TECHNICAL UNIVERSITY OF CRETE

Electromagnetic Brain Source Analysis with Statistical and Deep Learning Approaches

Author:
Athanasios Delatolas

Supervisor:
Prof. Michalis Zervakis

A thesis submitted in fulfillment of the requirements for the Diploma in Electrical and Computer Engineering

in the

Digital Image & Signal Processing (DISPLAY) Laboratory
School of Electrical and Computer Engineering

June 22, 2022

Thesis Committee

Professor Michalis Zervakis (Supervisor)
Professor Athanasios Liavas
Professor Carsten H. Wolters

TECHNICAL UNIVERSITY OF CRETE

Abstract

Digital Image & Signal Processing (DISPLAY) Laboratory
School of Electrical and Computer Engineering

Electrical and Computer Engineering

Electromagnetic Brain Source Analysis with Statistical and Deep Learning Approaches

by Athanasios Delatolas

Electroencephalography (EEG) is a well-established non-invasive recording method for the brain's functional activity. EEG uses an array of electrodes placed on the scalp to record electrical potential signals. EEG provides high temporal but low spatial resolution of brain activity. To gain insight about the spatial dynamics of the EEG, one has to solve the inverse problem of Source Analysis, that is, to find the neural sources that give rise to the recorded EEG activity. There are many existing numerical methods for solving the inverse problem but most of them strongly rely on priors and require significant amount of computational time. Recently, neural networks have been proposed to resolve these issues but their training is based on suboptimal forward modeling and they cannot localize EEG recordings in various brain anatomies. Here, we present a new CNN architecture which is independent of the modeled brain source space and its training is based on realistic and skull-conductivity calibrated head modeling. The performance of our CNN is validated with simulated EEG data and real EEG somatosensory evoked potentials for the first neurological component at 20 ms (P20/N20 response) from three healthy subjects. Our network has localized the P20/N20 component at the subject-specific Brodmann area 3b. Finally, the results suggest that our CNN outperforms the traditional numerical methods.

Acknowledgements

After six amazing years at the Technical University of Crete, I want to express my deepest gratitude to the people that supported me throughout these years.

First and foremost, I would like to thank my supervisor, Prof. Michalis Zervakis for exposing me to very fascinating research areas. Without his guidance and support, the completion of this thesis wouldn't be possible. I am also grateful to Dr. Marios Antonakakis for his crucial suggestions and insights in both the theoretical and technical part of this study. Prof. Zervakis and Dr. Antonakakis tremendously helped me with my master's and doctoral applications.

I also owe a lot to Prof. Michael Paterakis and Prof. Apostolos Dollas for their advice and recommendations during my graduate applications. Moreover, I would like to thank Prof. Athanasios Liavas and Prof. Carsten H. Wolters for their time and support to the thesis committee.

Finally, this journey wouldn't be possible without the help of my family and friends. My parents unconditionally supported me through every decision I made and gave me the opportunity to pursue all of my goals. I also want to express my gratitude for my close friends and my brother, who were there for me both in good and in bad times.

Contents

List of Figures

List of Tables

List of Abbreviations

BEM	Boundary element method
CSF	Cerebrospinal fluid
CNN	Convolutional neural network
EEG	Electroencephalography
EMEG	Combined Electroencephalography,Magnetoencephalography
EW	Electric-Wrist
FEM	Finite element method
MEG	Magnetoencephalography
MNE	Minimum norm estimate
MRI	Magnetic resonance imaging
SEF	Somatosensory evoked fields
SEP	Somatosensory evoked potentials
S1	Primary somatosensory cortex
SNR	Signal-to-noise ratio
T1w MRI	T1-weighted magnetic resonance imaging
sLORETA	Standardized low-resolution brain electromagnetic tomography
WHO	World Health Organization

Chapter 1

Introduction

1.1 The importance of Source Analysis

Electroencephalography (EEG) [1] and Magnetoencephalography (MEG) [2, 3] are two non-invasive recording methods for the brain's functional activity. EEG uses an array of electrodes placed on the scalp to record voltage fluctuations, whereas MEG uses sensitive magnetic detectors called superconducting quantum interference devices (SQUIDs) [4] to measure the same primary electrical currents that generate the electric potential distributions recorded in EEG. Since EEG and MEG capture the electromagnetic fields produced by neuronal currents, they provide high temporal but low spatial resolution of the neuronal activity.

Source analysis consists of two problems. Firstly, the forward problem, which is to build a head volume conductivity model for describing how the electrical signals of the brain signal source are transmitted to the scalp electrodes [5]. Secondly, the inverse problem, that is, to localize the active brain-sources based on the solution of the forward problem [6].

It is vital to gain insight about the spatial dynamics of neuronal activity (i.e. solution of the inverse problem) because accurate localization of active brain-sources holds promise to enable novel treatments. The information acquired through the brain source localization is helpful to diagnose different brain disorders such as epilepsy, schizophrenia, depression, Alzheimer and Parkinson disease. Among them, the most common is epilepsy; according to World Health Organization (WHO) statistics, around 50 million people in the world have been diagnosed as epileptic [7].

The ultimate goal of source analysis is to provide real-time localization of the neuronal activity in order to facilitate treatments of the aforementioned diseases.

1.2 Aim & Innovation

Deep Learning [8] offers a promising new approach to significantly improve source localization in real time. In this thesis, we develop and present a novel deep learning solution to localize neural sources, and assess its accuracy and robustness with real EEG-recordings [9, 10]. We aim to provide

real-time source localization, that is, to be able to localize EEG recordings from different brain anatomies with very low computational time.

The first step to achieve our goal is to to model realistically the geometrical and electromagnetic features of the head. A six-compartment (skin, skull compacta, skull spongiosa, cerebrospinal fluid, white and gray matter) anatomical template with 50,460 dipoles was used [10]. Then, we solve the forward problem using the Finite Element Method (FEM) [11] because of its high flexibility to accurately model the electromagnetic field propagation in geometrically challenging inhomogeneous and anisotropic head volume conductors such as the human head [12, 13].

Having calculated a realistic leadfield matrix, we aim to create a CNN that can localize EEG signals from any recording system and to any brain-anatomy. Thus, our CNN takes as input a topography of the EEG-signal and estimates the location of the seed dipole which raised the recorded EEG signal. In more detail, a topography can be generated from any number of electrodes and the output layer of our neural network has only three neurons where each neuron corresponds to the coordinates in the three-dimensional source space.

While we focus on EEG, the same approaches are directly extendable to MEG, enabling a portable and affordable solution to source localization. The conceptual novelties of the proposed Convolutional Neural Network (CNN) [14] are :

1. Localization accuracy in real EEG-recordings. We tested our CNN on three different subjects (see Chapter 6).

2. Better performance from other deep learning approaches as we not only used FEM [11] but also a six-compartment head model [10] with 50,460 dipoles.

3. Very low computational time (once trained) for the estimation of the source location.

4. Because of the fact that the proposed CNN takes as input a topography and not the EEG-signal, one can use a topography generated from a different EEG-recording system than ours.

The proposed CNN paves the way for real-time source localization as it localizes correctly the EEG recorded activity for different brain anatomies. In particular, our neural network is tested in three different subjects. Each subject has its own anatomy, source space and leadfield matrix. Thus, we can correctly localize EEG data regardless of the geometrical features of the head. Even though pioneer studies suggest that their neural networks may work for various anatomies with a transfer learning approach, we managed to solve the problem of individual brain anatomies using FSL Registration [15, 16, 17]. More specifically, the MRI of each subject is interpolated with the spline interpolation (computational time: 3-5 minutes) using the FSL tool. Hence, the predicted three-dimensional coordinates of our neural network are referring to the interpolated MRI.

Even though all the numerical methods can also localize EEG to different brain anatomies, they require the leadfield matrix of each subject (i.e. the solution of the forward problem). Since, the solution of the forward problem needs approximately 40 minutes, numerical methods cannot provide real-time EEG source localization. Due to the fact that our CNN does not need the leadfield matrix of each subject, it can provide real-time source localization for various anatomies.

Finally, another advantage of the numerical methods is that they take into account the inter-subject variability of skull conductivity which causes the most significant influence on the EEG localization [18]. Therefore, it is worth investigating a new deep learning approach that can overcome the problem of the inter-subject variability.

1.3 Related Work

The first pioneer studies that proposed an artificial neural network as a solution to the inverse problem are [19, 20, 21]. However, these studies are limited by the sample size, network depth and computational power at that time, leading to a poor performance. In recent years, several deep learning methods have been proposed for the inverse problem. Among them, two Multi-Layer Perceptron(MLP) networks [22, 23] and various CNNs such as [24, 25, 26].

The aforementioned neural network-based inverse solutions are based on suboptimal forward modeling and they cannot localize EEG recordings in various brain anatomies. However, they have been tested in a distributed dipole model with more than two dipole sources.

There is a long history to study numerical algorithms for the EEG inverse problem. These methods are divided into three categories: non-parametric methods, parametric methods [27] and Bayesian methods [28]. A non-parametric and one of the most famous numerical solutions is minimal norm estimation (MNE) [29]. According to [2], MNE provides best estimates when less a priori information about the source distribution is available. Moreover, low resolution electromagnetic tomography (LORETA) [30] localizes the 3D solutions properly as compared with previous minimum norm approach. The first variation on LORETA is standardized LORETA (sLORETA) [31]. This method utilizes MNE to estimate the current density and it furthers standardizes the current density with the expected standard deviation. Another variation of LORETA is exact (eLORETA) [32]. This method gives more importance to the deeper sources. Moreover, another promising inverse methods is beamforming. The basic idea of beamforming is the application of spatial filtering on the measured data to distinguish signals arriving from a region of interest and suppress those originating elsewhere [33].

Even though, these methods are verified to many studies [27, 34], they still have limitations. In particular, although MNE offers good results in terms of resolution and current estimation, it is unable to address the issue of deep source localization in the outermost cortex [35]. In addition, unlike

the MNE solution, LORETA can localize the boundary and deep sources. However, the solution provided has low spatial resolution, which is an undesirable feature when we are dealing with pattern-recognition applications of brain source localization [35]. Furthermore, localization accuracy of sLORETA and eLORETA methods is better than LORETA, but their spatial resolution is not appropriate [34]. Finally, drawbacks of the beamforming approach are the susceptibility to imprecisions in the forward model and that correlated sources are often not found [24].

1.4 Thesis Organization

This work is organized in 7 chapters:

In **chapter** 1, we outline the necessity for Source Analysis. We then explain the problem we are trying to address and its importance. Finally, we highlight pioneer studies and state the novelties of this study.

In **chapter** 2, the reader is introduced to the fundamental and theoretical knowledge of the topics covered in this thesis. Firstly, we present a short description of the human brain anatomy and secondly, we summarize the principals of Electroencephalography (EEG).

In **chapter** 3, the mathematical background of Source Analysis is introduced. On the one hand, the forward problem is concerned with the computation of scalp potentials and external fields for a specific set of neural current sources. On the other hand, the aim of the EEG inverse problem is to find the location inside the brain of the signal that is responsible for the measured EEG data.

In **chapter** 4, we briefly discuss the principals of Deep Learning. Having initially introduced the multilayer perceptron (MLP) model, we examine the convolutional neural network. We also present three important ideas convolution leverages, those are, sparse interactions, parameter sharing and equivariant representations [8].

In **chapter** 5, we explain the pipeline that implements our CNN. Briefly, given a six-compartment head model that models realistically the geometrical and electromagnetic features of the head, we solve the Forward Problem using FEM [11]. Having calculated the leadfield matrix, we generate multiple training data for our neural network. Finally, we present the architecture of our CNN.

In **chapter** 6, we compare the performance of the deep learning model against the popular scanning localization algorithms: sLORETA and Dipole Scan. To access the localization results of our CNN, we use both simulated and real EEG recordings.

In **chapter** 7, we draw our conclusions, summarize our results and outline our contribution. We then talk about the drawbacks of our approach and suggest potential areas for future work.

Chapter 2

Background

This chapter provides a brief introduction to both the anatomy of brain and the electroencephalography (EEG) recording system.

2.1 Anatomy of the Human brain

The human brain is the most complex organ with 10^{12} neurons, which are interconnected via axons, dendrites and 10^{15} synaptic connections. Moreover, brain only weighs around 1500g [36] . This complex structure allows it to release/absorb quintillion of neurotransmitter and neuromodulator molecules per second. Brain development starts at a primary age of 17–18 weeks of parental development and generates the electrical signals until death [1].

Brain is divided into several anatomical and functional regions. It is surrounded by many bones which together form the **skull** and it is surrounded by the **Cerebrospinal fluid (CSF)**. CSF flows inside the brain through the cerebral ventricles. Moreover, brain's tissues can be broken down into two major classes: **gray** and **white matter**.

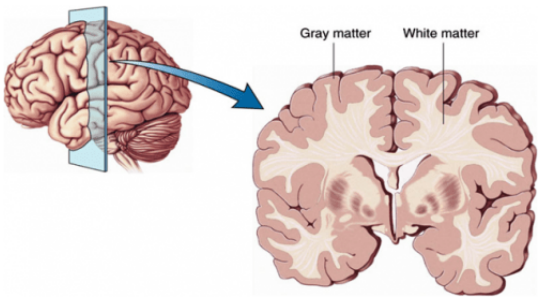


Figure 2.1: Gray and White matter

The brain regions formed by gray matter are responsible for processing information and establishing connections with white matter. The gray matter is mostly composed of unmyelinated neurons. The white matter is composed of myelinated neurons, which are used as connectors to the gray matter. Because myelinated neurons transmit nerve signals faster, white matter functions to increase the speed of signal transmission between the connections.

Furthermore, according to embryonic developments, the human brain can be divided into three regions anatomically: the **forebrain** (or *prosencephalon*), the **midbrain** (or *mesencephalon*), and the **hindbrain** (or *rhombencephalon*) [37].

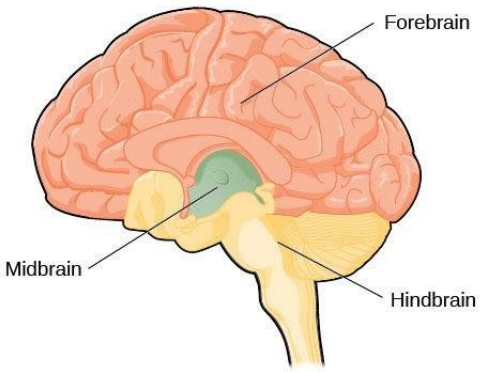


Figure 2.2: The three regions of the brain

2.1.1 Forebrain

By far the largest region of the brain is the forebrain, which contains the entire **cerebrum** and several structures directly nestled within it - the thalamus, hypothalamus, the pineal gland and the limbic system. When we picture the iconic shape of the human brain, the majority of what's visible is the cerebrum with its wrinkly, pinkish-grey outer appearance. It makes up around 85% of the brain and consists primarily of gray matter, divided into two hemispheres.

The cerebrum is where most of the important brain functions happen, such as thinking, planning, reasoning, language processing, and interpreting and processing inputs from our senses, such as vision, touch, hearing, taste and smell. The outer layer of the cerebrum is called the cerebral cortex.

Each hemisphere it is traditionally divided into four lobes - frontal, parietal, occipital and temporal. Communications between the two hemispheres are maintained by a fibrous bridge called the corpus callosum.

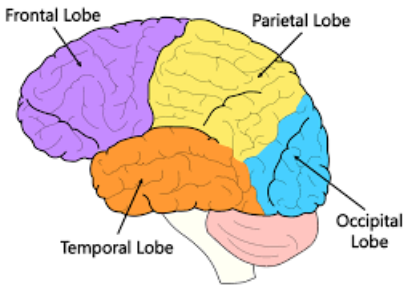


Figure 2.3: Lobes of the brain

- The frontal lobe, which lays at the front of the brain, controls executive functions like thinking, planning, and problem solving, as well as memory, attention and movement.
- The parietal lobe, which sits behind the frontal lobe, deals with the perception and integration of stimuli from the senses.
- The temporal lobe, which runs along the side of the brain under the frontal and parietal lobes, deals with the senses of smell, taste and sound and the formation and storage of memories.
- The occipital lobe, which is at the back of the brain, is concerned with vision.

2.1.2 Hindbrain

The hindbrain is one of the three major regions of our brains, located at the lower back part of the brain. It includes most of the brainstem and a dense coral-shaped structure called the **cerebellum**. The brainstem is one of the most important parts of the entire central nervous system, because it connects the brain to the spinal cord and coordinates many vital functions, such as breathing and heartbeat. There are three main parts of the hindbrain - pons, cerebellum, and medulla oblongata

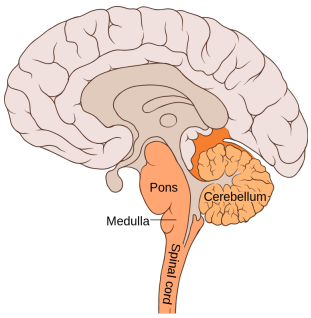


Figure 2.4: Hindbrain

- The pons gets its name from the Latin word for "bridge", and it connects the rest of the brainstem to the cerebral cortex. Bulbous in shape, it sits right underneath the midbrain and serves as a coordination centre for signals and communications that flow between the two brain hemispheres and the spinal cord.
- Behind the pons and the rest of the brainstem sits a structure called the cerebellum (Latin for "little brain"). Just like the cortex, it has two hemispheres, with a dense layer of gray matter surrounding an inner region of white matter. It also contains special neurons called Purkinje cells, capable of processing many signals at once due to their highly complex dendrite branches.

- The lower part of both the brainstem and the overall hindbrain is the medulla oblongata, where the brain transitions to the spinal cord. It is only about 3cm long, but the medulla is an indispensable nerve tract which contains the control centres for our autonomic vital functions - heart rate, blood pressure, breathing - and many involuntary reflexes such as swallowing and sneezing.

2.1.3 Midbrain

Located towards the base of your brain is a small but important region called the midbrain, which serves as a vital connection point between the other major regions of the brain - the forebrain and the hindbrain.

2.1.4 Primary Somatosensory Cortex

The primary somatosensory cortex (S1) is located in the postcentral gyrus of the parietal lobe and controls the somatic sensory information of the body. As it is shown in Figure 2.5, each part of the human body occupies a region at S1.

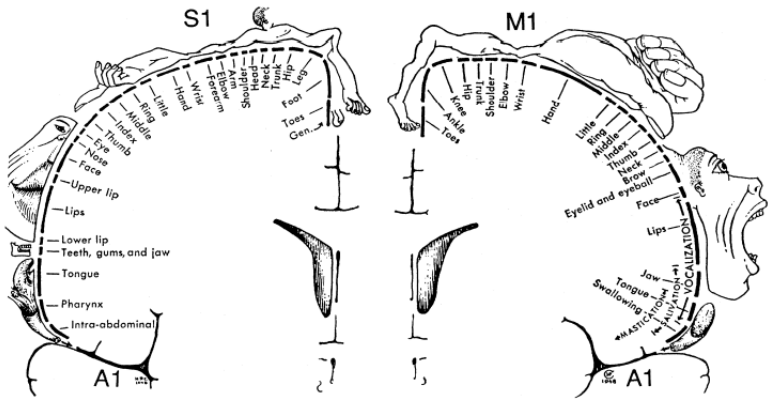


Figure 2.5: Representations of different body parts on the somatosensory (S1). Adapted from [2].

The sensory signals transfer through the afferent nerves from the receptors to the spinal cord, brainstem, thalamus and finally to the primary somatosensory cortex.

Moreover, S1 is divided into areas: 3 (which is subdivided into 3a and 3b), 1, and 2. Area 3 is generally considered the primary area of the somatosensory cortex. Area 3 receives the majority of somatosensory input directly from the thalamus, and the initial processing of this information occurs there. Area 3b specifically is concerned with basic processing of touch sensations, while area 3a responds to information from proprioceptors. Area 3b is densely connected to areas 1 and 2. Thus, while area 3b acts as a primary area for touch information, that information is then also sent to areas 1 and 2 for more complex processing.

2.1.5 The P20/N20 Component

As described in [33], the first response recorded by EEG/MEG is located in the ventroposterior thalamus at 14 ms after the stimulus onset. The first transient response is located in the area 3b at 20 ms post-stimulus. The response at 20 ms post-stimulus is called the P20/N20 component, for which the term "P" refers to a positivity and the "N" to a negativity of the EEG over the frontocentral and parietooccipital lobes, respectively. The advantage of knowing the exact cortical position of the P20/N20 component makes S1 a candidate brain network for the investigation of the behavior of experimental and modeling source analysis parameters. Thus, the P20/N20 component is especially appropriate for sensitivity investigations of EEG and MEG source analysis.

2.2 Electroencephalography (EEG)

While in this thesis we focus on EEG, there are other non-invasive brain measurements such as magnetoencephalography (MEG) [2, 3], magnetic resonance imaging (MRI) and the combination of electroencephalography and magnetoencephalography EMEG.

2.2.1 EEG definition

Electroencephalography (EEG) [1] is defined as the non-invasive neuroimaging technique having high temporal, low spatial resolution. EEG records the brain activity by measuring electrical signals generated by pyramidal neurons located in cortical region of brain with the help of electrode arrangement. EEG is a neuroimaging technique that was developed by German physicist and psychiatrist Hans Berger in 1924, to measure brain activity with a set of electrodes. The set of electrodes is placed on the surface of the scalp to measure the electrical potential of the patient under observation [38, 35].

EEG can also be defined as a functional neuroimaging technique with high temporal resolution (in milliseconds), which measures potential differences as linear functions of source strengths and nonlinear functions of dipole locations [39]

EEG recordings can be used for direct, real-time monitoring of spontaneous and evoked brain activity, which enables the spatio-temporal localization of neuronal activity.

An example of eeg recording is shown in the figure 2.6

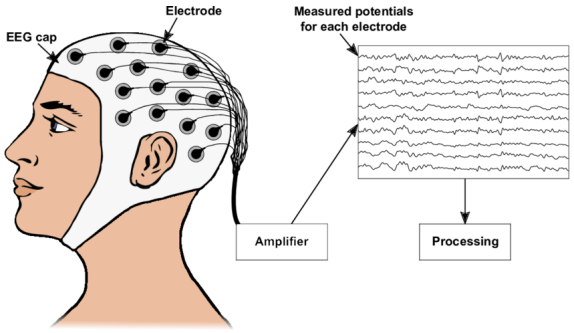


Figure 2.6: EEG recording, adapted from [40]

2.2.2 EEG system

In the early years of development, EEG signals were recorded using galvanometers. However, with the advancement in modern electronic circuitry, signals are now recorded using a set of electrodes, differential amplifiers for each channel, and filters. With the increase in the number of channels, the computational complexity is increased proportionally.

In particular, the EEG system consists of electrodes, amplifiers, filters and analogue to digital converters (ADCs) to store the signal into computer or a needle (pen)-type register to draw the signal on the paper.

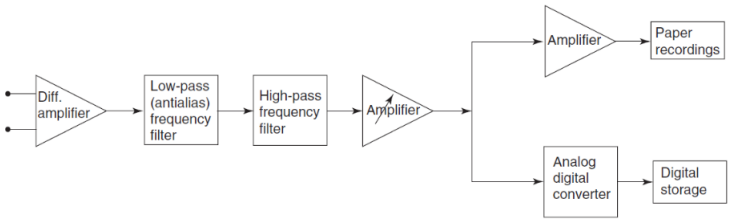


Figure 2.7: EEG system adapted from Wiley Encyclopedia of biomedical engineering [41]

The electrodes record the electrical activity produced by neurons. EEG machines use an array of electrodes because the brain produces different signals from different brain regions. The number of electrodes corresponds with the number of channels an EEG machine has. The more channels, the higher the resolution of EEG data captured. EEG signals are typically very small, around 10 microVolts or less. To produce accurate measurements, the signals from the electrodes are passed to an amplifier (as can be seen in figure 2.7) that stabilizes the signals and magnifies them to a level that can be measured accurately.

2.2.3 Electrodes placement

In 1958, a standard electrode placement system for the measurement of brain signals using EEG was proposed [42, 43, 44]. The system was called the 10-20 electrode placement system and it only consisted of 21 electrodes. In this setting, odd electrodes are placed on the left while even on the right. For an increased number of electrodes, one can simply place the additional electrode with equidistance between the electrodes. An extension of the 10-20 system is the 10-10 system. It is common practice that the electrodes connected to ear lobes (A1 and A2) are taken as reference electrodes. However, in modern instrumentation, the choice of a reference does not play a significant role in the measurement.

The standard nomenclature of the electrodes is based on their location. In particular, the first letters specify the location in the head: A = earlobe, O = occipital, P = parietal, F = frontal, and Fp = frontal polar, while the remaining name comes is a number or the letter 'z'. if it is an even number, then the electrode is located on the right side, if it is an odd number on the left side, while the letter 'z' (stands for zero) implies a central position.

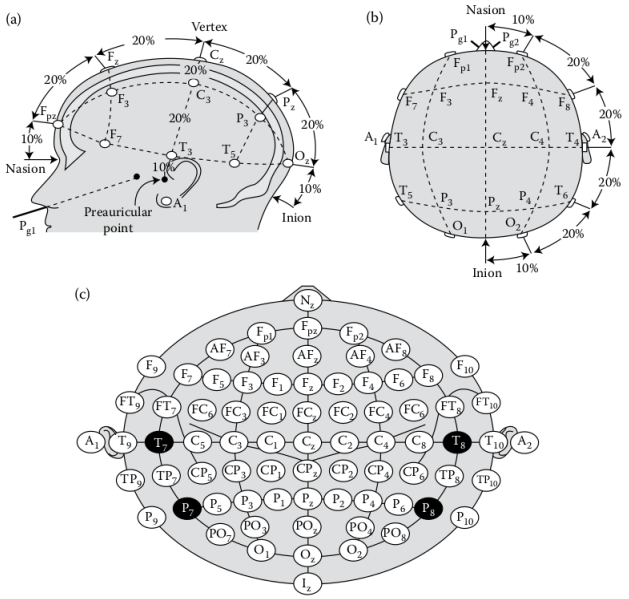


Figure 2.8: Diagrammatic view of the 10-20-electrode system.

A diagrammatic view for the 10-20 electrode system showing distance measurements for various head regions with 75 electrodes along with reference electrodes is shown in figure 2.8. Pictures (a) and (b) show three-dimensional measures while (c) shows a two-dimensional view of the electrode setup.

Chapter 3

Source Analysis

As mentioned in the chapter 2, EEG provides high temporal but low spatial resolution of the neuronal activity. It is vital to gain insight about the spatial dynamics of neuronal activity (i.e. solution of the inverse problem) because accurate localization of active brain-sources holds promise to enable novel treatments.

The aim of Source Analysis techniques is to find the brain areas responsible for the generation of the recorded signals from the EEG-electrodes. Especially in epilepsy, Source Analysis techniques are used in order to localize more accurately the epileptogenic zone, alongside with MRI, and facilitate the epilepsy surgery. Briefly, Source Analysis consists of two problems. Firstly, the EEG forward problem, which is to build a head volume conductivity model for describing how the electrical signals of the dipoles are transmitted to the scalp electrodes. Secondly, the EEG inverse problem, that is, to estimate the most possible dipoles which could generate the recorded signals from the electrodes. Solving the forward problem is a prerequisite to the inverse problem.

Due to the fact that the number of dipoles is significantly larger than the number of electrodes, without further constraints, this problem problem can be characterized as ill-posed [27]. The goal of this thesis is to solve the inverse problem using a neural network.

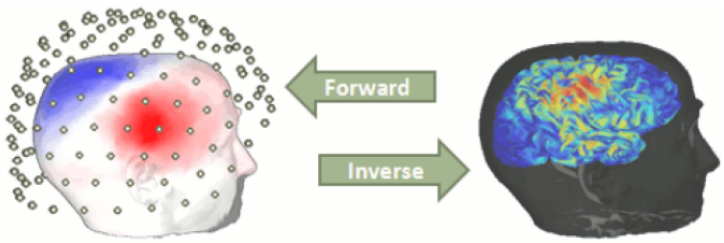


Figure 3.1: Forward vs Inverse

Before we delve deeper into the forward and inverse problem, it is vital to define the dipole and the source space.

Definition 1 (Dipole) *A source located in the brain that generates electrical activity is called dipole.*

Definition 2 (Source Space) *The source space (or source model) is a set with all candidate positions of the dipoles inside the brain (e.g. gray matter).*

Definition 3 (Dipole's parameters) *A dipole is determined according to the position in the source space and moment (amplitude and orientation).*
If d is the amplitude and $\overrightarrow{e}$ *is the orientation, then the moment of the dipole is*

$$m = d\overrightarrow{e} \tag{3.1}$$

3.1 Forward Problem

The forward problem is concerned with the computation of the channels measurements (values of the electrodes) for a specific set of neural current sources (dipoles). To put it differently, given the amplitude and the orientation of all dipoles in the source space and the EEG channel measurements must be calculated. Even though this problem does not occur in reality -we always know the channel measurements-, it is a prerequisite for the inverse problem.

Moreover, the EEG electrodes are located on the scalp while the source model is located inside the head. Thus, a head model is required which is a simulation of the geometrical and electromagnetic features of the head. The head modeling is carried out with various numerical techniques [11, 45]. However, prior to this, it is necessary to understand the mathematical background behind it.

As previously stated, each source corresponds to a single known position in source space, which is described in the Cartesian coordinate system. Therefore, in order to determine a dipole in a certain location with the Cartesian coordinate system, only the amplitude d is unknown as the orientation of each dipole is known $\{\overrightarrow{e_x}, \overrightarrow{e_y}, \overrightarrow{e_z}\}$. For simplicity reasons, we assume that the source is determined by one amplitude and one orientation. Hence, the moment of the k_{th} dipole is $m_k = d_k\overrightarrow{e_k}$ instead of $m_k = \begin{bmatrix} d_{k\text{x}}\overrightarrow{e_x} \\ d_{k\text{y}}\overrightarrow{e_y} \\ d_{k\text{z}}\overrightarrow{e_z} \end{bmatrix}$.

For p dipoles and q electrodes and for each time point:

$$\mathbf{X}_{\text{est}} = \mathbf{LS} \tag{3.2}$$

where, $\mathbf{X}_{\text{est}} \in \mathbb{R}^{q\times t}$ is the EEG channel measurements over time ; $\mathbf{L} \in \mathbb{R}^{q\times p}$ is the leadfield matrix that describes the flow of electrical current of each dipole through every electrode [27] ; $\mathbf{S} \in \mathbb{R}^{p\times t}$ and represents the electrical current of each dipole over time.

For a specific time point t_k, the above equation is:

$$X_{\text{est}_k} = \underbrace{\begin{bmatrix} X_{\text{est}}(elec_1)_k \\ \vdots \\ X_{\text{est}}(elec_q)_k \end{bmatrix}}_{q \times 1} = \underbrace{\begin{bmatrix} l(r_1, r_{\text{source}_1}) & \cdots & l(r_1, r_{\text{source}_p}) \\ \vdots & \ddots & \vdots \\ l(r_q, r_{\text{source}_1}) & \cdots & l(r_q, r_{\text{source}_p}) \end{bmatrix}}_{q \times p} \underbrace{\begin{bmatrix} d_1 \overrightarrow{e_1} \\ \vdots \\ d_p \overrightarrow{e_p} \end{bmatrix}}_{p \times 1} \tag{3.3}$$

where, $r_1 \cdots r_q$ are the positions of each electrode; $r_{\text{source}_1} \cdots r_{\text{source}_p}$ are the positions of each dipole in the source space.

Generally a noise or perturbation matrix $\mathbf{n}$ [27] is added:

$$\mathbf{X}_{\text{est}} = \mathbf{LS} + \mathbf{n} \tag{3.4}$$

The noise is modeled as zero mean Gaussian random variable: $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_q)$ where σ is the standard deviation of the noise and $\mathbf{I}_q$ is the identity matrix.

3.1.1 Maxwell's equations

Understanding Maxwell's equations is required to go deeper into the mathematical basis of the Forward Problem. The interrelation between various electrical and magnetic quantities was developed by James Maxwell in 1861. The set of equations that was later termed Maxwell's equations provides the relationship between electromagnetic field and the charge density and current density [35].

The electromagnetic field is the combination of the electric field $\mathbf{E}(V/m^2)$ and the magnetic field $\mathbf{B}(T)$ which can be computed by Maxwell's equations and the Continuity equation [2, 46, 13]. In particular, the Maxwell's equations are:

$$\nabla \mathbf{E} = \frac{\rho}{\epsilon} \tag{3.5}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.6}$$

$$\nabla \mathbf{B} = 0 \tag{3.7}$$

$$\nabla \times \mathbf{B} = \mu\left(\mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t}\right) \tag{3.8}$$

and the Continuity equation is:

$$\nabla \mathbf{J} = -\frac{\partial \rho}{\partial t} \tag{3.9}$$

where, ρ is the volume charge density in C/m^3, $\mathbf{J}$ is the current density in A/m^2, μ is magnetic permeability in H/m and ϵ is the electrical permittivity

of the medium in F/m. Moreover, Maxwell's equations and the Continuity equation can be simplified under the assumptions:

- The magnetic permeability μ for the human tissues is equal to vacuum permeability μ_0
- $\frac{\partial \mathbf{B}}{\partial t}$, $\frac{\partial \mathbf{E}}{\partial t}$ can be neglected while computing $\mathbf{E}$ and $\mathbf{B}$

3.1.2 Quasi-static approximation

For neuroscientists, the frequency range is crucial because most brain research are conducted at frequencies between 0.1 and 100 Hz. Therefore, the physics is well defined using quasi-static approximations for Maxwell's equations [35]. This approximation is also valid as the brain signals are generated at lower frequencies only. The time-derivative component of Maxwell's equations is neglected at these lower frequencies.

Using Ohm's law, $\mathbf{J} = \sigma\mathbf{E}$ (where σ is the conductivity of the medium), and the aforementioned assumptions, the equation (3.8) becomes

$$\nabla \times \mathbf{B} = \mu_0(\sigma\mathbf{E} + \epsilon\frac{\partial \mathbf{E}}{\partial t}) \tag{3.10}$$

Thus, for a quasi-static approximation:

$$\frac{\epsilon}{\sigma} << 1 \tag{3.11}$$

and

$$\left|\epsilon\frac{\partial \mathbf{E}}{\partial t}\right| << |\sigma\mathbf{E}| \tag{3.12}$$

Moreover, the current density $\mathbf{J}$ is divided into primary current density $\mathbf{J}^{\mathrm{p}}$ and the volume current density $\mathbf{J}^{\mathrm{v}} = \sigma\mathbf{E}$. Using the definition related to brain volume, $\mathbf{J}^{\mathrm{p}}$ is the current destiny due to the neuronal activity in the brain. This current density is spatially bounded in a volume. However, the volume current density ($\mathbf{J}^{\mathrm{v}}$) flows due to the electric field in the volume under observation. Thus, the total current destiny is given as follows:

$$\mathbf{J} = \mathbf{J}^{\mathrm{p}} + \mathbf{J}^{\mathrm{v}} = \mathbf{J}^{\mathrm{p}} + \sigma\mathbf{E} \tag{3.13}$$

Magnetic fields and electric currents behave as stationary at all times since the estimated frequency range of neural sources is less than 1kHz. Consequently, the quasi-static approximation of Maxwell's equations can be used. By differentiating equation (3.13) and taking into account that $\nabla\mathbf{J} = 0$ in the quasi-static approximation:

$$\begin{aligned} \nabla\mathbf{J} &= \nabla\mathbf{J}^{\mathrm{p}} + \sigma\nabla\mathbf{E} \Rightarrow \\ \nabla\mathbf{J}^{\mathrm{p}} &= -\sigma\nabla\mathbf{E} \end{aligned} \tag{3.14}$$

Provided that, $\nabla \times E = 0$ (quasi-static approximation) and $\mathbf{E} = -\nabla V$ (V stands for potential), the equation (3.14) can be written as follows:

$$\nabla \mathbf{J}^{\mathrm{p}} = -\nabla\big(\sigma(\nabla V)\big) \tag{3.15}$$

To solve the Forward Problem, the potential V must be estimated using analytical or numerical techniques. Afterwards, the computation of the magnetic field ($\mathbf{B}$) is straightforward. The whole forward problem is summarized in the computation of the magnetic field $\mathbf{B}$ outside the head from a given primary current distribution $\mathbf{J}^{\mathrm{p}}$ within the brain [2].

Brain sources are represented as mathematical point dipoles:

$$\mathbf{J}^{\mathrm{p}} = m_0 \cdot \delta(x - x_0) \tag{3.16}$$

where m_0 is the moment of the dipole , δ is the Dirac Function and x_0 is the position of the dipole. Moreover, given that the magnetic potential $\mathbf{A}$ can be written as:

$$\mathbf{A}(x) = \frac{\mu}{4\pi} \int_{\Omega} \frac{\mathbf{J}^{\mathrm{p}}(\tilde{x}) - \sigma(\tilde{x})\nabla V(\tilde{x})}{|x - \tilde{x}|}\, d\tilde{x} \tag{3.17}$$

By combining the above equations, the magnetic flux Ψ measured within a surface area S and circumference l can be written as:

$$\begin{aligned} \Psi &= \int_S \mathbf{B}\, d\tilde{S} \\ &= \oint_l A(x) d\tilde{x} \\ &= \frac{\mu}{4\pi}\left[\oint_l \int_{\Omega} \frac{\mathbf{J}^{\mathrm{p}}(y)}{|x - y|}\, dy dx + \oint_l \int_{\Omega} \frac{-\sigma(y)\nabla V(y)}{|x - y|}\, dy dx \right] \end{aligned} \tag{3.18}$$

where the first term is the primary magnetic flux and the second term is the secondary magnetic flux.

3.1.3 Conductivity

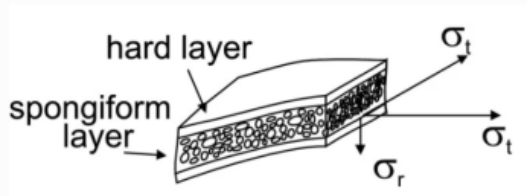
According to Ohm's law, the current density $\mathbf{J}$ is given by the electric field $\mathbf{E}$ as follows:

$$\mathbf{J} = \sigma \mathbf{E} \tag{3.19}$$

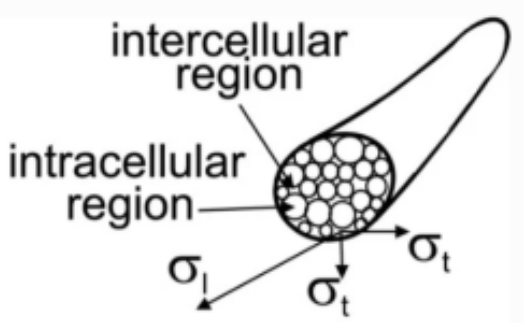
where $\sigma \in \mathbb{R}^{3\times 3}$ and

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \tag{3.20}$$

with units (S/m)



(a) The three layers of the skull. Radial conductivity (σ_r) is 10 times smaller than the tangential conductivity (σ_t). Adapted from [45].



(b) White matter consists of different nerve fibers (axons grouped in bundles). The conductivity along the nerve bundle is 9 times larger than perpendicular to the nerve bundle. Adapted from [45].

(c) The different colors indicate the primary fiber orientation (red: left-right, green: anterior-posterior and blue superior-inferior) Adapted from [47].

Figure 3.2: Anisotropy properties of the conductivity of skull and white matter.

In terms of conductivity, human head compartments are divided into two categories: isotropic and anisotropic [48]. Isotropic conductivity denotes that the current flow is the same in every direction (conductivity is equal in all directions [45]). Isotropic conductivity characterizes grey matter, scalp, and cerebrospinal fluid (CSF), and the position dependent conductivity tensor, σ, is reduced to a position dependent conductivity scalar. Anisotropic conductivity denotes conductivity inequality across the different directions since the electric field can induce a current density component perpendicular to it with the appropriate σ in the equation (3.19) [45]. Skull and white matter are compartments of anisotropic conductivity; The skull is formed by two hard layers with low conductivity and a spongiform layer with higher conductivity between them. White matter consists of different nerve fibers with higher conductivity in the direction along them. Anisotropic conductivity tissues put the forward problem calculation and, by extension, the inverse problem calculation at jeopardy [12, 13].

3.1.4 Boundary Conditions

During head modeling, two boundary conditions are set for separate parts of the head. They're known as *Neumann* and *Dirichlet* boundary conditions, and they're used at the intersections of two separate areas. The difference between two regions is checked by their respective conductivities (σ_1 and σ_2) and the unit normal vector $\mathbf{e}_n$ to the interface between the regions.

The Neumann boundary condition states that the charges are not accumulated on the intersections. Rather they are traveling after leaving one intersection. In other words, because the head is a pure resistive medium, there is continuity in current from one interface to the other. Hence, all the current leaving a region with conductivity σ_1 through the interface enters into a neighboring region with conductivity σ_2. Mathematically,

$$\mathbf{J}_1 \cdot \mathbf{e}_n = \mathbf{J}_2 \cdot \mathbf{e}_n \tag{3.21}$$

and

$$(\sigma_1 \nabla V_1) \cdot \mathbf{e}_n = (\sigma_2 \nabla V_2) \cdot \mathbf{e}_n \tag{3.22}$$

A special case of the Neumann boundary condition is for homogeneous medium, which is called the homogeneous Neumann boundary condition [49]. This states that due to the low conductivity of air at the outer surface of the human head, no electricity can travel from the head into the air. Hence, mathematically it can be defined as follows:

$$\mathbf{J}_1 \cdot \mathbf{e}_n = 0 \tag{3.23}$$

and

$$(\sigma_1 \nabla V_1) \cdot \mathbf{e}_n = 0 \tag{3.24}$$

The other boundary condition for the forward problem solution is the Dirichlet boundary condition. It is limited to internal interfaces and explains the potential at the boundary; hence, it states that the electric potential shows continuity across the interfaces, such that

$$V_1 = V_2 \tag{3.25}$$

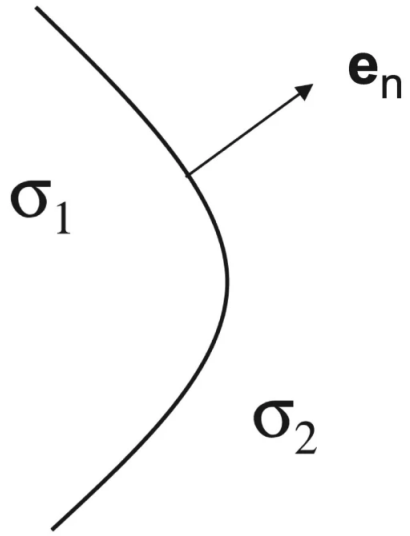


Figure 3.3: Boundary between two compartments with conductivities σ_1 and σ_2. Adapted from [45].

The equation (3.25) represents the Dirichlet boundary condition. Besides this, a reference electrode with zero potential is also assigned such that

$$V_{\text{ref}} = 0 \tag{3.26}$$

Hence, for the forward problem solution with different head models, which are to be used for the calculation of potential due to dipole sources, the aforementioned boundary conditions are taken into consideration. This will lead to a proper solution with less errors and more resolution

3.2 Head Volume Conductor Models

Simple and realistic head models are the two types of head models available. The former uses analytical homogeneous single spheres or multi-spheres, whilst the later uses numerical solutions to approximate head geometry, such as Boundary Element Method and Finite Element Method (FEM). MRI scan can be involved on both categories in order to assist head modeling for more accurate results [50].

Despite the fact that spherical models are simple to use and computationally light, they result in inaccurate source estimations due to crude approximations of the human head. Realistic head models solve Maxwell's equations with numerical methods. The Boundary Element Method (BEM) and the Finite Element Method (FEM) are the two main numerical techniques. In these approaches MRI scan of the subject's head is necessary in order to extract the geometric tessellations of the various envelopes forming the head tissues [50].

3.2.1 Boundary Element Method (BEM)

One of the numerical methods used to solve Poisson's equation in a realistic head volume conductor model is the boundary element method (BEM) [51]. It is a numerical technique for solving linear partial differential equations defined over certain limits of various domains once they have been converted from differential to integral form (weak formulation). The structural information of the subject is taken using MR images. The boundary integral equations are formed from the forward problem. Therefore, BEM uses triangles (boundary elements) as building blocks and surface boundaries are defined upon each triangle.

3.2.2 Finite Element Method (FEM)

FEM is a very useful numerical tool, which is used for solving boundary value problems that are defined by a differential equation with a set of boundary conditions [52]. FEM also uses as a building block, elementary volumes such as tetrahedron or hexahedrons [12, 45, 53, 54]. The potential difference is calculated on the vertices of the building blocks. FEM, also, allows anisotropic tissue conductivity, such as the three-layered skull bone (compacta-spongiosa-compacta), CSF and the white matter. Moreover, FEM's calculation is a laborious and complex task, therefore it requires a high computational time.

An important decision for performing FEM in reasonable amount of time is to choose a proper method to treat the singularity introduced by the mathematical dipole. Proposed approaches are "partial integration direct potential", "subtraction" and "Venant direct" approach [33].

In the present thesis, we used the Venant direct source modeling approach due to its high numerical accuracy and high computational efficiency when used in combination with EEG and MEG transfer matrices and an algebraic multigrid preconditioned conjugate gradient (AMG-CG) solver [10]. Further details about the solution of the Forward problem can be found in Chapter 5.2.

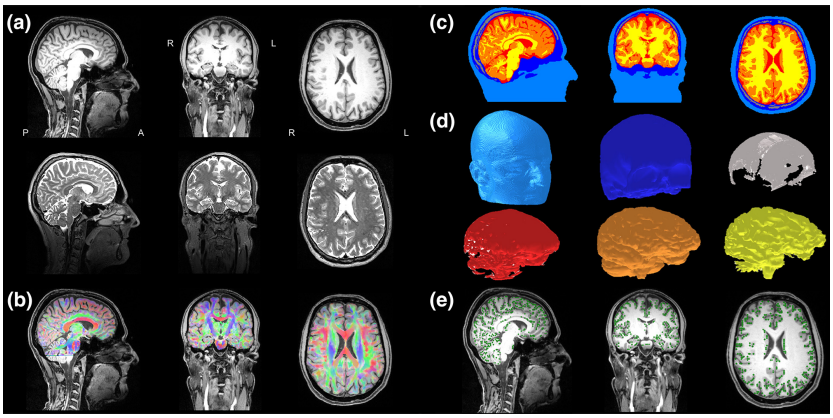


Figure 3.4: Six-compartment anisotropic realistic head model using FEM and source space. Adapted from [10].

3.3 Inverse Problem

Modeling is used in numerous domains such as engineering, physics, and applied mathematics to forecast the effects or outcomes of a collection of known variables. This is known as forward modeling or forward problem. When solving an inverse problem, we typically start with the results or measured data and then go on to infer what parameters 'caused' those measurements. Unlike in a forward problem, the inverse problem has no unique solution i.e., multiple (in theory infinite) set of parameters might explain the same measurement data.

The aim of the EEG inverse problem is to find the location of the signal that is responsible for the measured EEG data. Such solutions are can be extremely useful and important both in clinical neuroscience, for example to aid in the localization an epileptogenic foci and basic neuroscience to enhance our understanding about regions responsible for various cognitive tasks.

The EEG inverse problem is an ill-posed problem because for all admissible output voltages, the solution is non-unique (since $p >> q$, see equation (3.2)) and unstable (the solution is highly sensitive to small changes in the noisy data) [27]. In more detail, a sensor topography could be generated by an infinite number of active brain sources. This is due to the existence of sources (silent sources) that do not produce detectable electromagnetic signals but contribute to the solution. Assumptions are required in order to restrict the solution space. Mainly, the assumptions determine the nature of the sources, for example their quantity, anatomical and neurophysiological constraints, the a priori probability density functions and covariance models. Source localization accuracy depends on variant reasons such as head and source modelling errors and of course biological and non-biological noise.

The two main categories of Inverse Problem's solutions are the non- parametric and parametric methods [27]. Briefly, in non-parametric models a number of dipole sources with predetermined locations and orientations are allocated alongside the brain volume or cortical surface. Regarding the parametric methods, the aim is to find the best dipole position and orientation. The model consists either of a single dipole in a spherical head model or of multiple dipoles in a realistic head model. Moreover, dipole orientations can be predetermined or varying. On the other hand, non-parametric techniques estimate dipole's moment in a predetermined source space, leading on a linear problem. In this thesis, only non-parametric methods have been used, so an attempt to briefly explain them is made. The mathematical derivation following is based on the Bayesian framework.

3.3.1 Bayesian Framework

Both the EEG channel measurements $\mathbf{X}_{est} \in \mathbb{R}^{q \times t}$ (remainder: q is the number of EEG channels while p is the number of dipoles) and the the electrical current of each dipole $\mathbf{S} \in \mathbb{R}^{p \times t}$ are considered as random variables and the noise is modeled as zero-mean Gaussian random variable $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_q)$ [27].

The likelihood density is the conditional probability density of $\mathbf{X}_{est}$ given $\mathbf{S}$:

$$p_{\text{li}}(\mathbf{X}_{est}|\mathbf{S}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{q}{2}} \exp\left(-\frac{1}{2\sigma^2}\left|\left|\mathbf{X}_{est} - \mathbf{LS}\right|\right|_2^2\right) \tag{3.27}$$

In order to estimate the sources $\tilde{\mathbf{S}}$, the conditional density of $\mathbf{S}$ given $\mathbf{X}_{est}$ need to be calculated, which is called posterior density $p_{\text{post}}(\mathbf{S}|\mathbf{X}_{est})$. Furthermore, the a-priori information of $\mathbf{S}$ is encrypted in its density probability $p_{\text{pr}}(\mathbf{S})$ and Bayes' rule is applied as follows [55]:

$$p_{\text{post}}(\mathbf{S}|\mathbf{X}_{est}) = \frac{p_{\text{li}}(\mathbf{X}_{est}|\mathbf{S}) \cdot p_{\text{pr}}(\mathbf{S})}{p(\mathbf{X}_{\text{est}})} \tag{3.28}$$

$p(\mathbf{X}_{\text{est}})$ operates as a normalizing constant and thereby it does not have an important role. Finally in order to estimate the sources $\tilde{\mathbf{S}}$, the the maximum a posteriori (MAP) or conditional mean (CM) could be utilized [55].

$$\tilde{\mathbf{S}}_{\text{MAP}} := \text{argmax}\, p_{\text{post}}(\mathbf{S}|\mathbf{X}_{est}) \tag{3.29}$$

$$\tilde{\mathbf{S}}_{\text{CM}} := \mathbb{E}\left[\mathbf{S}|\mathbf{X}_{\text{est}}\right] = \int S p_{\text{post}}(S|\mathbf{X}_{\text{est}})\, dS \tag{3.30}$$

In order to link the aforementioned methods with other widely used techniques *Gibbs distributions* are used in the prior probability:

$$p_{\text{pr}}(\mathbf{S}) \propto \exp\left(-\frac{\lambda}{2\sigma^2}\mathcal{P}(\mathbf{S})\right) \tag{3.31}$$

where λ is the regularization parameter and $\mathcal{P}(\mathbf{S})$ is the penalty function. Finally the equation (3.29) can be written:

$$\tilde{\mathbf{S}}_{\text{MAP}} := \underset{\mathbf{s}\in\mathbb{R}^p}{\text{argmax}}\left\{\exp\left(-\frac{1}{2\sigma^2}\left|\left|\mathbf{X}_{est} - \mathbf{Ls}\right|\right|_2^2 + \frac{\lambda}{2\sigma^2}\mathcal{P}(\mathbf{s})\right)\right\} \tag{3.32}$$

3.3.2 Dipole Scanning

In this method a single point or a limited group of points that explain the largest amount of topographical variance are assumed. In other words, it assumes that the source of brain activity consists of one or a small number of dipoles. This method runs a grid search (for each position in the source space) based on the assumed dipole(s) for a time window or a time point, and finds the dipole model parameter that minimizes the difference between the simulated and measured distributions [56]. The resulting Equivalent dipole model is interpreted as the source of cortical activity.

3.3.3 Minimum Norm Estimate (MNE)

This solution to the EEG inverse problem was proposed by M. S. Hämäläinen and R. J. Ilmoniemi in 1984. This is the most generalized approach to estimate 3D source distribution in the absence of any a priori information. The only assumption it takes into account for a solution is that the current distribution should have a minimum overall intensity (smallest L2-norm) [57]. The idea to use minimum norm estimates was first provided by M. S. Hämäläinen et al. in "Interpreting Magnetic Fields of the Brain: Minimum Norm Estimates" [29]. According to Hämäläinen et al. [2], MNE provides best estimates when less a priori information about the source distribution is available. In the aforementioned articles [2, 29], the magnetic fields were discussed at length; however, the same analogy can also be derived for electric fields.

In more detail, MNE uses the Tikhonov regularization by setting $\mathcal{P}(\mathbf{S}) = ||\mathbf{S}||_2^2$. The estimated electrical current of each dipole $\tilde{\mathbf{S}}_{\text{MNE}}$ can be obtained by the minimization of

$$||\mathbf{X}_{\text{est}} - L\mathbf{S}||_2^2 + \lambda||\mathbf{S}||_2^2 \tag{3.33}$$

which alternatively it can be written as:

$$\tilde{\mathbf{S}}_{\text{MNE}} = \left(\mathbf{L}^\intercal\mathbf{L} + \lambda\mathbf{I}_p\right)^{-1}\mathbf{L}^\intercal\mathbf{X}_{\text{est}} \tag{3.34}$$

The above equation is valid since the number of dipoles (p) is larger than the number of the EEG channels (q).

3.3.4 Standardized low resolution brain electromagnetic tomography (sLORETA)

MNE localization approach presented drawbacks related to the magnitutde and depth of the sources. Therefore, a motivation for a new method was arisen. With the introduction of low resolution brain electromagnetic tomography (LORETA) [30] the problem of extensively large errors seemed to be resolved. This method has good accuracy in localization of the sources even when they are located deep in the head.

sLORETA [31] is the first variation on LORETA. This method is based on standardization of current density that is supposed to be estimated for source localization. The current density estimate is carried out using the MNE approach, and it is further standardized using its expected standard deviation, which is hypothesized to have originated exclusively by noise in measurements.

For each dipole k sLORETA yields the estimate of standardized current density power:

$$\tilde{\mathbf{S}}_{\text{MNE, k}}^\intercal\left\{\left[\mathbf{V}_{\tilde{\mathbf{S}}_{\text{pp}}}\right]\right\}^{-1}\tilde{\mathbf{S}}_{\text{MNE, k}} \tag{3.35}$$

where $\tilde{\mathbf{S}}_{\text{MNE, k}} \in \mathbb{R}^{3\times 1}$ is the current density estimate at the k_{th} voxel given by the minimum norm estimate and $\left\{ \left[\mathbf{V}_{\tilde{\mathbf{S}}_{\text{pp}}} \right] \right\} \in \mathbb{R}^{3\times 3}$ is the k_{th} diagonal block of MNE variance.

Chapter 4

Deep Learning

Modern deep learning provides a very powerful framework for supervised learning. By including more layers and units within a layer, a deep network can represent functions of increasing complexity. Deep learning can be used to do most tasks that require translating an input vector to an output vector and are simple to complete quickly, given sufficiently large models and large datasets of labeled training examples. Other tasks that aren't as simple as linking one vector to another, or that are challenging enough to necessitate time for thought and reflection in order to complete the the task, remain beyond the scope of deep learning for now [8].

Deep feedforward networks, also often called feedforward neural networks, are the quintessential deep learning models. The goal of a feedforward network is to approximate some function f^*. For example, for a classifier, $y = f^*(\mathbf{x})$ maps an input $\mathbf{x}$ to a category y. A feedforward network defines a mapping $\mathbf{y} = f(\mathbf{x}; \boldsymbol{\theta})$ and learns the value of the parameters $\boldsymbol{\theta}$ that result in the best function approximation.

These models are called feedforward because information flows through the function being evaluated from $\mathbf{x}$, through the intermediate computations used to define f, and finally to the output $\mathbf{y}$. There are no feedback connections in which outputs of the model are fed back into itself.

Feedforward networks are of extreme importance to machine learning practitioners. They form the basis of many important commercial applications. For example, the convolutional networks used for object recognition from photos are a specialized kind of feedforward network. Feedforward networks are a conceptual stepping stone on the path to recurrent networks, which power many natural language applications.

4.1 Multilayer perceptron (MLP)

Perceptrons were developed [58] in the 1950s and 1960s by the scientist Frank Rosenblatt. Multi layer perceptron (MLP) is a supplement of feed forward neural network. It consists of three types of layers—the input layer, output layer and hidden layer, as shown in Figure 4.1. The input layer receives the input signal to be processed. The required task such as prediction and classification is performed by the output layer. The true computational engine of the MLP is an arbitrary number of hidden layers inserted between

the input and output layers. In a MLP, as a feedforward network, the data flows in the forward direction from input to output layer.

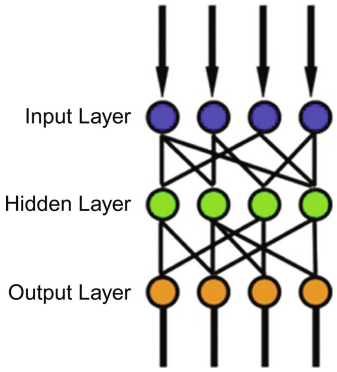


Figure 4.1: Schematic representation of a MLP with single hidden layer.

A directed acyclic graph defines how the functions are combined together in a MLP. Thus, the above MLP can be mathematically described as:

$$f(\mathbf{x}) = f^{(3)}\left(f^{(2)}\left(f^{(1)}(\mathbf{x})\right)\right) \tag{4.1}$$

where $f^{(1)}$ is the first layer, $f^{(2)}$ is the second layer and so on.

4.1.1 Artificial Neuron

The most fundamental unit of a deep neural network is called an artificial neuron, which takes an input, processes it, passes it through an activation function (see subsection 4.1.2) like the ReLU, return the activated output.

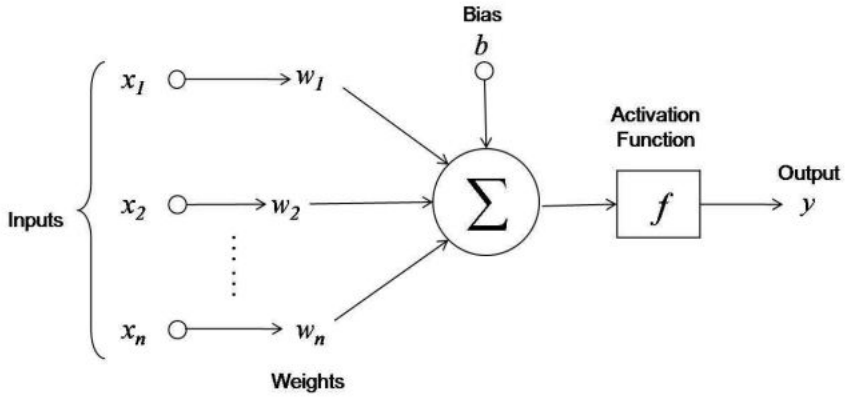


Figure 4.2: An Artificial neuron

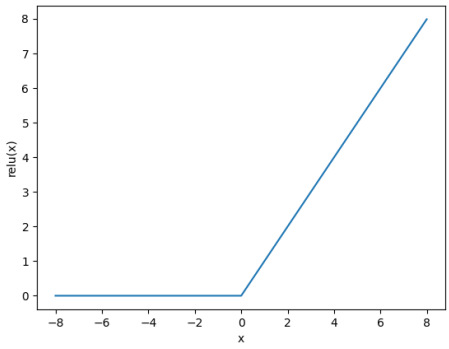
The output of the above neuron is:

$$y = f\left(\sum_{i=1}^{n} x_i \cdot w_i + b_i\right) \tag{4.2}$$

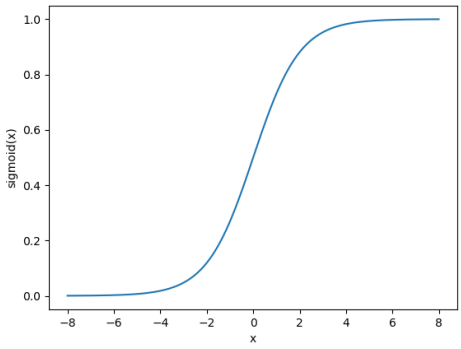
A layer in the MLP shown in Figure 4.1 consists of multiple neurons. The values of both vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^\intercal$ and $\mathbf{b} = [b_1, b_2, \ldots, b_n]^\intercal$ are assigned during the training process.

4.1.2 Activation Function

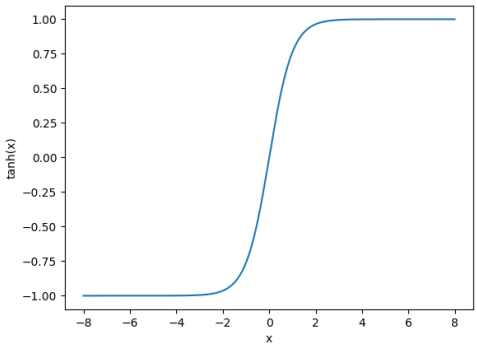
The activation functions are inspired by human neural firing, i.e., it either fires or does not. They have a crucial influence in a neural network. The activation functions are used to create nonlinear input-output relations. This nonlinearity, paired with a large number of neural nodes and layers, resembles the structure of a human brain, which is why it's termed a neural network. There are many activation functions (some of them presented in Figure 4.3):



(a) relu



(b) sigmoid



(c) tanh

Figure 4.3: Some activation functions

The role of the activation function is to transform and abstract the data into a more classifiable plane. Generally, the data is very tightly clustered; it is the job of the activation function which transforms the data into a different plane which helps in observing the effects of different dimensions in the given problem [59].

4.1.3 Gradient-Based Learning

The process of designing and training a neural network is similar to that of any other machine learning model with gradient descent [8]. The most significant distinction between linear models (such as SVM) and neural networks is that a neural network's nonlinearity leads most interesting loss functions to become non-convex. This means that neural networks are usually trained by using iterative, gradient-based optimizers that merely drive the loss function to a very low value, rather than the linear equation solvers used to train linear regression models or the convex optimization algorithms with global convergence guarantees used to train logistic regression or SVMs.

Convex optimization converges starting from any initial parameters (in theory—in practice it is very robust but can encounter numerical problems). Stochastic gradient descent applied to non-convex loss functions has no such convergence guarantee, and is sensitive to the values of the initial parameters. For feedforward neural networks, it is important to initialize all weights to small random values. The biases may be initialized to zero or to small positive values.

Of course, we can use gradient descent to train models like linear regression and support vector machines, and this is frequent when the training set is extremely large. Training a neural network is similar to training any other model from this perspective. Computing the gradient is slightly more complicated for a neural network, but can still be done efficiently and exactly.

As with other machine learning models, to apply gradient-based learning we must choose a loss function, and we must choose how to represent the output of the model.

4.1.4 Loss Function

An important aspect of the design of a deep neural network is the choice of the loss function. From the training data, a deep learning neural network learns to map a set of inputs to a set of outputs. We cannot calculate the perfect weights for a neural network; there are too many unknowns. Instead, the learning problem is recast as a search or optimization problem, and an algorithm is used to navigate the space of possible weight settings that the model could employ to create good or adequate predictions.

Typically, a neural network model is trained using the stochastic gradient descent optimization algorithm and weights are updated using the backpropagation of error algorithm. The "gradient" in gradient descent refers to an error gradient. The model with a given set of weights is used to make predictions and the error for those predictions is calculated. The gradient descent algorithm aims to alter the weights so that the following evaluation reduces the error, thereby travelling down the gradient (or slope) of error.

In the context of an optimization algorithm, the function used to evaluate a candidate solution (i.e. a set of weights) is referred to as the objective function. We might try to maximize or minimize the objective function, which means we're looking for a potential solution with the highest or lowest score.

Typically, with neural networks, we seek to minimize the error. As such, the objective function is often referred to as a cost function or a loss function and the value calculated by the loss function is referred to as simply "loss."

Importantly, the choice of loss function is directly related to the activation function used in the output layer of your neural network. These two design elements are connected. Think of the configuration of the output layer as a choice about the framing of your prediction problem, and the choice of the loss function as the way to calculate the error for a given framing of your problem. [8]. In particular, some widely used configurations are:

- *Regression Problem*: A problem where you predict a real-value quantity.
  - **Output Layer Configuration:** One node with a linear or relu (depending on the problem) activation unit.
  - **Loss Function:** Mean Squared Error (MSE).
- *Binary Classification Problem*: A problem where you classify an example as belonging to one of two classes. The problem is framed as predicting the likelihood of an example belonging to class one, e.g. the class that you assign the integer value 1, whereas the other class is assigned the value 0.
  - **Output Layer Configuration:** One node with a sigmoid activation unit.
  - **Loss Function:** Cross-Entropy, also referred to as Logarithmic loss.
- *Multi-Class Classification Problem*: A problem where you classify an example as belonging to one of more than two classes. The problem is framed as predicting the likelihood of an example belonging to each class.
  - **Output Layer Configuration:** One node for each class using the softmax activation function.
  - **Loss Function:** Cross-Entropy, also referred to as Logarithmic loss.

4.2 Convolutional neural network (CNN)

Convolutional networks also known as convolutional neural networks or CNNs, are a specialized kind of neural network for processing data that has a known, grid-like topology. Examples include time-series data, which can be thought of as a 1D grid taking samples at regular time intervals, and image data, which can be thought of as a 2D grid of pixels. Convolutional networks have been tremendously successful in practical applications. *Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers* [8].

4.2.1 Layers of a CNN

As shown in Figure 4.4, in CNN the features are detected through the use of filters which are also known as kernels [60]. A filter is just a matrix of values, called weights that are trained to detect specific features. The purpose of the filter is to carry out the convolution operation, which is an element-wise product and sum between two matrices.

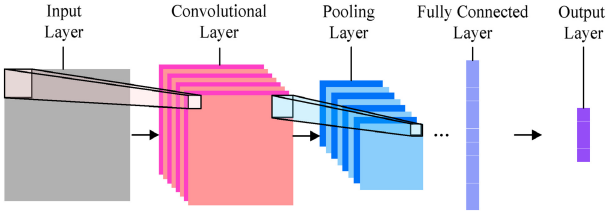


Figure 4.4: Schematic representation of a convolutional neural network

The training of the CNN is fastened by reducing the amount of redundancy present in the input feature. Hence, the amount of memory consumed by the network is also reduced. One common method to achieve this is max pooling, in which, a window passes over input data and the maximum value within the window is pooled into an output matrix. The algorithm is made efficient for feature extraction by concatenating multiple convolution layers and max pooling operations. The data is processed through these deep layers to produce the feature maps which is finally converted into a feature vector by passing through a MLP. This is referred to as a Fully-Connected Layer that performs high-level reasoning in the developed model.

If the k-th feature map at a given layer is represented as $\mathbf{h}^k$ whose filters are determined by the weights $\mathbf{W}^k$ and bias $\mathbf{b}^k$, then the feature map is obtained as follows for tanh activation function:

$$\mathbf{h}_{ij}^k = \tanh\left(\left(\mathbf{W}^k * \mathbf{x} \right)_{ij} + \mathbf{b}^k \right) \tag{4.3}$$

where, $\mathbf{x}$ is the input and ij is the pixel of an image.

The output of the fully-connected layer carries probabilities of each class or the estimation for a regression problem. The weights updation and optimization of the algorithm is done through the back propagation of gradients [60].

4.2.2 Benefits of CNNs

Convolution leverages three important ideas that can help improve a machine learning system: **sparse interactions**(also referred to as **sparse connectivity** or **sparse weights**), **parameter sharing** and **equivariant representations** [8]. Moreover, convolution provides a means for working with inputs of variable size. We now describe each of these ideas in turn.

Matrix multiplication by a matrix of parameters with a separate parameter indicates the interaction between each input unit and each output unit is used in traditional neural network layers. This means that every output unit depends on every input unit. Convolutional networks, however, typically have **sparse interactions**. This is accomplished by making the kernel smaller than the input. When processing a picture, for example, the input image may have thousands or millions of pixels, but we can discover small, important features like edges using kernels that only take up tens or hundreds of pixels. This means that we need to store fewer parameters, which both reduces the memory requirements of the model and improves its statistical efficiency. It also means that computing the output requires fewer operations. If there are m inputs and n outputs, then matrix multiplication requires $m \times n$ parameters and the algorithms used in practice have $\mathcal{O}(m \times n)$ runtime (per sample). It is possible to get good performance on the machine learning problem while keeping k several orders of magnitude less than m in many practical situations. For graphical demonstrations of sparse connectivity, see Figure 4.5.

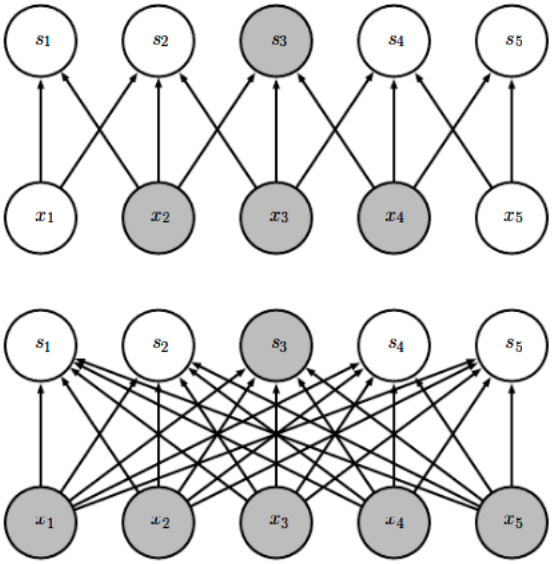


Figure 4.5: Sparse connectivity.

In the Figure 4.5, we highlight one input unit, x_3, and also highlight the output units in output **s** that are affected by this unit. (Top)When **s** is formed by convolution with a kernel of width, only three outputs are affected **x**. (Bottom)When **s** is formed by matrix multiplication, connectivity is no longer sparse, so all of the outputs are affected by x_3.

Parameter sharing refers to using the same parameter for more than one function in a model. In a traditional neural net, each element of the weight matrix is used exactly once when computing the output of a layer. It is multiplied by one input element and then never used again. As a synonym for parameter sharing, one can say that a network has **tied weights**, because the value of the weight applied to one input is tied to the value of a weight applied elsewhere. Each member of the kernel is used at every point of the input in a convolutional neural network (except perhaps some of the boundary pixels, depending on the design decisions regarding the boundary). The parameter sharing used by the convolution operation means that rather than learning a separate set of parameters for every location, we learn only one set. This does not affect the runtime offorward propagation ($\mathcal{O}(\mathrm{k} \times \mathrm{n})$) but it does further reduce the storage requirements of the model to k parameters. For a graphical depiction of how parameter sharing works, see Figure 4.6.

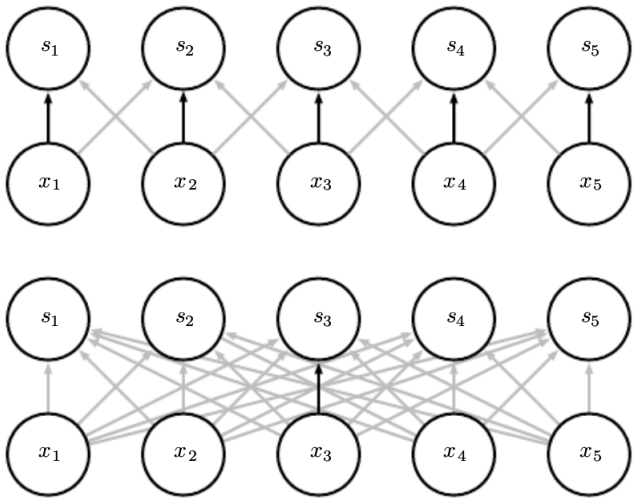


Figure 4.6: Parameter sharing

In the Figure 4.6, black arrows indicate the connections that use a particular parameter in two different models. (Top)The black arrows indicate uses of the central element of a 3-element kernel in a convolutional model. This single parameter is used across all input locations because of parameter sharing. (Bottom)The single black arrow indicates the use of the central element of the weight matrix in a fully connected model. This model has no parameter sharing so the parameter is used only once.

Finally, in the case of convolution, the particular form of parameter sharing causes the layer to have a property called **equivariance** to translation. To say a function is equivariant means that if the input changes, the output changes in the same way. Specifically, a function $f(x)$ is equivariant to a

function g if:

$$f(g(x)) = g(f(x)) \tag{4.4}$$

In the case of convolution, if we let g be any function that translates the input, i.e., shifts it, then the convolution function is equivariant to g. For example, let I be a function giving image brightness at integer coordinates. Let g be a function mapping one image function to another image function, such that $I' = g(I)$ the image function with $I'(x, y) = I(x - 1, y)$. This moves all the pixels of I one unit to the right. If we apply this transformation to I, then apply convolution, the result will be the same as if we applied convolution to I' , then applied the transformation g to the output.

Convolution creates a form of timeline when processing time series data, showing when distinct features arise in the input. If we move an event later in time in the input, the exact same representation of it will appear in the output, just later in time. Similarly with images, convolution creates a 2-D map of where certain features appear in the input. If we move the object in the input, its representation will move the same amount in the output. This is useful for when we know that some function of a small number of neighboring pixels is useful when applied to multiple input locations. For example, when processing images, it is useful to detect edges in the first layer of a convolutional network. The same edges appear more or less everywhere in the image, so it is practical to share parameters across the entire image.

4.3 Improving the way neural networks learn

Successfully applying deep learning techniques requires more than just a good knowledge of what algorithms exist and the principles that explain how they work. In this section, a suite of techniques is introduced which can be used to improve on our implementation of backpropagation, and so improve the way our networks learn.

4.3.1 Batch normalization

Batch normalization [61] is one of the most exciting recent innovations in optimizing deep neural networks and it is actually not an optimization algorithm at all. Instead, it is a method of adaptive reparametrization, motivated by the difficulty of training very deep models.

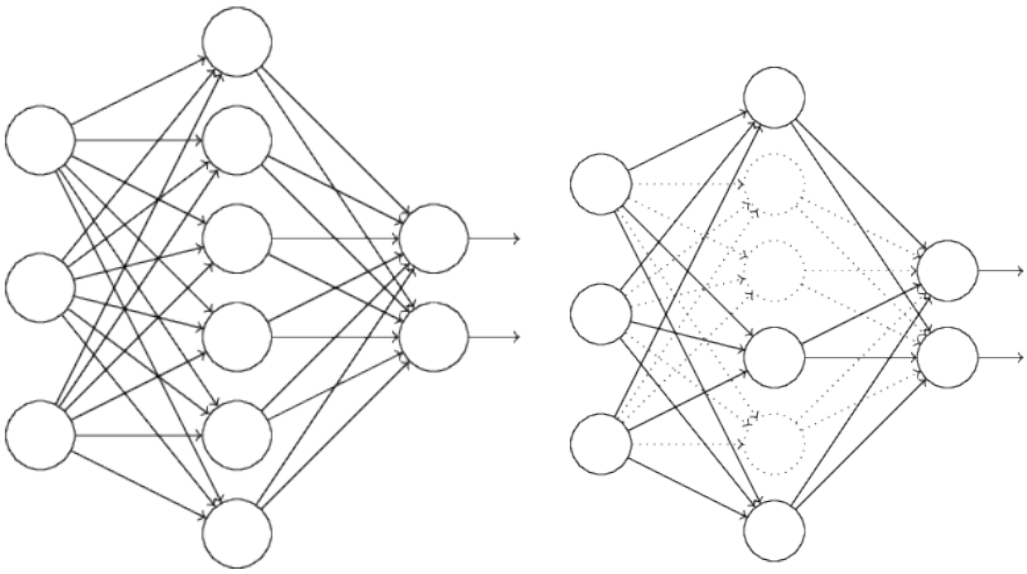
Batch normalization provides an elegant way of reparametrizing almost any deep network. The reparametrization significantly reduces the problem of coordinating updates across many layers. Batch normalization can be applied to any input or hidden layer in a network. Let $\mathbf{H}$ be a minibatch of activations of the layer to normalize, arranged as a design matrix, with the activations for each example appearing in a row of the matrix. To normalize (standardize) $\mathbf{H}$, we replace it with:

$$\mathbf{H}' = \frac{\mathbf{H} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{4.5}$$

where μ is a vector containing the mean of each unit and σ is a vector containing the standard deviation of each unit.

4.3.2 Dropout

Dropout is a radically different technique for regularization [62]. In dropout we modify the network itself. Suppose we're trying to train the network in Figure 4.7a:



(a) No dropout **(b)** With dropout

Figure 4.7: Graphical explanation of the dropout layer

In particular, suppose we have a training input $\mathbf{x}$ and corresponding desired output $\mathbf{y}$. Ordinarily, we'd train by forward-propagating x through the network, and then backpropagating to determine the contribution to the gradient. With dropout, this process is modified. We start by randomly (and temporarily) deleting a proportion of the hidden neurons (usually 50%) in the network, while leaving the input and output neurons untouched. After doing this, we'll end up with a network along the following lines. Note that the dropout neurons, i.e., the neurons which have been temporarily deleted, are still ghosted in Figure 4.7b. We forward-propagate the input $\mathbf{x}$ through the modified network, and then backpropagate the result, also through the modified network.

By repeating this process over and over, our network will learn a set of weights and biases. Of course, those weights and biases will have been learnt under conditions in which a proportion of the hidden neurons were dropped out. When we actually run the full network, all the neurons will be active. To compensate for that, we halve the weights outgoing from the hidden neurons.

Heuristically, when we dropout different sets of neurons, it's rather like we're training different neural networks. And so the dropout procedure is

like averaging the effects of a very large number of different networks. The different networks will overfit in different ways, and so, hopefully, the effect of dropout will be to reduce overfitting.

Chapter 5

Implementation

In this thesis, we develop and present a novel deep learning solution to localize neural sources, and assess its accuracy and robustness with both real [9] and simulated EEG-recordings. To model more realistically the geometrical and electromagnetic features of the head, a six-compartment head model was used [10]. Then, we solve the forward problem using the Finite Element Method (FEM) [11] because of its high flexibility to accurately model the electromagnetic field propagation in geometrically challenging inhomogeneous and anisotropic head volume conductors such as the human head [12, 13]. With the solution of the Forward problem we generate simulated EEG-recordings to train a Convolutional neural network (CNN) that solves the inverse problem. While we focus on EEG, the same approaches are directly extendable to MEG, enabling a portable and affordable solution to source localization.

We utilized Python and Matlab to implement the above functionality. In particular, we used the libraries: Duneuro [63], Tensorflow [64], Keras [65], Pandas [66] and Fieldtrip [67].

5.1 Forward Modeling

In order to produce a more realistic result of the Inverse Problem, the Forward Problem must be modeled as realistically as possible. Thus, we used a source space with $p = 50,460$ dipoles and a set of $q = 74$ electrodes based on the 10-10 system. The precision of our source space is in the order of millimeters. The dipoles must be imported to a realistic head model. We used a six-compartment head model [10] with the following compartments:

- skin
- skull compacta
- skull spongiosa
- cerebrospinal fluid
- white matter
- gray matter

Furthermore, we set the isotropic conductivities as follows:

- for skin: 0.43 S/m
- for skull compacta: 0.0031 S/m
- for skull spongiosa: 0.01116 S/m
- for cerebrospinal fluid: 1.79 S/m

The anisotropic conductivity tensors for the compartments gray and white matter were determined as described in [10].

Moreover, the source space in the center of the gray matter compartment without restrictions to source orientation. This ensures that all dipoles are located inside the gray matter and sufficiently far away from the neighboring tissue compartments to fulfill the so-called *Venant* condition, that is, for each dipole, the closest mesh node should only belong to elements, which are labeled as gray matter [10].

5.2 Solution of the Forward Problem

As described in the Chapter 3, in the Forward problem we have to build a head volume conductivity model for describing how the electrical signals of the brain signal source are transmitted to the scalp electrodes. In other words, given the neural sources and the recorded potentials at the EEG-electrodes the leadfield matrix must be calculated (see Equation (3.2)).

Despite the computational cost of the Finite Element Method, we choose to utilize FEM because of its high flexibility to accurately model the electromagnetic field propagation in geometrically challenging inhomogeneous and anisotropic head volume conductors such as the human head [12, 13]. With the help of Duneuro [63] we computed the EEG leadfield matrix (and thereby solved the forward problem) using the standard (CG-) FEM approach [11] with the *Venant* source model using the transfer matrix approach [68].

The *Venant* direct approach is used due to its high accuracy and computational efficiency [33]. When using FEM, an increased number of elements is needed for modeling all the complex geometries of the modeled volume but the computational costs can be reduced significantly due to the use of sparse matrices [68]. The use of the transfer matrix approach enables to solve the Forward Problem for a head model with millions of elements.

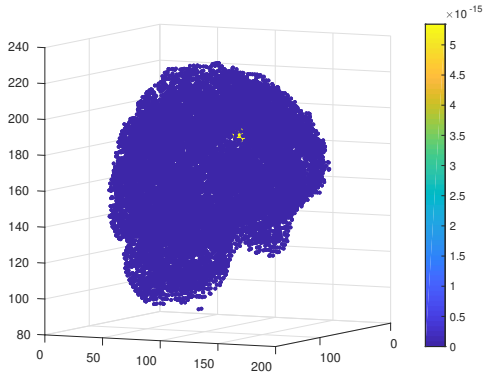
In FEM, three-dimensional elements such as hexahedra or tetrahedra are used to model the head volume. In this thesis, hexahedral meshes are used because it is easier and faster to obtain the mesh out of the labeled MRI. For hexahedral meshes, the voxels in the labeled MRI volume can be directly converted to mesh elements. In contrast to tetrahedral meshes, this direct conversion is not possible and further elaboration of the segmented model is necessary to create tissue compartments that are not intersect each other.

5.3 Simulation of EEG Recordings

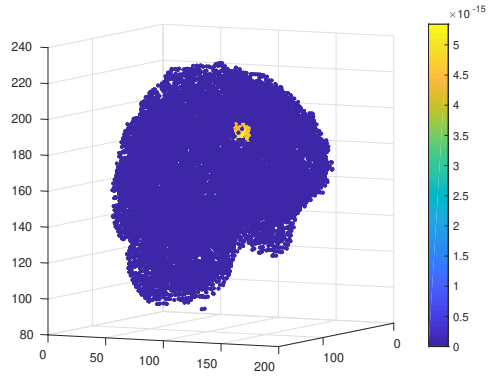
Having solved the Forward Problem, we can now proceed to the solution of the Inverse Problem, that is, to estimate the most possible source activity which could generate the scalp EEG signals. Since the Inverse Problem is solved using a neural network, we must generate the training data. To train a deep learning model and evaluate its performance on source localization, we need to know the ground truth of the underlying neural sources generating EEG data. Since this information is unavailable in real EEG measurements of human participants, we performed simulations with an actual EEG sensor array, a realistic anatomy and source configurations.

Our neural network is designed to operate on single time instances of EEG data with a single source. Thus, we simulate the electrical activity as described in [24].

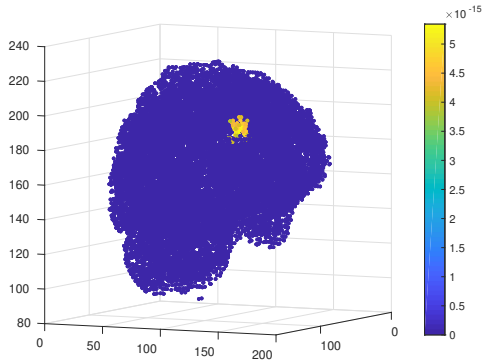
Each simulation contained one dipole cluster, which can be considered as a smooth region of brain activity. A dipole cluster was generated by selecting a random dipole in the cortical source model and then adapting a region growing approach as described in [69]. In more detail, we recursively included all surrounding neighbors starting from a single seeding location, thereby creating a larger source extent with each iteration. The number of iterations define the neighborhood order s, where the first order s_1 entails only the single selected dipole. Each seed location was assigned a dipole moment between 5 and 10 Nano-Ampere Meter (nAm). The neighboring dipoles were assigned attenuated moments based on the distance to seeding location. The attenuation followed a Gaussian distribution with a mean of the seeding dipole moment and a standard deviation of half the radius of the dipole cluster, yielding smooth source patches. The generation of this spatial pattern is graphical represented in the source model:



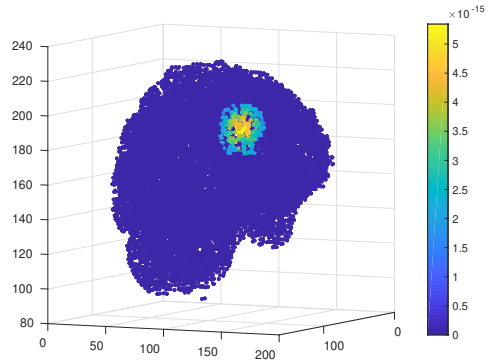
(a) Step 1 (seed dipole)



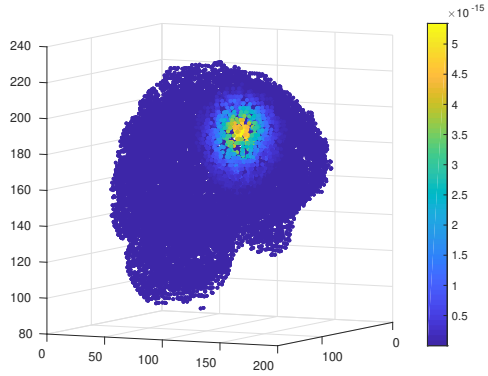
(b) Step 2



(c) Step 3



(d) Step 5



(e) Final step (simulated electrical activation)

Figure 5.1: Neural source simulation

The simulated electrical activation $\mathbf{S} \in \mathbb{R}^p$ of $p = 50,460$ dipoles was then projected the leadfield matrix $\mathbf{L} \in \mathbb{R}^{q \times p}$ in order to calculate the potential at the 74 EEG electrodes $\mathbf{X}_{\text{est}} \in \mathbb{R}^q$. To generate realistic training data, we added Gaussian white noise at a specific SNR level (see Equation (3.4)). In more detail, the SNR is set based on the power of the neural sources to be 15 dB:

$$\text{SNR} = 10 \cdot \log\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{5.1}$$

where P_{signal} is the power of the simulated electrical activation and P_{noise} is the power of the additive Gaussian noise (unknown variable in the above equation). From the simulated 74 channel measurements, one can create their topography. In the figure below, the topography of the simulated source (see Figure 5.1) is shown. The dots represent the EEG-electrodes placed on the scalp.

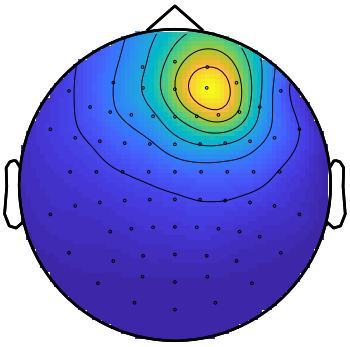


Figure 5.2: Simulated topography

With the aforementioned algorithm 300, 000 training samples (electrical currents of neural sources and their respective topographies) were produced. Since each simulation contained one dipole cluster, we produce single source EEG-snapshots.

5.4 CNN for the EEG Inverse Problem

Deep neural networks have been considered a potential tool for inverse problems. Theoretically, it has been proven that deep neural networks are able to fit any distribution. In practice, many network structures have been proposed, aiming to obtain multi-scale source information from the original data. In this way, the deep learning models can adapt to more complex distributions and thus have greater potential to generate more realistic source distributions compared with traditional numerical algorithms. Nevertheless, the potential of deep learning is still a new and growing in Source Analysis. We therefore focus on developing a novel deep learning framework for real-time EEG source localization.

5.4.1 Architecture

The design and training of the neural network (see Figure 5.3) was accomplished using the Tensorflow [64] and Keras [65] libraries, which are based on Python 3. In particular, the CNN takes as input an EEG topography and predicts the location of the electrical activation (that is responsible for the EEG recorded signals) in the three-dimensional source space. Thus, our CNN can

be mathematically described as:

$$\Phi : \mathbb{R}^{67\times 67} \rightarrow \mathbb{R}^3 \tag{5.2}$$

since a topography is a 67×67 picture (see Figure 5.2) and the output layer is a vector with the x,y,z coordinates in the three-dimensional source space of the activated seed dipole (see Section 5.3). Training was accomplished on the Google Colab Pro.

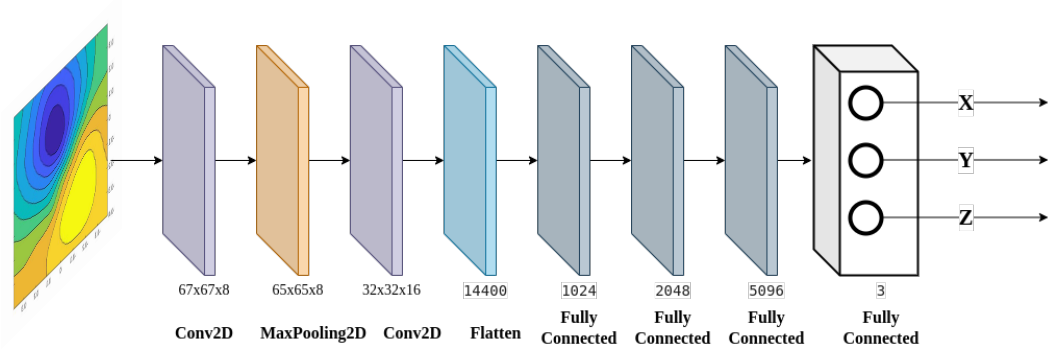


Figure 5.3: CNN architecture

The architecture of our CNN (see Figure 5.3) is inspired by three pioneer studies [22, 23, 24] (see Section 5.5 for differences between the neural networks). The input topography passes through two 2D-Convolution layers which are separated with a max pooling layer. Moving on, there are three fully-connected layers with 1024,2048,5096 neurons respectively. Finally, the output layer is also a fully connected layer with only three neurons that correspond to the coordinates in the three-dimensional source space. In more detail,

- The first convolution layer has only 8 filters F_i , $i = \{1, \ldots, 8\}$ of size 3×3
- The second layer is a two-dimensional max pooling layer with a pool size of 2×2
- The third layer is a convolution layer with 16 filters F_i , $i = \{1, \ldots, 16\}$ of size 3×3
- The output of the third layer is a tensor $\mathbf{G} \in \mathbb{R}^{30\times 30\times 16}$ which is reshaped from the fourth layer (flatten) to a vector $\tilde{\mathbf{g}} \in \mathbb{R}^{14400}$ to enable a connection to the next fully connected layer. The flattened vector $\tilde{\mathbf{g}}$ consists of 14400 output nodes and each node is connected to every neuron of the following fully connected layer.
- The next three fully connected layers have $1024, 2048, 5096$ neurons respectively and each neuron of each layer is connected to every neuron of the following layer. Moreover, after each one these layer there are also Batch Normalization [61] and Dropout [62] layers. The rate of the Dropout layer is set to 0.25.

- All the 5096 neurons of the seventh layer are connected to all three neurons of the output layer. The output layer estimates the position of the electrical activation inside the head. Thus, each layer corresponds to a coordinate in the three-dimensional source space.

To recapitulate, our CNN maps a 67×67 topography to the location of the dipole that is responsible for the EEG recorded signal. Thus, it solves the EEG Inverse Problem.

5.4.2 Deep Network Training

Convolution filters, weights and biases were optimized using the Stochastic Gradient Descent (SGD) algorithm [70] with a learning rate $\lambda = 0.001$ and batch size of 32. The proposed convolutional neural network was trained with datasets of 300 thousands simulated snapshots, generated at a fixed SNR level (15 dB), yet, as discussed in Chapter 6 the trained CNN operates well in a wide range of SNR levels.

We tried out various loss functions for regression problems. The loss functions below yielded the better results. The term $\mathbf{y}$ denotes the true values, $\tilde{\mathbf{y}}$ the predicted values and N the length of both the vectors with actual and predicted values. In more detail,

- **Mean Square Error (MSE)**: The loss is the mean overseen data of the squared differences between true and predicted values:

$$L(\mathbf{y}, \tilde{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \tilde{y}_i\right)^2 \tag{5.3}$$

- **Mean absolute error (MAE)**: The loss is the mean overseen data of the absolute differences between true and predicted values:

$$L(\mathbf{y}, \tilde{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^{N} \left|y_i - \tilde{y}_i\right| \tag{5.4}$$

- **Huber loss**: is a combination of the mean squared error function and the absolute value function. The intention behind this is to make the best of both worlds.

$$L_\delta(y_i, \tilde{y}_i) = \begin{cases} \frac{1}{2} \cdot \left(y_i - \tilde{y}_i\right)^2, & \text{if } \left|y_i - \tilde{y}_i\right| < \delta \\ \delta \cdot \left(\left|y_i - \tilde{y}_i\right| - \frac{1}{2}\delta\right), & \text{otherwise} \end{cases} \tag{5.5}$$

We ultimately decided to use the Mean absolute error (MAE) (5.4) as it shown to perform well on image segmentation tasks and allowed for a fast convergence of our CNN compared to others.

We decided to use Rectified Linear Units (ReLUs) [71] as activation functions after each layer. ReLUs have shown to exhibit the best performance in

our preliminary tests compared to alternatives (e.g. sigmoid, tanh). ReLU activation functions were used in the output layer. Typical CNNs for regression use linear activation functions, however, since predictions are by definition non-negative in our application, ReLUs appeared to us as an appropriate alternative.

To recapitulate, we trained our CNN (see Figure 5.3) with the the Stochastic Gradient Descent (SGD) algorithm [70] as optimization algorithm, the Mean absolute error (MAE) as a loss function over 500 epochs with a batch size of 32. The training was accomplished on Google Colab Pro and it took 12 hours.

5.4.3 Novelties of the proposed CNN

While we focus on EEG, the same approaches are directly extendable to MEG, enabling a portable and affordable solution to source localization. The conceptual novelties of the proposed Convolutional Neural Network (CNN) are :

1. Localization accuracy in real EEG-recordings. We tested our CNN on three different subjects (see Chapter 6).
2. Better performance from other deep learning approaches as we not only used FEM [11] but also a six-compartment head model [10] with 50,460 dipoles.
3. Very low computational time (once trained) for the estimation of the source location.
4. Because of the fact that the proposed CNN takes as input a topography and not the EEG-signal, one can use a topography generated from a different EEG-recording system than ours.

5.5 A comparison of our proposed CNN with other neural network-based inverse solutions

The architecture of our CNN is inspired by three pioneer studies [22, 23, 24]. More specifically, we compare our CNN with the other neural networks in terms of:

- **Forward Modeling**: The ConvDip-study [24] used a head model with 5,124 dipoles and solved the Forward Problem with the Boundary Element Method. In the DeepMEG [22], they solved the Forward Problem using the BrainStorm [72]. Finally, both in the ESBN [23] and in this study the Forward Problem is solved using the Finite Element Method.
- **Neural Network Input Layer:** All neural networks use a snapshot of the EEG-recordings. Both DeepMEG [22] and ESBN [23] use the recorded values from the EEG-electrodes as input. In comparison, ConvDip [24]

and our CNN use a topography which is generated from the EEG-recorded values.

- **Neural Network Output Layer:** Even though, all the aforementioned networks use a Fully Connected layer in the output, there is a key difference. In ConvDip [24] the output layer consists of 5,124 neurons and each one of them corresponds to a dipole in the source space. Thus, ConvDip estimates the amplitude of each dipole. A source space with only 5,124 dipoles is not as detailed and realistic as our source space with 50,460 dipoles (>10 times larger). In comparison, in both the DeepMEG [22] and our CNN, the output layer consists of three neurons (for a single source) and thereby computes the source coordinates in the three-dimensional source space.

- **Depth of the network:** All the pioneer neural nets have less layers than ours. More specifically, while both DeepMEG [22] and ESBN [23] have four layers, ConvDip [24] has three. The proposed CNN has 7 (without the flatten layer). Moreover, ConvDip in comparison with our CNN does not use pooling layers.

- **Evaluation on real EEG-recordings:** While our CNN, ESBN [23] and ConvDip [24] were tested on real EEG-recordings, DeepMeg [22] is not. Finally, our CNN is tested on three different subjects with different anatomies (see Chapter 6) whereas, both the ESBN [23] and ConvDip [24] were tested in recordings from a single participant.

Chapter 6

Evaluation

In this Chapter, we evaluate the performance of our CNN and compare it to state-of-the-art inverse algorithms, namely sLORETA and Dipole Scanning. Firstly we assess the performance of the neural network using simulated data (as described in Chapter 5). Note that the evaluation set was not part of the training set of our CNN, hence it is unknown to the model. We further validate the performance of our network on real data recorded during the doctoral of Dr. Antonakakis [33]. The real EEG-recordings are from three different subjects and thereby three different anatomies.

We use the localization error [55] as metric to quantify EEG source localization performance. The localization error (LE) can be quantified as the Euclidean distance between truly activated source r_{true} and the reconstructed peak source r_{peak} in three dimensional source space.

$$\text{LE} = ||r_{\text{true}} - r_{\text{peak}}||_2 \tag{6.1}$$

The localization error requires the location of the seed dipole (r_{true}) that generated the recorded activity (from the EEG-electrodes). In the case of simulated data, the seed dipole of the dipole cluster (see Section 5.3) is known and the LE can be used. Many studies use the Spatial Dispersion (SD) as a metric to access the quality of the source localization in real-data. As described in [73], Spatial Dispersion SD is quantified as:

$$\text{SD} = \frac{\sum_{k=1}^{p} d_{jk}\tilde{\mathbf{x}}_k}{\sum_{k=1}^{p} \tilde{\mathbf{x}}_k} \quad , \quad d_{jk} = ||r_k - r_j||_2 \tag{6.2}$$

which is the l_1 norm of the distance vector $\mathbf{d} = \left\{d_{jk}\right\}_{k=1,\dots,p}$, containing the distances d_{jk} between the position of the peak reconstruction amplitude r_j and all sources in the source space $\{\mathbf{r}_k\}_{k=1,\dots,p}$, weighted by the source estimate $\tilde{\mathbf{x}}$ normalized to its l_1 norm.

Spatial Dispersion requires to estimate the amplitude of each dipole as ESBN [23] and ConvDip [24]. Thus, it can not be used in our neural network as it predicts the location of the seed dipole in the three-dimensional source space. Moreover, SD does not account for localization at all and it only measures the extent of the electrical activation.

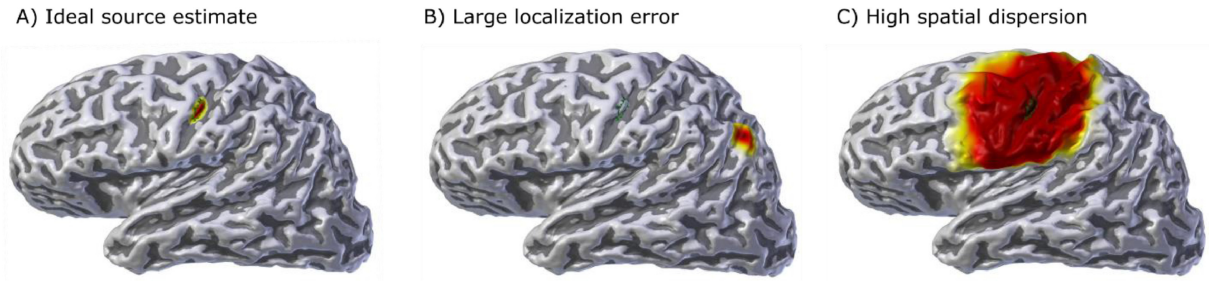


Figure 6.1: Visualization of spatial resolution metric. Adapted from [73].

In the above figure, the contour outlines the activated patch. What is shown in the subfigure:

A) is the ideal source estimate

B) is the source estimate with large localization error (quantified as high LE) but low spatial dispersion (SD)

C) is the source estimate with high spatial dispersion (high SD)

Hence, Spatial Dispersion is not reliable if it's the only metric which accesses the performance of the model in real EEG recordings. As it will be described in Section 6.2, we can evaluate the performance of our convolutional neural network with the predicted location as the EEG-recordings are product of a very specific experiment and thereby we can anticipate where the seed dipole should be located.

6.1 Evaluation with simulated data

Firstly, we assessed the performance of our convolutional neural network using simulated data as described in the data generation workflow (see Section 5.3). Note that the evaluation set was not part of the training set of our CNN, hence it is unknown to the model.

6.1.1 Evaluation for various SNR levels

To assess localization accuracy in different realistic scenarios, we conducted simulations with different SNR levels. While our CNN is trained with 15 dB SNR data, in the evaluation we used SNR levels ranging from -10 dB to 20 dB. We compared the performance of our CNN against the popular sLORETA [31] and Dipole Scan localization algorithms. For each SNR $5,000$ samples (EEG and sources data) were simulated. The localization error for each SNR is shown in the Figure 6.2.

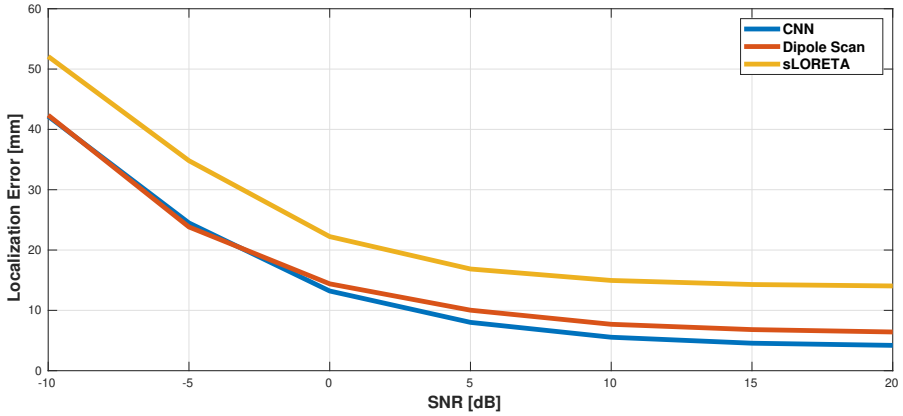
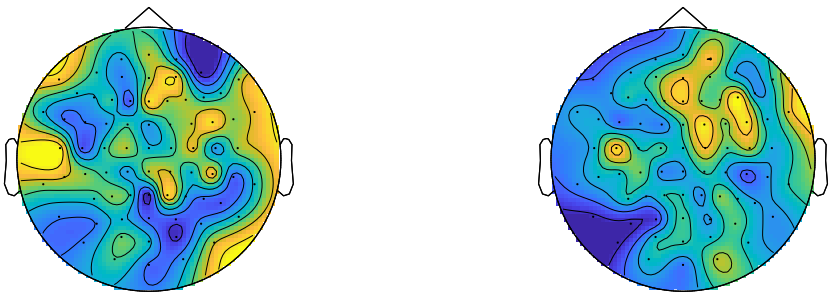


Figure 6.2: Localization Error for various SNR levels

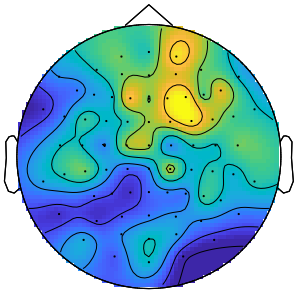
It appears that, our CNN consistently outperformed both the sLORETA and Dipole Scanning localization algorithms with the exception of low SNR levels (< -4.5 dB) where Dipole Scanning had a slightly better accuracy. As expected, all methods consistently improved their localization performance with increasing SNR values.

For visualization purposes, we display a simulated topography for each signal-to-noise ratio. All the simulation have the same seed dipole and different SNR.

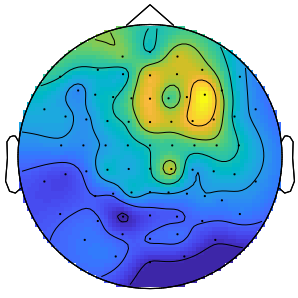


(a) SNR $= -10$dB

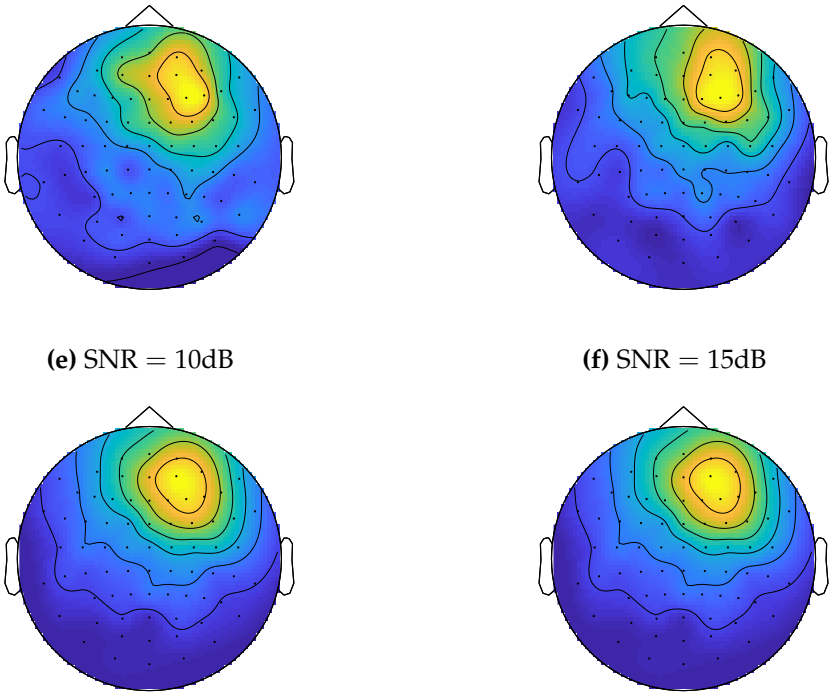
(b) SNR $= -5$dB



(c) SNR $= 0$dB



(d) SNR $= 5$dB



(e) SNR = 10dB

(f) SNR = 15dB

(g) SNR = 20dB

(h) Without noise

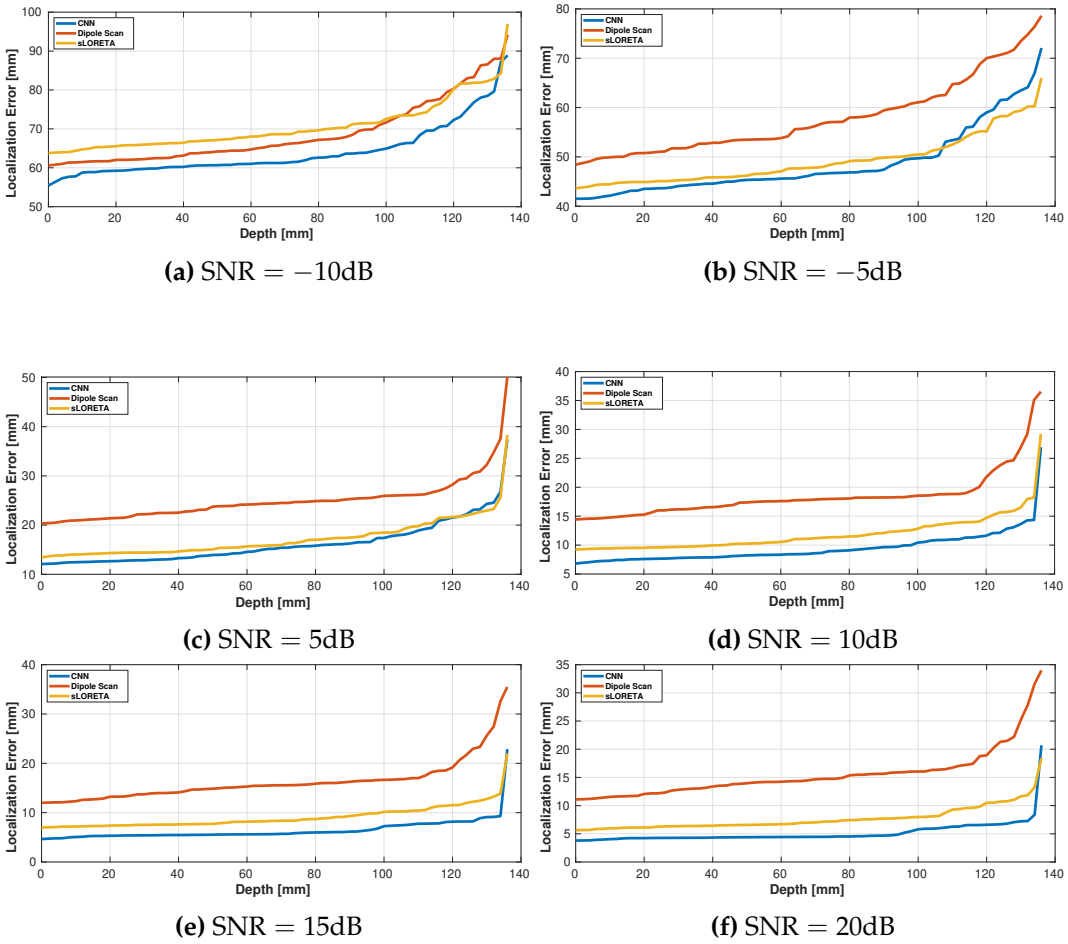
Figure 6.3: Topographies for various SNR levels

Even though, our CNN is trained with 15dB SNR data (see Figure 6.3f), it operates well and outperforms the traditional approaches in a wide range of SNR levels. Moreover, the Figures 6.2 and 6.3 prove the generalization ability of our convolutional neural network.

Generalization is defined as the ability of an algorithm to perform well on unseen examples. In statistical learning terms an algorithm $\mathcal{A} : \mathcal{X} \rightarrowtail \mathcal{Y}$ is learned using a training dataset $\mathcal{S} = \left\{ (x_1, y_1) \dots (x_N, y_N) \right\}$ of size N where $x_i \in \mathcal{X}$ is a data sample and $y_i \in \mathcal{Y}$ is the corresponding label (for example, source location coordinates).

6.1.2 Influence of the depth of the source

The influence of the depth of the simulated source is tested. The larger the depth is, the weaker the source affects the EEG-signal. Thus, the larger the depth is, the more difficult the localization is. We compared the performance of the Inverse methods (CNN, sLORETA, Dipole Scanning) for all the depths in the source models and different SNR levels (see Figure 6.4).



(a) SNR $= -10$dB **(b)** SNR $= -5$dB

(c) SNR $= 5$dB **(d)** SNR $= 10$dB

(e) SNR $= 15$dB **(f)** SNR $= 20$dB

Figure 6.4: Localization Error for various SNR levels and depths

A further surprising result is that our CNN does not depend on the depth of the source cluster for high SNR levels. Our neural network is capable to correctly localize even deep sources with nearly no drop-off. In particular, while Dipole Scan has the worst localization results and strongly depends on the depth of the source, both our convolutional neural network and sLORETA slightly depend on the depth of the source cluster. Furthermore, for all Signal-to-noise ratios except -5dB, our CNN yields the less localization error over all the spectrum of the depths. Finally, as expected, as the SNR increases the dependence of the localization methods on the depth of the neural source decrease and thereby they yield better results.

To recapitulate, our CNN outperforms the traditional approaches in terms of reconstructing deep sources (see Figure 6.4) and robustness to noise (see Figure 6.2). Even though our neural network is trained with simulated data which have 15dB SNR, it can localize accurately EEG topographies with various SNR levels. Thus, our CNN has a great generalization ability and it is neither underfitted nor overfitted.

6.2 Evaluation with real data

To evaluate the performance of our CNN and the other inverse algorithms in a realistic set-up, we used data of real EEG recordings. As a comparison with other neural networks that solve the Inverse Problem, we used EEG recordings from three different subjects and thereby three different anatomies. To the best of the author's knowledge, all existing Deep Learning studies for Source Localization (including this thesis) are trained with data generated from a specific MRI and thus a particular anatomy of the human head.

Pioneer studies suggest that their neural networks may work for various anatomies with a transfer learning approach. In particular, a transfer learning approach where the neural network is trained on one subjects' anatomy and fine-tuned for each additional subject with new training data of the individual anatomies. Fine tuning could be achieved by replacing the output-layer, lowering the learning rate and retraining for only few epochs.

Nevertheless, we managed to solve the problem of individual brain anatomies using FSL Registration [15, 16, 17]. More specifically, the MRI of each subject is interpolated with the spline interpolation using the FSL tool. Hence, the predicted three-dimensional coordinates of our neural network are referring to the interpolated MRI. To this extent, we overcome the suggestion of transfer learning and we present a novel solution without raising the problem of computation time as the spline interpolation needs approximately three to five minutes.

The general idea of a spline is this: on each interval between data points, represent the graph with a simple function. The simplest spline is something very familiar; it is obtained by connecting the data with lines. Since linear is the most simple function of all, linear interpolation is the simplest form of spline. The next simplest function is quadratic. If we put a quadratic function on each interval then we should be able to make the graph a lot smoother. If we were really careful then we should be able to make the curve smooth at the data points themselves by matching up the derivatives. This can be done and the result is called a quadratic spline. Using cubic functions or 4^{th} degree functions should be smoother still. So, where should we stop? There is an almost universal consensus that cubic is the optimal degree for splines.

6.2.1 Data Acquisition

The real EEG data were recorded during the PhD [33] of Dr. Antonakakis and specifically are from this paper [10]. The EEG recordings of one of the participants in the study can be found here [9].

As described in [10], five *right-handed* subjects (three of them were 27 one 32 and the last 49 years old; 2 females) participated in the study. Somatosensory evoked potentials (SEP) and fields (SEF) were simultaneously acquire in a magnetically shielded room using 80 AgCl sintered ring electrodes (EASYCAP GmbH, Herrsching, Germany, 74 EEG channels plus additional six channels to detect eye movements). For the detection of cardiac activity, electrocardiography (ECG) was additionally measured. Prior to the

measurements, the electrode positions of the EEG cap were digitized using a Polhemus device (FASTRAK, Polhemus Incorporated, Colchester, VT).

A MAGNETOM Prisma 3.0 T (Release D13, Siemens Medical Solutions, Erlangen, Germany) was used for the acquisition of MRI datasets. A 3D-T1-weighted (T1w) was measured with fast gradient-echo pulse sequence (TFE) using water selective excitation to avoid fat shift (TR/TE/FW = 2300/3.51 ms/8°, inversion prepulse with TI =1.1 s, cubic voxels of 1 mm edge length); 3D-T2w turbo spin-echo pulse sequence (TR/TE/FA = 3200/408 ms/90°, cubic voxels, 1 mm edge length) and DTI using an echo-planar imaging sequence (TR/TE/FA = 9500/79 ms/90°, cubic voxels, 1.89 mm edge length), with one volume with diffusion sensitivity b = 0s / mm^2 (i.e., flat diffusion gradient) and 20 volumes with b = 1,000s / mm^2 in different directions, equally distributed on a sphere.

An additional volume with flat diffusion gradient, but with reversed spatial encoding gradients was scanned and utilized for susceptibility artifact correction. During T1w-MRI measurement, gadolinium markers were placed at the same nasion, left and right distal outer ear canal positions for landmark-based registration of EEG to MRI.

6.2.2 Stimulation

The somatosensory stimulation was conducted with the following experiment [10]. The median nerve at the *right wrist* was stimulated with monophasic square-wave electrical pulses having a duration of 0.5 ms. The stimulus strength was increased until a clear movement of the thumb was visible. This type of stimulation is abbreviated as EW stimulation.

The data were acquired with a sampling rate of 1,200 Hz and online filtered with a 300 Hz low pass filter. The experiment consisted of 1,200 trials. The stimulus onset asynchrony (SOA) varied randomly from 350-450 ms to avoid habituation and to allow obtaining clear prestimulus intervals for signal-to-noise ratio (SNR) determination.

6.2.3 Expected Localization

Principally, in real EEG-recordings, as opposed to simulated, we cannot know the location of the dipole cluster that produced the recorded electrical activity. As we lack ground truth for real EEG sources, we cannot use the localization error to quantify the performance of the Inverse Methods. However, as described in Section 6.2.1, the EEG-recordings were generated by a very specific experiments with particular parameters, thus, we can know the approximate location of the seed dipole.

In the study [74] participated five *right-handed* male volunteers ages 21 to 36. The participants were stimulated with an air-puff-derived tactile stimulator which provides a light, superficial pressure stimulus to the skin surface. About 40 points in the right hemibody (including the tongue, lips, hand, arm, trunk, leg, and foot) were stimulated in a randomly determined order (see Figure 6.5).

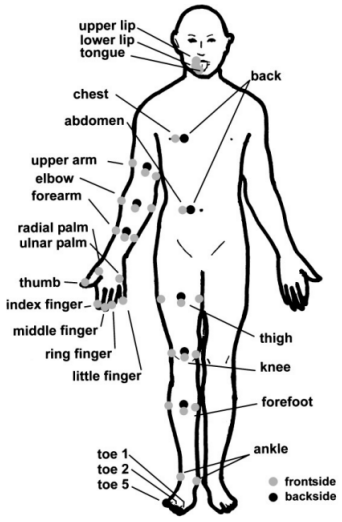


Figure 6.5: Forty-three stimulated points of the right hemibody.
Adapted from [74]

Having collected EEG recordings for each stimulation, Source analysis, based on a single moving equivalent current dipole (ECD) model n a spherical volume conduction, were applied to magnetic field distribution studies. The location (x, y, and z positions), orientation, and dipole moment of the best-fitting single ECD were estimated for each stimulation.

Overall, an ECD following chest stimulation were located in the left post-central gyrus. The ECD location to the tongue stimulation was the most inferior, followed by the lips, fingers, and arm, which were gradually shifted to superior and medial, along the central sulcus (as shown in Figure 6.6).

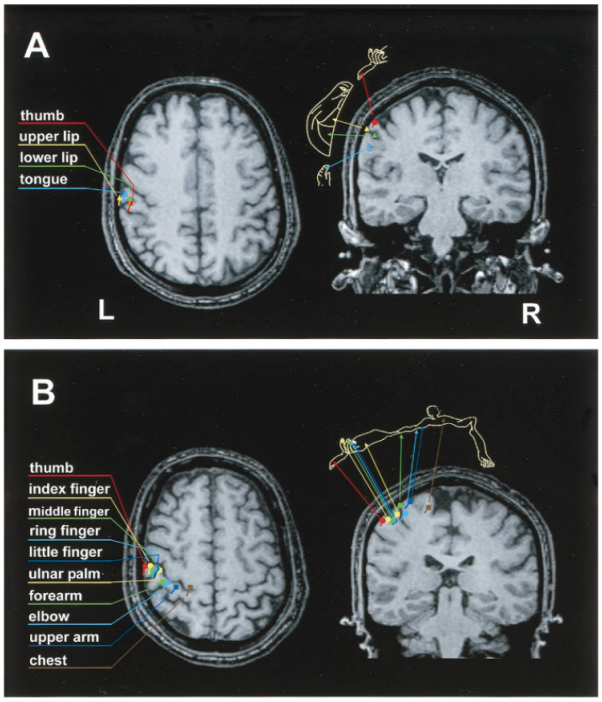


Figure 6.6: Expected localization of each stimulus. Adapted from [74]

Furthermore, the localizations regarding the right thumb and index are further verified in [75]. In this study, twenty participants underwent tactile stimulation in all the fingers (D1 to D5 where D1 is the thumb, D2 the index, etc.) of the dominant right hand.

In more detail, studies have revealed the cytoarchitectonic subdivisions of primary sensory cortex (S1), namely areas 3a, 3b, 1 and 2, outline the cortex in the postcentral gyrus (see Figure 6.7).

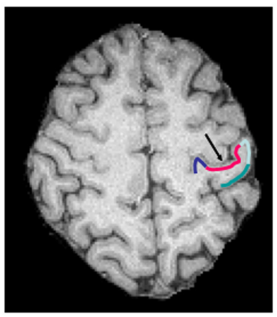


Figure 6.7: Areas of S1 as defined in cytoarchitectonic studies. Adapted from [75]

In Figure 6.7, area 3a occupies the fundus of the central sulcus (dark blue), area 3b the anterior wall of the postcentral gyrus (red), area1 its crown (light blue) and area 2 its posterior wall (green). The black arrow indicates the central sulcus.

The activation in the contralateral somatosensory cortex during tactile stimulation of the fingers of the right hand is shown in the Figure 6.8. The

first column shows transverse anatomical image with z-coordinate indicated. Subsequent columns show the activation patterns in S1 overlayed on magnified T1-weighted images for each finger. The location of the peak voxel in area 3b is indicated by blue crosshairs.

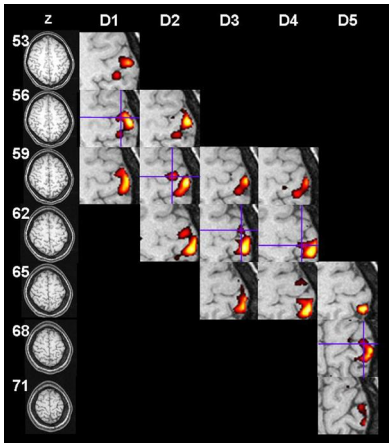


Figure 6.8: Activation on the S1. Adapted from [75]

These two experiments[74, 75] are pertinent to ours because:

- The age of participants in both studies is in the same range.
- The participants in both studies are right-handed.
- The stimulus of the right wrist is implicitly included in both [74] and [75], as the stimulus of the right wrist produces the stimulation of both the thumb and the index.

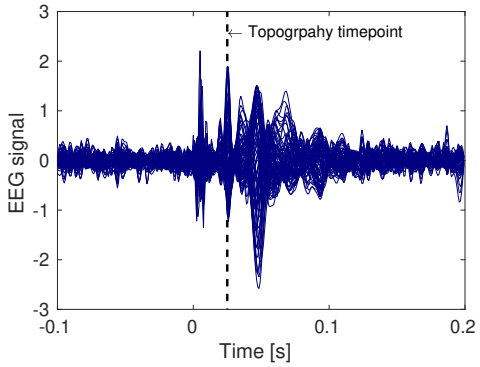
Moreover, the results from both [74] and [75] are trustworthy as they have been verified from numerous subjects and brain anatomies. Therefore, we can use the produced localization as references.

To recapitulate the information from both studies [74, 75], the stimulation of the right wrist in a right-handed subject, generates an electrical activation in the primary sensory cortex (S1) and in particular in area 3b which is the anterior wall of the postcentral gyrus (red in Figure 6.7). In the next section, we will show that our CNN localize this type of EEG recording to the expected location.

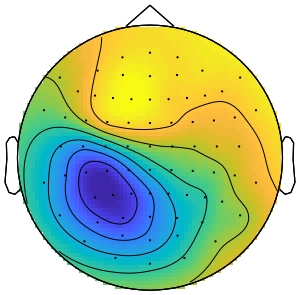
6.2.4 Results

To evaluate the performance of our CNN and the other inverse algorithms in a realistic set-up, we used the real EEG data. The data were recorded as described in Section 6.2.1. Based on the preprocessed clean EEG signal, we compare the performance of our neural network with the traditional numerical methods (sLORETA and Dipole Scan). We tested our convolutional neural network in three out of the five participants in [10].

The *first subject* is a 49 years old male. The recorded EEG signal and the produced topography (which is the input of our CNN) are shown in the Figure 6.9



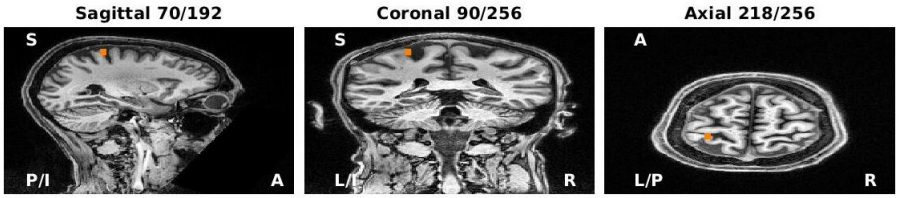
(a) EEG Signal



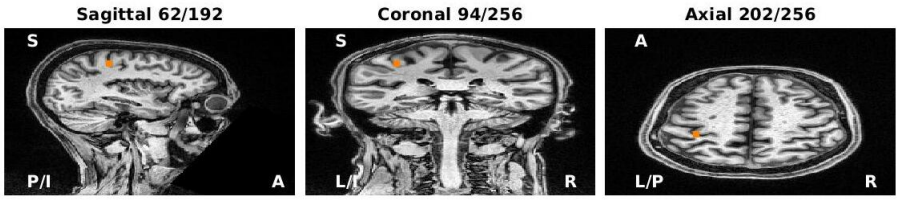
(b) Topography

Figure 6.9: Recorded data from subject 1

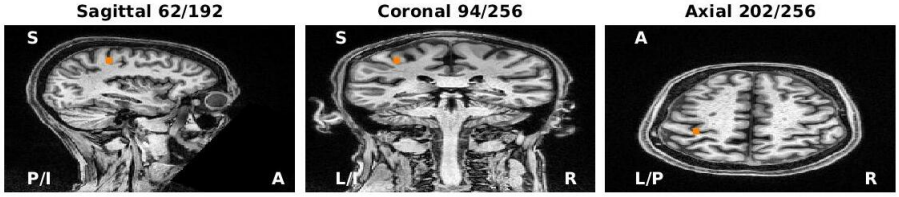
The vertical black dotted-line in the left plot represents the timepoint in which the topography was created. In this timepoint,the measurements of the 74 EEG-electrodes are the input in the traditional algorithms while the topography is the input of our CNN. The source localization on the MRI of each method is shown in the Figure 6.10.



(a) CNN



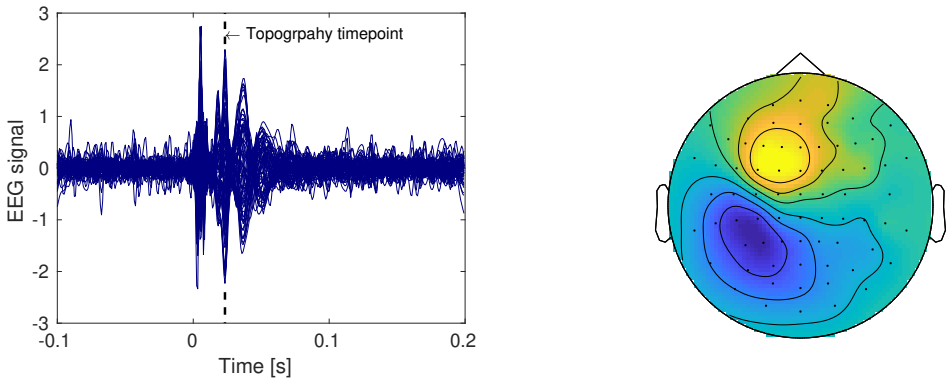
(b) sLORETA



(c) Dipole Scan

Figure 6.10: Source Localization for subject 1

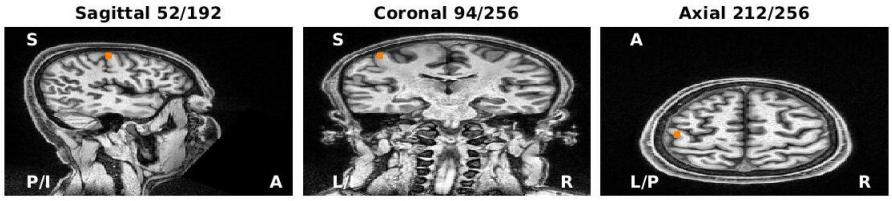
The *second subject* is a 27 years old female. The recorded EEG signal and the produced topography (which is the input of our CNN) are shown in the Figure 6.11



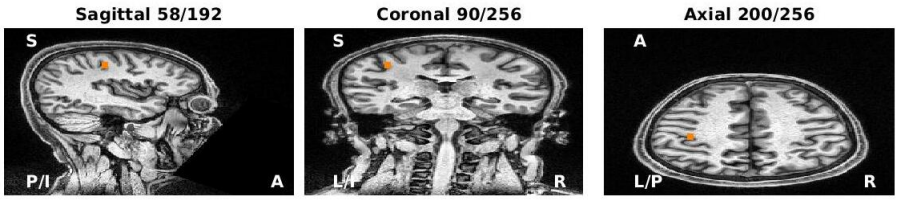
(a) EEG Signal **(b)** Topography

Figure 6.11: Recorded data from subject 2

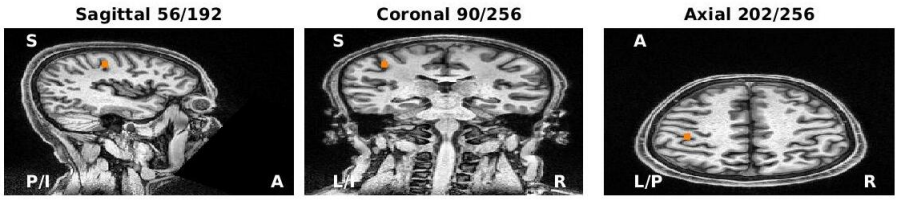
The source localization on the MRI of each method is shown in the Figure 6.12.



(a) CNN



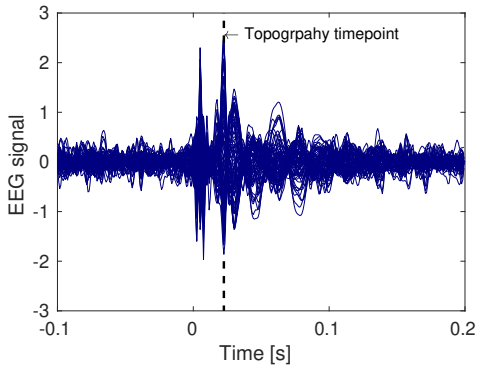
(b) sLORETA



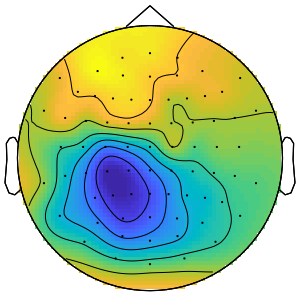
(c) Dipole Scan

Figure 6.12: Source Localization for subject 2

The *third subject* is a 27 years old male. The recorded EEG signal and the produced topography (which is the input of our CNN) are shown in the Figure 6.13



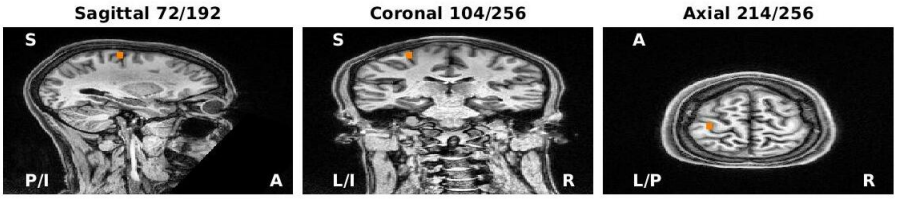
(a) EEG Signal



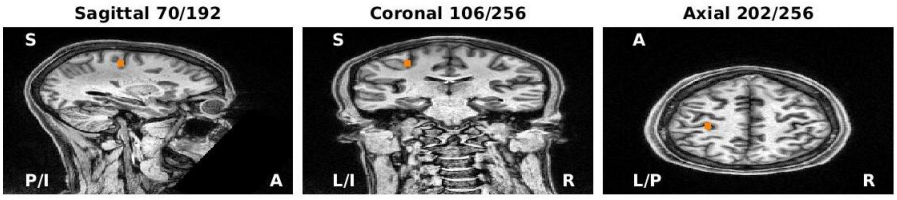
(b) Topography

Figure 6.13: Recorded data from subject 3

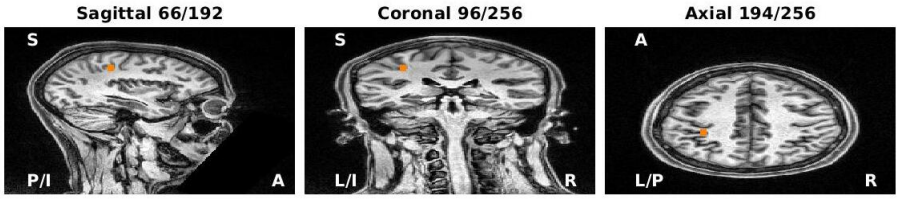
The source localization on the MRI of each method is shown in the Figure 6.14.



(a) CNN



(b) sLORETA



(c) Dipole Scan

Figure 6.14: Source Localization for subject 3

As it can be seen from Figures 6.10,6.12,6.14, our CNN generates the expected localization (see Section 6.2.3) ,for all the participants, which is the primary sensory cortex (S1). Moreover, while sLORETA also correctly localizes the electrical activity, Dipole Scan estimates inaccurately a deeper (in the head) location. Furthermore, our CNN predicts a location in area 3b whereas sLORETA localizes the EEG measurements in area 2 of the S1 (see Figure 6.7). Based on Section 6.2.3, area 3b is the optimal location for this experiment.

Finally, we compare the execution time of each algorithm. In particular, table 6.1 shows the mean localization time of each method for all three subjects.

Method	Localization time (seconds)
CNN	0.17 $\pm$0.001
sLORETA	0.38 $\pm$0.02
Dipole Scan	0.57 $\pm$0.03

Table 6.1: Mean Execution time

The aforementioned computational time does not take into account the pre/post-processing steps of each method, it only measures the time each algorithm requires. In more detail, both sLORETA and Dipole Scan require the leadfield matrix of each subject, thus they need the solution of the Forward Problem. Since we solve the Forward Problem with FEM, it needs approximately 40 minutes to be computed. In comparison, our neural network needs the interpolated MRI (approx 3 minutes).

6.3 A comparison of our results with other neural network-based localizations

In the Section 5.5, we compared our approach with [22, 23, 24] in terms of implementation. We now compare these studies in terms of results. In particular, we discuss the localizations in both the simulated and the real data:

- **Simulated:** In the DeepMEG [22] study, the performance in the simulated data was evaluated with the Localization Error (LE) as in this study. Both ConvDip [24] and ESBN [23] utilized many more metrics to access the performance of their networks. More specifically, ConvDip utilized Area under the ROC curve (ROC), LE, Mean squared error (MSE), Normalized mean squared error (nMSE). ESBN utilized LE, area under curve (AUC) and Spatial Dispersion (SD). Finally, all the networks including ours, outperformed the traditional numerical algorithms in simulated data.

- **Real data:** Firstly, DeepMEG [22] is not evaluated with real data. In comparison, our CNN, ESBN [23] and ConvDip [24] are evaluated with real data. In particular, both ESBN and ConvDip are evaluated with real data from one participant. Moreover, ConvDip and ESBN use spatial dispersion to evaluate the performance of their networks on real data. As it was explained before (see Figure 6.1), spatial dispersion (sd) is not reliable when it is the only metric. In contrast, our CNN is validated with EEG recordings from three participants and with the location of the predicted source not the sd. Our network has localized the P20/N20 component at the subject-specific Brodmann area 3b, making possible the EEG localization for various brain anatomies

Chapter 7

Discussion

To recapitulate, in this thesis we discussed the problem of real-time source localization.

We have presented a CNN that aims to find the location of the seed dipole that is responsible for the recorded data from the EEG-electrodes placed on the scalp. Initially, to model more realistically the geometrical and electromagnetic features of the head, a six-compartment (skin, skull compacta, skull spongiosa, cerebrospinal fluid, white and gray matter) head model was used [10]. Then, we solve the forward problem using the Finite Element Method (FEM) [11] because of its high flexibility to accurately model the electromagnetic field propagation in geometrically challenging inhomogeneous and anisotropic head volume conductors such as the human head [12, 13]. Having calculated the leadfield matrix with the help of Duneuro [63], we generate simulated EEG recordings. Those synthetic EEG-recordings was later on used in order to train the proposed CNN. Finally, we access the accuracy and robustness of our CNN with both real [9, 10] and simulated EEG-recordings.

7.1 Result Summary & Contribution

As we previously discussed, we addressed the Inverse Problem of Source Analysis by combining a six-compartment head model, the Finite Element Method, EEG data generation and a convolutional neural network. Our results showed that our method can correctly localize real EEG recordings in any brain anatomy.

Even though, the traditional analytical approaches can also localize EEG recordings in any anatomy, the existing Deep Learning approaches have been tested in a single participant, thereby a single brain anatomy. Our neural network is tested in three different subjects. Each subject has its own anatomy, source space and leadfield matrix. Thus, we can correctly localize EEG data regardless of the geometrical features of the head.

Also, within the limited scope of our experiments, our approach seems to compare very favorably to approaches as Dipole Scan and sLORETA. Our neural network outperforms the aforementioned methods. In particular, to assess localization accuracy in different realistic scenarios, we conducted simulations with different SNR levels. Even though our CNN is trained with 15 dB SNR data, it can correctly localize EEG data in a wide range of SNR levels.

Moreover, our model does not depend on the depth of the source cluster, i.e., our CNN is capable to correctly localize even deep sources with nearly no drop-off.

Moreover, our CNN yields very low computational time once trained. More specifically, it needs 0.17 seconds to localize EEG-data in an extremely detailed source space with 50,460 dipoles. In comparison, sLORETA needs 0.38 seconds and Dipole Scan needs 0.57 seconds in the same detailed source space. The aforementioned computational time does not take into account the pre/post-processing steps of each method, it only measures the time each algorithm requires. In more detail, both sLORETA and Dipole Scan require the leadfield matrix of each subject, thus they need the solution of the Forward Problem. Since we solve the Forward Problem with FEM, it needs approximately 40 minutes to be computed. In comparison, our neural network needs only interpolation which requires around 3 minutes.

Finally, due to the fact that our neural network is trained with a realistic six-compartment head model and an extremely detailed source space, it can accurately localize the recorded electrical activity regardless the anatomy of the brain.

7.2 Drawbacks of our Approach

Despite the advantages of our approach, it is worthy to note the drawbacks. First of all, the dipole orientations in the head model need to be constrained, as the free orientation can greatly influence the source imaging results. Our method does not estimate dipole orientations because the simulated EEG-data do not have this piece of information. Thus, our model does not learn about orientations.

Furthermore, our approach is under the assumption that brain activity is always smooth. In particular, we only generate synthetic source data that follow a Gaussian distribution. Hence, during training, our CNN learns that the brain activity is always smooth and follows a Gaussian distribution.

Moreover, our model does not estimate the location of the seed dipoles in a distributed dipole model with more than one source. Our CNN has three neurons in the output layer and each one of them corresponds to a coordinate in the three-dimensional source space. Thus, it can only predict the location of a single dipole in the source space.

Finally, even though we managed to overcome the problem of different brain anatomies and we tested our model in three subjects, we have to solve the problem of the inter-subject variability of skull conductivity which causes the most significant influence on the EEG localization [18]. More specifically, while each head model has its own skull conductivity, our approach assumes that all the subjects have the same skull conductivity. Additional research is important to offer a trainned model with representatitive head tissue conductivities beyond one subject.

7.3 Future Work

Based on the drawbacks of our approach, there are still many disciplines worth exploring:

- **Dipole Orientations**: As mentioned before, our model does not estimate the orientation of the reconstructed source. Thus, it is worth investigating an architecture that can predict the orientation of the source cluster.
- **Distribution of brain activity**: In our work, we generate synthetic source data that follow a Gaussian distribution. In reality that is not the case, thus we could simulate samples of brain activity which a proportion of them still follows a Gaussian distribution and the rest of them follow a Random Markov Field.
- **Distributed dipole model**: Our approach predicts the location of a single dipole in the source space. Hence, it cannot estimate the location of multiple sources from a topography of the EEG data. Although further research is required, our CNN could operate in a distributed dipole model with a simple extension. In particular, in the output layer, it could have three neurons for each source. For instance, if the location of two sources was to be predicted then the neural network would have six output neurons. The first three would correspond to the coordinates of the first dipole while the rest to the coordinates of the second dipole.

Bibliography

[1] Saeid Sanei and Jonathon A Chambers. *EEG Signal Processing*. en. Newy York: Wiley-Interscience, May 2013.

[2] Matti Hämäläinen et al. "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain". In: *Rev. Mod. Phys.* 65 (2 1993), pp. 413–497. DOI: 10.1103/RevModPhys.65.413. URL: https://link.aps.org/doi/10.1103/RevModPhys.65.413.

[3] *MEG: An introduction to methods.* MEG: An introduction to methods. New York, NY, US: Oxford University Press, 2010, pp. xii, 436–xii, 436. ISBN: 978-0-19-530723-8 (Hardcover). DOI: 10.1093/acprof:oso/9780195307238.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780195307238.001.0001.

[4] R. Kleiner et al. "Superconducting quantum interference devices: State of the art and applications". In: *Proceedings of the IEEE* 92.10 (2004), pp. 1534–1548. DOI: 10.1109/JPROC.2004.833655.

[5] M.S. Hamalainen and J. Sarvas. "Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data". In: *IEEE Transactions on Biomedical Engineering* 36.2 (1989), pp. 165–171. DOI: 10.1109/10.16463.

[6] Jae-Hyun Cho et al. "Influence of the head model on EEG and MEG source connectivity analyses". In: *NeuroImage* 110 (2015), pp. 60–77. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2015.01.043. URL: https://www.sciencedirect.com/science/article/pii/S1053811915000683.

[7] *World Health Organization Epilepsy*. https://www.who.int/news-room/fact-sheets/detail/epilepsy. Accessed: 2022-02-07.

[8] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[9] Maria Carla Piastra et al. *The WWU DUNEuro reference data set for combined EEG/MEG source analysis*. The research related to this dataset was supported by the German Research Foundation (DFG) through project WO1425/7-1 and the EU project ChildBrain (Marie Curie Innovative Training Networks, grant agreement 641652). June 2020. DOI: 10.5281/zenodo.3888381. URL: https://doi.org/10.5281/zenodo.3888381.

[10] Marios Antonakakis et al. "The effect of stimulation type, head modeling, and combined EEG and MEG on the source reconstruction of the somatosensory P20/N20 component". In: *Human Brain Mapping* 40.17 (2019), pp. 5011–5028. DOI: https://doi.org/10.1002/hbm.24754. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.24754. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.24754.

[11] Carsten Wolters et al. "Geometry-Adapted Hexahedral Meshes Improve Accuracy of Finite-Element-Method-Based EEG Source Analysis". In: *IEEE transactions on bio-medical engineering* 54 (Aug. 2007), pp. 1446–53. DOI: 10.1109/TBME.2007.890736.

[12] C H Wolters et al. "Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling". en. In: *Neuroimage* 30.3 (Dec. 2005), pp. 813–826.

[13] Johannes Vorwerk et al. "A guideline for head volume conductor modeling in EEG and MEG." In: *NeuroImage* 100 (2014), pp. 590–607. ISSN: 1095-9572. URL: https://www.unboundmedicine.com/medline/citation/24971512/A_guideline_for_head_volume_conductor_modeling_in_EEG_and_MEG_.

[14] M.V. Valueva et al. "Application of the residue number system to reduce hardware costs of the convolutional neural network implementation". In: *Mathematics and Computers in Simulation* 177 (2020), pp. 232–243. ISSN: 0378-4754. DOI: https://doi.org/10.1016/j.matcom.2020.04.031. URL: https://www.sciencedirect.com/science/article/pii/S0378475420301580.

[15] Mark W Woolrich et al. "Bayesian analysis of neuroimaging data in FSL". en. In: *Neuroimage* 45.1 Suppl (Nov. 2008), S173–86.

[16] Stephen M Smith et al. "Advances in functional and structural MR image analysis and implementation as FSL". en. In: *Neuroimage* 23 Suppl 1 (2004), S208–19.

[17] Mark Jenkinson et al. "FSL". en. In: *Neuroimage* 62.2 (Sept. 2011), pp. 782–790.

[18] Marios Antonakakis et al. "Inter-Subject Variability of Skull Conductivity and Thickness in Calibrated Realistic Head Models". en. In: *Neuroimage* 223 (Sept. 2020), p. 117353.

[19] Gert Van Hoey et al. "EEG dipole source localization using artificial neural networks." In: *Physics in medicine and biology* 45 4 (2000), pp. 997–1011.

[20] Sung Chan Jun, Barak A Pearlmutter, and Guido Nolte. "Fast accurate MEG source localization using a multilayer perceptron trained with real brain noise". In: *Physics in Medicine & Biology* 47.14 (2002), p. 2547.

[21] Sung Chan Jun and Barak A Pearlmutter. "Fast robust subject-independent magnetoencephalographic source localization using an artificial neural network". In: *Human brain mapping* 24.1 (2005), pp. 21–34.

[22] Dimitrios Pantazis and Amir Adler. *MEG Source Localization via Deep Learning*. 2020. arXiv: 2012.00588 [eess.SP].

[23] Chen Wei et al. "Edge sparse basis network: a deep learning framework for EEG source localization". In: *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2021, pp. 1–8.

[24] Lukas Hecker et al. "ConvDip: A Convolutional Neural Network for Better EEG Source Imaging". In: *Frontiers in Neuroscience* 15 (June 2021). DOI: 10.3389/fnins.2021.569918.

[25] Alexandra Razorenova et al. "Deep Learning for Non-invasive Cortical Potential Imaging". In: Dec. 2020, pp. 45–55. ISBN: 978-3-030-66842-6. DOI: 10.1007/978-3-030-66843-3_5.

[26] Rui Sun et al. "SIFNet: Electromagnetic Source Imaging Framework Using Deep Neural Networks". In: *bioRxiv* (2020). DOI: 10.1101/2020.05.11.089185. eprint: https://www.biorxiv.org/content/early/2020/07/09/2020.05.11.089185.full.pdf. URL: https://www.biorxiv.org/content/early/2020/07/09/2020.05.11.089185.

[27] Roberta Grech et al. "Review on solving the inverse problem in EEG source analysis". In: *Journal of NeuroEngineering and Rehabilitation* 5.1 (2008), p. 25. ISSN: 1743-0003. DOI: 10.1186/1743-0003-5-25. URL: https://doi.org/10.1186/1743-0003-5-25.

[28] David Wipf and Srikantan Nagarajan. "A unified Bayesian framework for MEG/EEG source imaging". eng. In: *NeuroImage* 44.3 (2009). 18602278[pmid], pp. 947–966. ISSN: 1095-9572. DOI: 10.1016/j.neuroimage.2008.02.059. URL: https://pubmed.ncbi.nlm.nih.gov/18602278.

[29] Matti S Hämäläinen and Risto J Ilmoniemi. "Interpreting magnetic fields of the brain: minimum norm estimates". In: *Medical & biological engineering & computing* 32.1 (1994), pp. 35–42.

[30] R D Pascual-Marqui, C M Michel, and D Lehmann. "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain". en. In: *Int J Psychophysiol* 18.1 (Oct. 1994), pp. 49–65.

[31] R.D. Pascual-Marqui. "Standardized low resolution brain electromagnetic tomography (SLORETA): Technical details". In: *Methods and findings in experimental and clinical pharmacology* 24 Suppl D (Feb. 2002), pp. 5–12.

[32] Munsif Ali Jatoi et al. "EEG based brain source localization comparison of sLORETA and eLORETA". en. In: *Australas Phys Eng Sci Med* 37.4 (Oct. 2014), pp. 713–721.

[33] Marios Antonakakis. "The effect of experimental and modeling parameters on combined EEG/MEG source analysis and transcranial electric stimulation optimization of somatosensory and epilepsy activity". PhD thesis. 2021. URL: https://www.db-thueringen.de/receive/dbt_mods_00049153.

[34] Shiva Asadzadeh et al. "A systematic review of EEG source localization techniques and their applications on diagnosis of brain abnormalities". In: *Journal of neuroscience methods* 339 (2020), p. 108740.

[35] MUNSIF ALI JATOI. *Brain source localization using EEG signal analysis*. CRC Press, 2019.

[36] Robert Plonsey. "Bioelectric phenomena". In: *Wiley encyclopedia of electrical and electronics engineering* (2001).

[37] *Inner Body*. https://www.innerbody.com/image/nerv02.html. Accessed: 2022-02-14.

[38] Katrina Wendel et al. "EEG/MEG Source Imaging: Methods, Challenges, and Open Issues". In: *Computational Intelligence and Neuroscience* 2009 (2009), p. 656092. ISSN: 1687-5265. DOI: 10.1155/2009/656092. URL: https://doi.org/10.1155/2009/656092.

[39] Munsif Jatoi et al. "A survey of methods used for source localization using EEG signals". In: *Biomedical Signal Processing and Control* 11 (May 2014), 42–52. DOI: 10.1016/j.bspc.2014.01.009.

[40] Sebastian Nagel. "Towards a home-use BCI: fast asynchronous control and robust non-control state detection". PhD thesis. Dec. 2019. DOI: 10.15496/publikation-37739.

[41] Metin Akay. *Wiley Encyclopedia of biomedical engineering*. Wiley-Interscience, 2006.

[42] G H Klem et al. "The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology". en. In: *Electroencephalogr Clin Neurophysiol Suppl* 52 (1999), pp. 3–6.

[43] Herbert H. Jasper. "Report of the committee on methods of clinical examination in electroencephalography". In: *Electroencephalography and Clinical Neurophysiology* 10 (1958), pp. 370–375.

[44] G H Klem et al. "The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology". en. In: *Electroencephalogr Clin Neurophysiol Suppl* 52 (1999), pp. 3–6.

[45] Hans Hallez et al. "Review on solving the forward problem in EEG source analysis". In: *Journal of neuroengineering and rehabilitation* 4.1 (2007), pp. 1–29.

[46] Carsten Wolters and Jan C de Munck. "Volume conduction". In: *Scholarpedia* 2.3 (2007), p. 1738.

[47] Marios Antonakakis. *EEG/MEG source reconstruction of electric wrist/Brailletactile/pneumatotactile somatosensory stimulation using realistic head volume conductor modeling*. Institute for Biomagnetism and Biosignalanalysis, University of Muenster, Germany. URL: http://www.sci.utah.edu/~wolters/PaperWolters/2017/AntonakakisTalkBACI2017.pdf.

[48] Hannah McCann, Giampaolo Pisano, and Leandro Beltrachini. "Variation in Reported Human Head Tissue Electrical Conductivity Values". In: *Brain Topography* 32.5 (2019), pp. 825–858. ISSN: 1573-6792. DOI: 10.1007/s10548-019-00710-2. URL: https://doi.org/10.1007/s10548-019-00710-2.

[49] K. A. Awada et al. "Closed-Form Evaluation of Flux Integrals Appearing in a Fem Solution of the 2D Poisson Equation with Dipole Sources". In: *Electromagnetics* 16.2 (1996), pp. 75–90. DOI: 10.1080/02726349608908463. eprint: https://doi.org/10.1080/02726349608908463. URL: https://doi.org/10.1080/02726349608908463.

[50] Sylvain Baillet. "Forward and Inverse Problems of MEG/EEG". In: Jan. 2014, pp. 1–8. ISBN: 978-1-4614-7320-6. DOI: 10.1007/978-1-4614-7320-6_529-1.

[51] J. Haueisen et al. "The Influence of Conductivity Changes in Boundary Element Compartments on the Forward and Inverse Problem in Electroencephalography and Magnetoencephalography - Der Einfluß der Änderung der Schalenleitfähigkeit bei Randelementemodellen auf die Vorwärtsrechnung und das inverse Problem in Elektroenzephalographie und Magnetoenzephalographie". In: 44.6 (1999), pp. 150–157. DOI: doi:10.1515/bmte.1999.44.6.150. URL: https://doi.org/10.1515/bmte.1999.44.6.150.

[52] Munsif Ali Jatoi et al. "Representing EEG source localization using finite element method". In: *2013 IEEE International Conference on Control System, Computing and Engineering*. IEEE. 2013, pp. 168–172.

[53] Michael Rullmann et al. "EEG source analysis of epileptiform activity using a 1 mm anisotropic hexahedra finite element head model". In: *NeuroImage* 44.2 (2009), pp. 399–410.

[54] Moritz Dannhauer et al. "Modeling of the Human Skull in EEG Source Analysis". In: *Human brain mapping* 32 (Sept. 2011), pp. 1383–99. DOI: 10.1002/hbm.21114.

[55] Felix Lucka et al. "Hierarchical Bayesian inference for the EEG inverse problem using realistic FE head models: depth localization and source separation for focal primary currents". In: *NeuroImage* 61.4 (2012), pp. 1364–1382.

[56] Chungki Lee et al. "Dipole Source Localization of Mouse Electroencephalogram Using the Fieldtrip Toolbox". In: *PLOS ONE* 8.11 (Nov. 2013), pp. 1–9. DOI: 10.1371/journal.pone.0079442. URL: https://doi.org/10.1371/journal.pone.0079442.

[57] Per Christian Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. USA: Society for Industrial and Applied Mathematics, 1999. ISBN: 0898714036.

[58] F. Rosenblatt. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Cornell Aeronautical Laboratory. Report no. VG-1196-G-8. Spartan Books, 1962. URL: https://books.google.ca/books?id=7FhRAAAAMAAJ.

[59] Shi Dong, Ping Wang, and Khushnood Abbas. "A survey on deep learning and its applications". In: *Computer Science Review* 40 (2021), p. 100379. ISSN: 1574-0137. DOI: https://doi.org/10.1016/j.cosrev.2021.100379. URL: https://www.sciencedirect.com/science/article/pii/S1574013721000198.

[60] S. Abirami and P. Chitra. "Chapter Fourteen - Energy-efficient edge based real-time healthcare support system". In: *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*. Ed. by Pethuru Raj and Preetha Evangeline. Vol. 117. Advances in Computers 1. Elsevier, 2020, pp. 339–368. DOI: https://doi.org/10.1016/bs.adcom.2019.09.007. URL: https://www.sciencedirect.com/science/article/pii/S0065245819300506.

[61] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *CoRR* abs/1502.03167 (2015). arXiv: 1502.03167. URL: http://arxiv.org/abs/1502.03167.

[62] Michael A. Nielsen. *Neural Networks and Deep Learning*. misc. 2018. URL: http://neuralnetworksanddeeplearning.com/.

[63] Sophie Schrader et al. "DUNEuro—A software toolbox for forward modeling in bioelectromagnetism". In: *PLOS ONE* 16.6 (June 2021), pp. 1–21. DOI: 10.1371/journal.pone.0252431. URL: https://doi.org/10.1371/journal.pone.0252431.

[64] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[65] Francois Chollet et al. *Keras*. 2015. URL: https://github.com/fchollet/keras.

[66] The pandas development team. *pandas-dev/pandas: Pandas*. Version latest. Feb. 2020. DOI: 10.5281/zenodo.3509134. URL: https://doi.org/10.5281/zenodo.3509134.

[67] Robert Oostenveld et al. "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data". In: *Computational Intelligence and Neuroscience* 2011 (2010), p. 156869. ISSN: 1687-5265. DOI: 10.1155/2011/156869. URL: https://doi.org/10.1155/2011/156869.

[68] Carsten Wolters, Lars Grasedyck, and Wolfgang Hackbusch. "Efficient Computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem". In: *Inverse Problems* 20 (Aug. 2004), pp. 1099–1116. DOI: 10.1088/0266-5611/20/4/007.

[69] Rasheda Arman Chowdhury et al. "MEG Source Localization of Spatially Extended Generators of Epileptic Activity: Comparing Entropic and Hierarchical Bayesian Approaches". In: *PLOS ONE* 8.2 (Feb. 2013), pp. 1–19. DOI: 10.1371/journal.pone.0055969. URL: https://doi.org/10.1371/journal.pone.0055969.

[70] Sebastian Ruder. "An overview of gradient descent optimization algorithms". In: *CoRR* abs/1609.04747 (2016). arXiv: 1609.04747. URL: http://arxiv.org/abs/1609.04747.

[71] Vinod Nair and Geoffrey E. Hinton. "Rectified Linear Units Improve Restricted Boltzmann Machines". In: *ICML*. 2010.

[72] François Tadel et al. "Brainstorm: A User-Friendly Application for MEG/EEG Analysis". In: *Computational Intelligence and Neuroscience* 2011 (2011), p. 879716. ISSN: 1687-5265. DOI: 10.1155/2011/879716. URL: https://doi.org/10.1155/2011/879716.

[73] John G. Samuelsson et al. "Spatial fidelity of MEG/EEG source estimates: A general evaluation approach". In: *NeuroImage* 224 (2021), p. 117430. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2020.117430. URL: https://www.sciencedirect.com/science/article/pii/S1053811920309150.

[74] Akinori Nakamura et al. "Somatosensory Homunculus as Drawn by MEG". In: *NeuroImage* 7.4 (1998), pp. 377–386. ISSN: 1053-8119. DOI: https://doi.org/10.1006/nimg.1998.0332. URL: https://www.sciencedirect.com/science/article/pii/S1053811998903329.

[75] Danielle van Westen et al. "Fingersomatotopy in area 3b: an fMRI-study". en. In: *BMC Neurosci* 5 (Aug. 2004), p. 28.