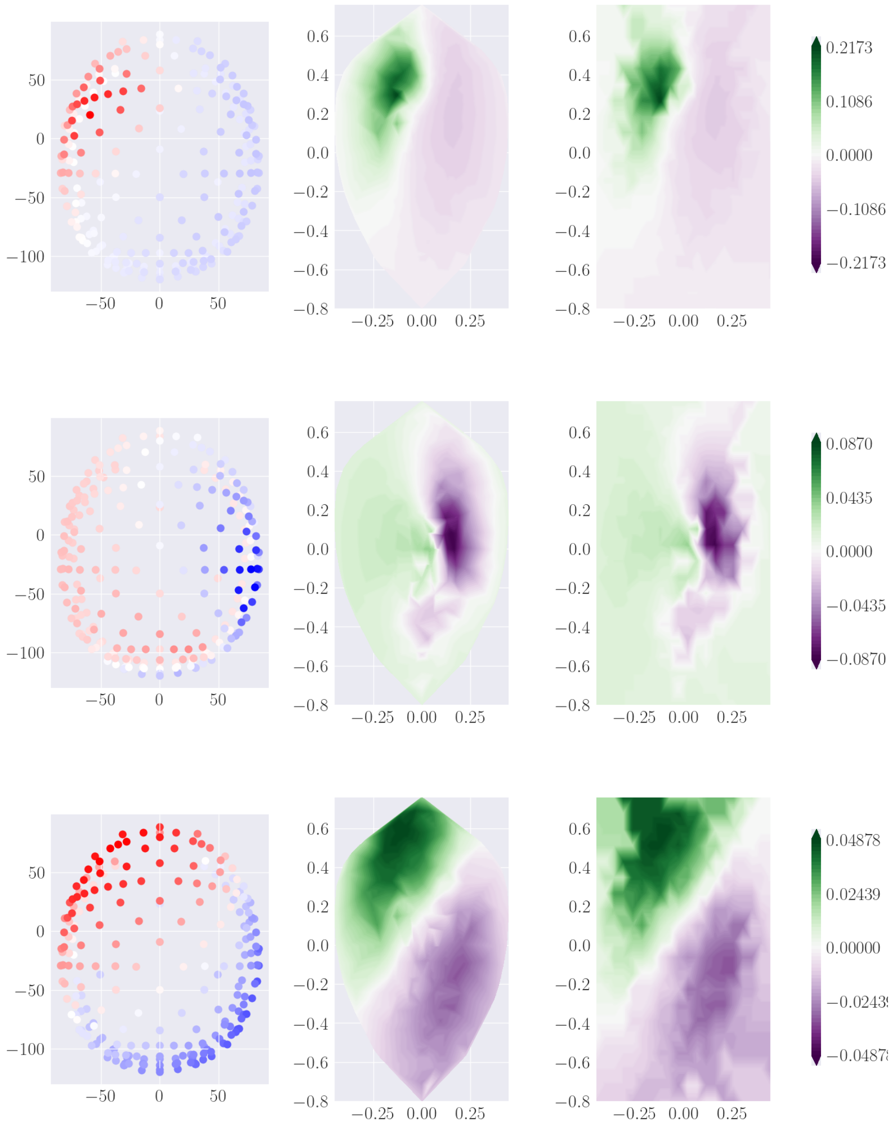

Figure 5.5:
Right: EEG measure for 3 different samples measured in μV.
Middle and Left: Illustration of the interpolation of the EEG data into two-dimensional matrix.

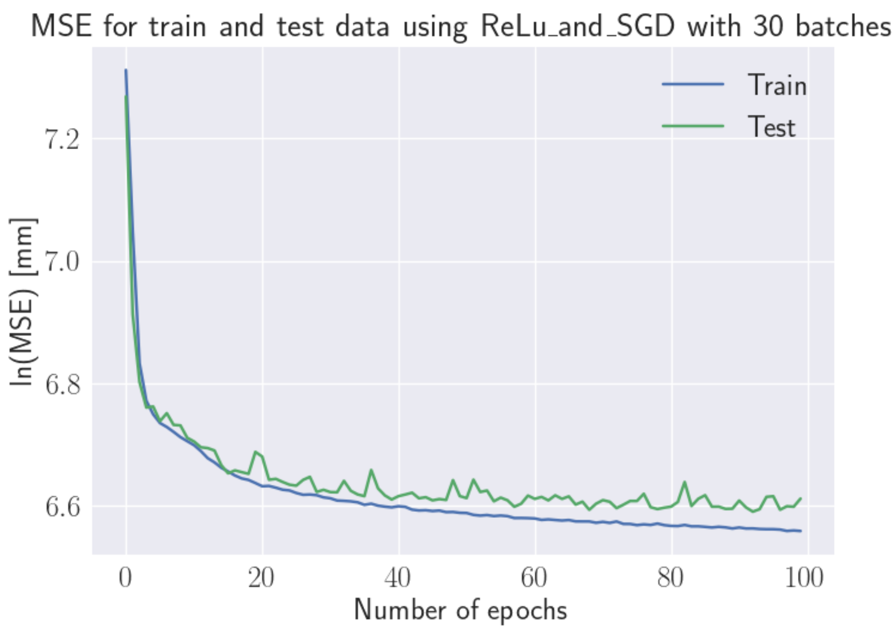

Figure 5.6: The validation accuracy for Convolutional Neural Network with 10 000 samples (20x20 matrix) with ReLU activation function.

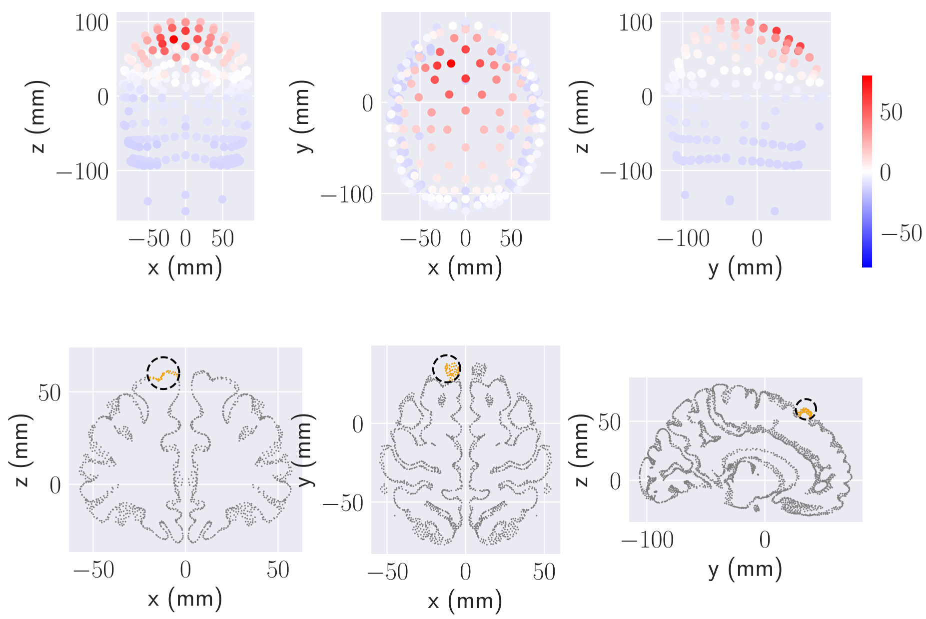

Figure 5.7: New text.

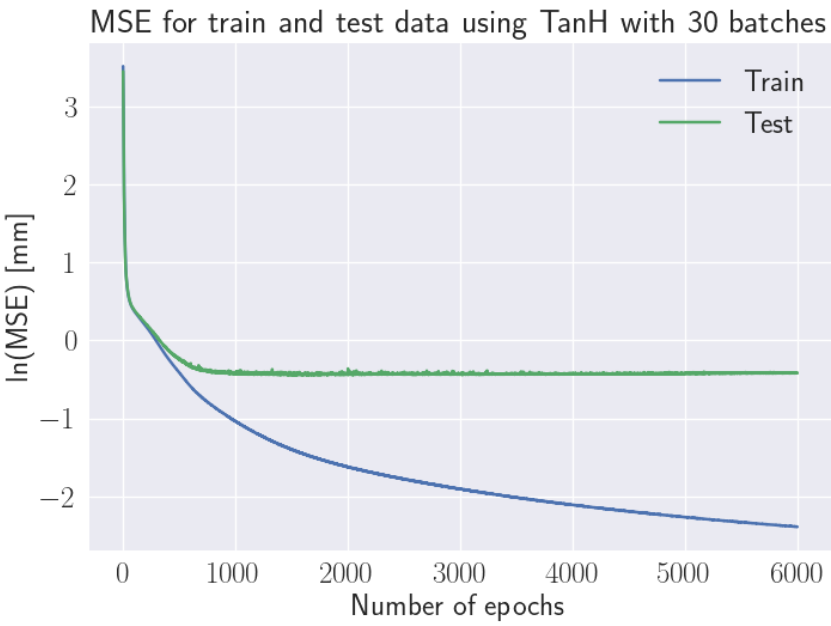

Figure 5.8: New text.