

COMPRESSED SENSING- AND DEEP LEARNING-BASED BRAIN SOURCE LOCALIZATION OF EEG SIGNALS

BY

ELMUBARAK M. IMAM

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

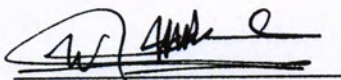
In

ELECTRICAL ENGINEERING

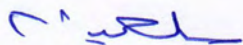
SEPTEMBER 2021

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN- 31261, SAUDI ARABIA
DEANSHIP OF GRADUATE STUDIES

This thesis, written by **ELMUBARAK M. IMAM** under the direction of his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**



Dr. Abdallah S. Alahmari
Department Chairman

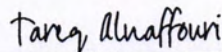


Prof. Suliman S. Al-Homidan
Dean of Graduate Studies

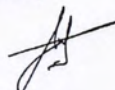
Date



Dr. Mudassir Masood
(Advisor)



Prof. Tareq Y. Al-Naffouri
(Member)



Prof. Mohamed A. Deriche
(Member)



©Elmubarak Imam
2021

*Dedicated to the ones I cherish the most, after ALLAH then his
messenger, my beloved mother and my father, who are always eager to
give their all to help me*

ACKNOWLEDGMENTS

In the name of Allah, the Most Gracious, the Most Merciful

All praise and glory go to Almighty ALLAH (Subhanahu wa Ta'ala), who provided me with the bravery and endurance to carry out this study. Allah's peace and blessings are upon His final Prophet Muhammad (Sallallahu-Alaihi-Wasallam) and all of his companions (Razi-Allahu-Anhum), who dedicated their lives to the success and expansion of Islam.

First and foremost, I would like to express my gratitude to my supervisor Dr. Mudassir Masood for his guidance and knowledge. I was motivated to conduct this study by his inventiveness and drive to tackle complex difficulties. His scholarly qualifications and expertise influenced my research and aided in developing my ideas. Dr. Masood's frequent interactions and numerous moments of mentorship crafted the accomplishment of my thesis. Dr. Masood has always been a source of knowledge in my professional life, in addition to learning. He is a man of ethics, and I learned gratitude, politeness, humility, and tolerance from him, among many other virtues. Because of his goodwill, I never had any trouble communicating with him about any matter, personal or professional, at whatever time of day—he was always friendly and sensitive.

I would like to appreciate the other two thesis committee members, Prof. Tareq Al-Naffouri and Prof. Mohamed Deriche, for their time and thoughtful remarks on my thesis. I would also like to express my special gratitude to Dr. Abdallah Alahmari, who served as our Dean and always directed and looked after me. His great advice enabled me to bring about positive change possible in some of the most critical aspects of my personal and professional life.

Gratitude is due to King Fahd University of Petroleum and Minerals (KFUPM) and the Saudi Ministry of Higher Education for granting me a scholarship to conduct my MS degree. Furthermore, I would like to extend my gratitude deep inside KFUPM to the Center for Energy and Geo Processing (CeGP) for providing the computational power necessary to simulate the Brain forward model and train the proposed Deep Learning modules.

A Special acknowledgment to the Donders Institute for Brain, Cognition and Behaviour whom we have compiled our work around their resources and tools, particularly the Fieldtrip toolbox.

I would like to thank my parents for their everlasting love, sacrifices, inspiration, and prayers, even though I was incredibly neglectful of them. Their unending encouragement and support in my pursuit of higher education enabled me to obtain my MS. I would like to dedicate this thesis to them. They have always offered their all to help me. My appreciation also extends to my sisters and brother, who have always supported me, believed in me, and, most significantly, looked after my parents while I was away. They have always given me a sense of security and assurance. Heartfelt thanks to the space technology specialist Rayan Imam, whom ALLAH (Subhanahu

wa Ta'ala) gifted me with as a sister, for giving continual technical assistance and managerial guidance. Despite her busy schedule, she sincerely devoted a significant amount of her time to me.

I would like to thank all of my friends and colleagues for making this journey so lively. Akram Musaad, Mohammed MohammedAltoum, Asim Karoum, Murad Almadani, Sidgy Alnajar, and Basit Alawode deserve particular gratitude. I would like to thank Dr. Omer Al-ebraheemy for his unwavering support, love, and wisdom in assisting us in understanding and following the true way, our faith Islam, as dictated by the Quran and Sunnah. It is also because of him that I see my time in Saudi Arabia as the most meaningful of my life. It is impossible to identify all of my KFUPM friends and colleagues who made my stay and work at the campus so entertaining and memorable. Thank you all!

Last, but not least, I would like to express my gratitude to all of my professors, from undergraduate to the present, for their tremendous intellectual influence and inspiration, especially Prof. Sharief Babikir.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
LIST OF TABLES	xi
LIST OF FIGURES	xii
ABSTRACT (ENGLISH)	xvi
ABSTRACT (ARABIC)	xix
CHAPTER 1 INTRODUCTION	1
1.1 Motivation	1
1.2 Problem Statement and Organization of the Thesis	3
CHAPTER 2 LITERATURE REVIEW	7
2.1 Introduction	7
2.2 Generative Models	8
2.3 Spatial and Spatio-Temporal Approaches	10
2.4 Compressed Sensing Techniques	12
2.5 Deep Learning Models	14
CHAPTER 3 EEG AS A NEUROIMAGING TECHNIQUE	19
3.1 Introduction	19
3.2 Brain Mapping	20
3.3 Source Modeling	22
3.3.1 Volume Conductor Model	24

3.3.2	Inverse Model	25
CHAPTER 4	COMPRESSED SENSING	26
4.1	Introduction	26
4.2	Reconstruction Algorithms	28
4.2.1	Convex Relaxation	29
4.2.2	Greedy Pursuit Algorithms	29
4.2.3	Bayesian Algorithms	30
CHAPTER 5	DATA GENERATION	32
5.1	Introduction	32
5.2	High Spatial Resolution MRI Image	32
5.3	Digitization	33
5.4	Anatomical Segmentation	34
5.5	Electrodes Placement	35
CHAPTER 6	PROPOSED APPROACHES	38
6.1	Introduction	38
6.2	Single-Stage Approach	39
6.3	Multi-Stage Approach	39
6.4	Methodology	40
6.4.1	Compressed Sensing Reconstruction Algorithms	40
6.4.2	Deep Learning Mapping Structures	43
CHAPTER 7	IMPLEMENTATION, RESULTS, AND DISCUSSION	46
7.1	Introduction	46
7.2	Compressed Sensing	47
7.2.1	Low-Resolution Source Model	47
7.2.2	High-Resolution Source Model	52
7.3	Deep Learning	59
7.3.1	Source Space Division	59
7.3.2	Learning Setups	64

7.3.3	Simplified Residual Networks Model	65
7.3.4	Simplified Sequence Models	71
7.3.5	Complete Whole-Brain Structure	77
CHAPTER 8	CONCLUSION AND FUTURE WORK	83
8.1	Conclusion	83
8.2	Future Work	85
8.2.1	Block Sparsity	86
8.2.2	Joint Sparsity	89
APPENDIX A	FULL BRAIN 56 REGIONS' RESNET MODELS'	
	PERFORMANCE CURVES	92
REFERENCES		112
VITAE		128

LIST OF TABLES

7.1	Low-resolution source model parameters	49
7.2	high-resolution source model parameters	55
7.3	Residual networks models summary	69
7.4	Sequence models summary	76
7.5	Deep Learning models accuracy summary	77
7.6	Full Brain Residual networks models summary	81
7.7	Full Brain Deep Learning models accuracy summary	82

LIST OF FIGURES

1.1	Active brain cortex parts during hand movement	2
1.2	Simulated brain divided into voxels	2
1.3	Propagation of EEG signals generated from sparse sources through the brain into non-invasive electrodes	3
1.4	Deep Learning module as a function mapping the inverse problem . . .	5
2.1	(a) Sort of feed-forward neural network used. (b) The network's fun- damental computational element	15
2.2	Spatio-temporal source decoding model using LSTM	17
3.1	The brain of leborgne, Broca's patient, the first brain with lesion that had related a region to a function	21
3.2	Functional map of the brain	22
3.3	Superposition of source activity	23
3.4	Source modelling	24
5.1	Raw MRI head image	33
5.2	CTF coordinates aligned MRI images	34
5.3	Head volume digitization	35
5.4	Head anatomy segmentation	36
5.5	10-05 system electrodes alignment with the head CTF coordinates system	37
6.1	Center points of coarse grid regions	41
6.2	Fine grid points of one region	42
7.1	Source points of low-resolution source space	48

7.2	Low-resolution one-component source space: One active source reconstruction using compressed sensing algorithms	50
7.3	Low-resolution one-component source space: two active sources reconstruction using compressed sensing algorithms	51
7.4	Low-resolution one-component source space: three active sources reconstruction using compressed sensing algorithms	53
7.5	Low-resolution one-component source space: multi-active sources reconstruction using compressed sensing algorithms. (a) five active sources, (b) seven active sources and (c) ten active sources	54
7.6	High-resolution one-component source space: One active source reconstruction using compressed sensing algorithms	56
7.7	High-resolution one-component source space: two active sources reconstruction using compressed sensing algorithms	57
7.8	High-resolution one-component source space: three active sources reconstruction using compressed sensing algorithms	58
7.9	High-resolution one-component source space: five active sources reconstruction using compressed sensing algorithms	60
7.10	Source points of high-resolution source space	61
7.11	Simulated regions, Region1 and Region2 are adjacent while Region3 is placed on the further side	63
7.12	Full brain volume division into regions colored in red, green and blue .	64
7.13	The implemented ResNet structure	66
7.14	Training performance curves of individual regions ResNets	69
7.15	(a) Training performance curves of coarse grid ResNet, and (b) Confusion matrix	70
7.16	Many-to-one Sequence model structure with LSTM cells	72
7.17	Training performance curves of individual regions LSTMs	74
7.18	(a) Training performance curves of coarse grid LSTM, and (b) Confusion matrix	76
7.19	Training performance curves of full brain coarse grid ResNet	80

8.1	Demonstration of block-sparse signals. (left) known and (right) unknown block partitions	88
8.2	Matrices representation of Multiple Measurement Vectors	89
A.1	ResNet model's Training performance curves of Region1	93
A.2	ResNet model's Training performance curves of Region2	93
A.3	ResNet model's Training performance curves of Region3	93
A.4	ResNet model's Training performance curves of Region4	94
A.5	ResNet model's Training performance curves of Region5	94
A.6	ResNet model's Training performance curves of Region6	94
A.7	ResNet model's Training performance curves of Region7	95
A.8	ResNet model's Training performance curves of Region8	95
A.9	ResNet model's Training performance curves of Region9	95
A.10	ResNet model's Training performance curves of Region10	96
A.11	ResNet model's Training performance curves of Region11	96
A.12	ResNet model's Training performance curves of Region12	96
A.13	ResNet model's Training performance curves of Region13	97
A.14	ResNet model's Training performance curves of Region14	97
A.15	ResNet model's Training performance curves of Region15	97
A.16	ResNet model's Training performance curves of Region16	98
A.17	ResNet model's Training performance curves of Region17	98
A.18	ResNet model's Training performance curves of Region18	98
A.19	ResNet model's Training performance curves of Region19	99
A.20	ResNet model's Training performance curves of Region20	99
A.21	ResNet model's Training performance curves of Region21	99
A.22	ResNet model's Training performance curves of Region22	100
A.23	ResNet model's Training performance curves of Region23	100
A.24	ResNet model's Training performance curves of Region24	100
A.25	ResNet model's Training performance curves of Region25	101
A.26	ResNet model's Training performance curves of Region26	101
A.27	ResNet model's Training performance curves of Region27	101

A.28 ResNet model's Training performance curves of Region28	102
A.29 ResNet model's Training performance curves of Region29	102
A.30 ResNet model's Training performance curves of Region30	102
A.31 ResNet model's Training performance curves of Region31	103
A.32 ResNet model's Training performance curves of Region32	103
A.33 ResNet model's Training performance curves of Region33	103
A.34 ResNet model's Training performance curves of Region34	104
A.35 ResNet model's Training performance curves of Region35	104
A.36 ResNet model's Training performance curves of Region36	104
A.37 ResNet model's Training performance curves of Region37	105
A.38 ResNet model's Training performance curves of Region38	105
A.39 ResNet model's Training performance curves of Region39	105
A.40 ResNet model's Training performance curves of Region40	106
A.41 ResNet model's Training performance curves of Region41	106
A.42 ResNet model's Training performance curves of Region42	106
A.43 ResNet model's Training performance curves of Region43	107
A.44 ResNet model's Training performance curves of Region44	107
A.45 ResNet model's Training performance curves of Region45	107
A.46 ResNet model's Training performance curves of Region46	108
A.47 ResNet model's Training performance curves of Region47	108
A.48 ResNet model's Training performance curves of Region48	108
A.49 ResNet model's Training performance curves of Region49	109
A.50 ResNet model's Training performance curves of Region50	109
A.51 ResNet model's Training performance curves of Region51	109
A.52 ResNet model's Training performance curves of Region52	110
A.53 ResNet model's Training performance curves of Region53	110
A.54 ResNet model's Training performance curves of Region54	110
A.55 ResNet model's Training performance curves of Region55	111
A.56 ResNet model's Training performance curves of Region56	111

THESIS ABSTRACT

NAME: Elmubarak M. Imam

TITLE OF STUDY: Compressed Sensing- and Deep Learning-based Brain
Source Localization of EEG Signals

MAJOR FIELD: Electrical Engineering

DATE OF DEGREE: September 2021

The brain source localization problem aims to identify the location of neurons sources responsible for generating the electrical signals within the brain. Correct localization of functional parts is still an open research problem. Non-invasive techniques utilizing high temporal resolution EEG and MEG signals have proved helpful. Due to the lack of large size real EEG datasets that also label the active regions responsible for these signals, EEG datasets were generated by simulating the EEG forward problem using the fieldtrip toolbox. MRI anatomical imaging was used to generate a realistic head model; then, FEM was used to formulate the leadfield matrix. Two types of simulation data were generated: a) Spatial EEG data and b) Spatial-Temporal EEG data. Two disciplines were applied to solve the inverse problem: 1) Compressed Sensing and 2) Deep Learning. Making advantage of the sparse nature of brain source activity, sparse

reconstruction techniques, like compressed sensing, were investigated. OMP, StOMP, and CoSAMP were applied to the spatial EEG data. For single-source activity, these algorithms were able to reconstruct the active source perfectly wherever it was, in the cortex of the brain or deep inside the brain. For multi-source activity, compress sensing algorithms performance was quite well on reconstructing cortex sources. However, their performance was not good due to the correlation between electromagnetic channels as sources go deep inside the brain. In general, inverse problems are considered complex functions, difficult to be formulated in a closed-form solution. Since neural networks are universal mapping techniques, deep learning structures were investigated. Within this thesis, deep learning architectures were only applied to single-active source problems. Residual networks were utilized to predict the spatial data source as they are good at capturing signatures. And Recurrent Neural Networks were applied to the Spatio-temporal data due to their ability to correlate temporal dynamic features. Furthermore, a multi-stage approach is proposed for deep learning techniques to increase prediction accuracy and reduce the number of labels per model, thus reducing the amount of Softmax computation. Two grids were formulated: coarse grid and fine grid. Three regions were implemented to verify the approach. ResNets of 5-layers were used for the spatial data, and LSTM models were utilized for Spatio-temporal data. Deep learning models gave high prediction accuracy. For ResNets, the overall spatial model test accuracy is 96%. For LSTM models, the overall Spatio-temporal model test accuracy is 95%. Deep learning techniques are a promising solution. Therefore, they were applied to the real problem, whole-brain volume. However, it was only implemented for the ResNets module due to the large dataset required to train the whole-brain volume

LSTM module. The overall whole-brain spatial localization accuracy of EEG signals is 92.7%.

Keywords— Brain Source Localization, EEG, Inverse problem, Sparse reconstruction, Compressed Sensing, Deep Learning, ResNet, LSTM.

ملخص الرسالة

الاسم: المبارك محمد المبارك امام

عنوان الدراسة: تحديد مصدر اشارات الدماغ القائم على الاستشعار المضغوط والتعلم العميق لإشارات EEG

التخصص: هندسة كهربائية

تاريخ الدرجة العلمية: سبتمبر 2021

الهدف من مشروع تحديد مصدر إشارات الدماغ هو تحديد موقع مصادر الخلايا العصبية المسؤولة عن توليد الإشارات الكهربائية داخل الدماغ. لا يزال التحديد الصحيح للأجزاء الوظيفية يمثل مسألة بحث مفتوحة. أثبتت التقنيات غير الجراحية التي تستخدم إشارات EEG و MEG ذات الدقة الزمنية العالية أنها مفيدة. نظرًا لعدم وجود مجموعات بيانات EEG حقيقية كبيرة الحجم والتي تؤسّم أيضًا المناطق النشطة المسؤولة عن هذه الإشارات، تم إنشاء مجموعات بيانات EEG عن طريق محاكاة مسألة إشارات EEG الأمامية باستخدام أداة Fieldtrip. تم استخدام التصوير التشريحي بالرنين المغناطيسي لتوليد نموذج رأس واقعي؛ ثم تم استخدام الـ FEM لصياغة مصفوفة المجال الأمامي. تم إنشاء نوعين من بيانات المحاكاة: (أ) بيانات EEG المكانية (ب) بيانات EEG المكانية-الزمانية. تم تطبيق تخصصين لحل المسألة العكسية: (1) الاستشعار المضغوط (Sensing Compressed) و (2) التعلم العميق (Deep Learning). بالاستفادة من الطبيعة المتفرقة لنشاط الدماغ، تم فحص تقنيات إعادة البناء المتفرقة، مثل الاستشعار المضغوط. تم تطبيق خوارزميات OMP و StOMP و CoSAMP على بيانات EEG المكانية. بالنسبة للنشاط أحادي المصدر، تمكنت هذه الخوارزميات من إعادة بناء المصدر النشط تمامًا أينما كان، في قشرة الدماغ أو في أعماق الدماغ. بالنسبة للنشاط متعدد المصادر، كان أداء خوارزميات الاستشعار المضغوط جيدًا في إعادة بناء مصادر القشرة. ومع ذلك، لم يكن أداءهم جيدًا بسبب الترابط بين القنوات الكهرومغناطيسية للمصادر العميقة داخل الدماغ. بشكل عام، تعتبر المسائل العكسية دوال معقدة، يصعب صياغتها في حل نهائي. نظرًا لأن الشبكات العصبية تعتبر تقنيات عامة لربط الدوال، تم دراسة بنيات التعلم العميق. ضمن هذه الأطروحة، تم تطبيق بنيات التعلم العميق فقط على المصدر الأحادي النشط. تم استخدام شبكات الباقي للتنبؤ بمصدر البيانات المكانية لأنها جيدة في التعرف على الأنماط. وتم تطبيق الشبكات العصبية المتكررة على البيانات المكانية-الزمانية؛ نظرًا لقدرتها على ربط السمات الديناميكية الزمنية. علاوة على ذلك، تم اقتراح طريقة متعددة المراحل لتقنيات التعلم العميق لزيادة دقة التنبؤ وتقليل عدد العلامات لكل نموذج، من ثم تقليل مقدار حساب الـ Softmax. تمت إعداد شبكتين: شبكة غير دقيقة وشبكة دقيقة. تم تجهيز ثلاث مناطق للتحقق من طريقتنا. تم

استخدام شبكات ResNets مكونة من 5 طبقات في حالة البيانات المكانية، واستخدمت نماذج LSTM في حالة البيانات المكانية-الزمانية. أعطت نماذج التعلم العميق دقة تخمين عالية. بالنسبة لشبكات ResNets ، بلغت الدقة الإجمالية للنموذج المكاني 96٪. وبالنسبة لنماذج LSTM ، بلغت الدقة الإجمالية للنموذج المكاني-الزمني 95٪. تعتبر تقنيات التعلم العميق حلاً واعدًا. لذلك، تم تطبيقها على المسألة الحقيقية، كامل الدماغ. غير أنه تم تنفيذه فقط لوحدة ResNets نظرًا للحاجة لمجموعة كبيرة من البيانات لتدريب وحدة LSTM لكامل الدماغ. بلغت الدقة الإجمالية للتحديد المكاني لإشارات EEG لكامل الدماغ 92,7٪.

CHAPTER 1

INTRODUCTION

1.1 Motivation

With all the advancements in technology, the human brain is still a vastly unexplored arena. The researchers, both belonging to the medical field and technological fields, are trying to understand the function of the human brain properly. It is well-known that the brain is functionally localized, i.e., different parts of the brain are related to or are responsible for specific activities. However, the correct localization of functional parts is still an open research problem. In this regard, non-invasive techniques utilizing EEG and MEG signals have proved helpful. These techniques allow us to observe brain activity in the form of electromagnetic signals emitted from the various regions of the brain. After that, machine learning and data analysis tools could be applied to this data to map the functional localization of the brain and then understand the meaning of that particular brain activity. Correct localization of brain activity is also essential to help understand the common neurodegenerative disorders. It is also expected to make the human-computer interface more efficient and useful.

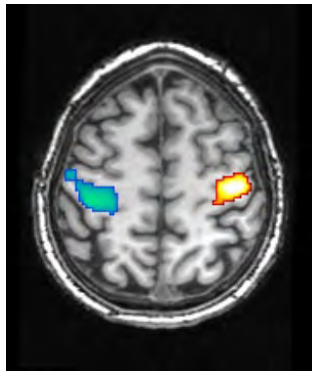


Figure 1.1: Active brain cortex parts during hand movement [1]

We know that several parts of the brain are responsible for different tasks. However, it is interesting to note that only a few parts of the brain are active at a given moment. This can also be interpreted in that each action activates only a small number of the brain parts (region). This could be explained with the help of Fig. 1.1, where only two areas of the brain are active.

For computational purposes, we can divide the whole-brain into several three-dimensional sectors with the help of a grid, as shown in Fig. 1.2. This means only a few sectors of the entire grid will be active at a given time. This brings us to the concept of sparsity in human brain activity.

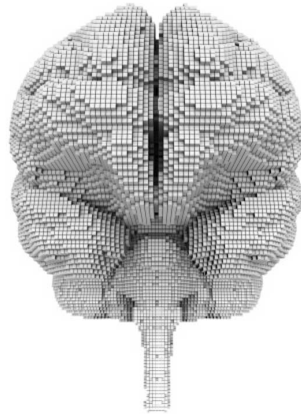


Figure 1.2: Simulated brain divided into voxels [2]

1.2 Problem Statement and Organization of the Thesis

In this thesis, we will benefit from:

- The sparse nature of brain activity to formulate a compressed sensing problem to localize the brain's active regions, as modeled in Fig. 1.3.
- The universal mapping feature of neural networks to develop a Deep Learning module to map the brain inverse problem function.



Figure 1.3: Propagation of EEG signals generated from sparse sources through the brain into non-invasive electrodes [3]

We can model the problem of brain source localization as a linear system as follows:

$$y = Ax + n \tag{1.1}$$

where $y \in \mathbb{R}^M$ is the recorded or received EEG/MEG signals, with $M \ll N$. Further,

$A \in \mathbb{R}^{M \times N}$ is a matrix that controls the signal recording mechanism. The nature of A depends on a number of factors including the brain characteristics and the nature of signal generated by a particular part of a brain. These characteristics are usually unknown and therefore the matrix A cannot be determined perfectly. This uncertainty in the values of A needs to be taken into account in any compressed sensing based solution to this problem. Moreover, $x \in \mathbb{R}^N$ is an unknown sparse vector that represents the several locations within a brain. The 3-dimensional location grid is flattened into a 1-dimensional vector that is sparse as only a few brain regions are active at a specific time. Compressed sensing can be used to find these active locations with the help of recorded non-invasive electrography signals (EEG/MEG). Finally, $n \in \mathbb{R}^M$ is the noise associated with the recorded signals, which is assumed to be Gaussian random vector with zero mean and covariance matrix C , $n \sim \mathcal{N}(0, C)$.

This is a compressed sensing problem, which could be solved by a number of algorithms. However, the special nature of the problem imposed by the nature of brain makes it impossible to use the already available algorithms in an efficient manner. Therefore, there is a need to come up with an efficient strategy to estimate the active locations of the brain. The aim of the compressed sensing part of the research is to design an efficient approach for the detection of active brain regions, and to compare this approach with other existing approaches in the literature.

On the other hand, Artificial Intelligence and data analysis tools could be applied to this data to map the functional localization of the brain and then understand the meaning of that particular brain activity. Taking advantage of the universal mapping feature of neural networks, the brain source localization problem could be modeled

as:

$$f : X \rightarrow Y \quad (1.2)$$

where f is the Deep Learning module as a function to map from $X \in \mathbb{R}^{M \times p}$, the scalp EEG/MEG signals set, to $Y \in \mathbb{R}^{N \times p}$, the corresponding active neurons within the total brain neurons set. p is the number of samples within each dataset. In other terms, Deep Learning structures shall be utilized to formulate the inverse model from the observed potential to the physiological source, as shown in Fig. 1.4.

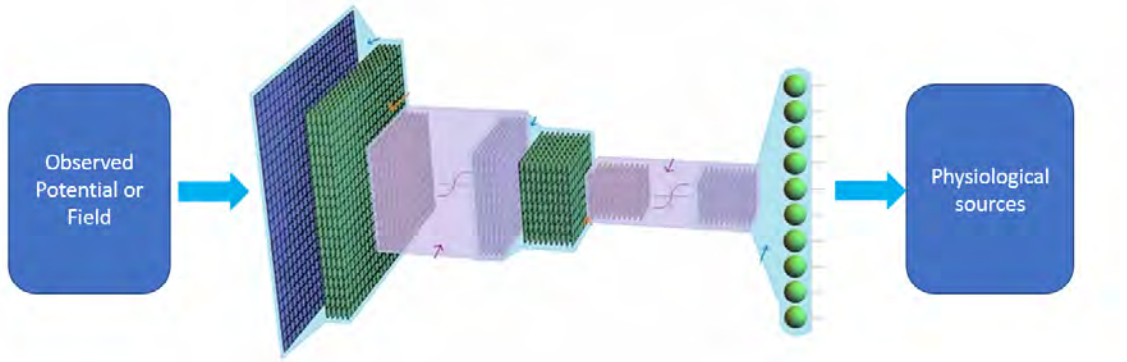


Figure 1.4: Deep Learning module as a function mapping the inverse problem [4]

We organized the thesis by introducing the research efforts exerted on the problem while illustrating in detail the most significant works in Chapter 2. We have separated these literature efforts into five divisions according to the discipline followed in the reconstruction model. Then, we presented a medical and physiological introduction of the brain and brain modeling, respectively, in Chapter 3. We illustrated how the brain is modeled by the forward problem as closely as possible to a real

one. Furthermore, forward problem neuroscience and electromagnetic computational steps, including MRI slicing, digitization and quantization, anatomical segmentation, and electrode placement, to generate the simulated electrophysiological signals and dataset has been mentioned in Chapter 5. Chapter 4 contains a brief demonstration of Compressed Sensing and its reconstruction techniques, primarily those related to our problem under study. Chapter 6 presented our proposed approaches: single-stage and multi-stage approaches and the methodology of applying them with compressed sensing and deep learning. To make reading and following-up easy, in Chapter 7, we have included implementation, results, and discussion, altogether. Training performance curves have been excluded from Chapter 7 and listed in the appendix since they will disturb reading fluency by their vast number of pages. Finally, in the last chapter, Chapter 8, we concluded our work and observations. Then we gave our recommendations to enhance our proposed solution further.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Localizing brain signals' sources is an active area where researchers developed various techniques and algorithms to map the active brain locations responsible for the sensed EEG signals. The contemporary trends and advances of these techniques are surveyed and illustrated in [3], [5], [6].

This chapter summarizes the methods and techniques used in literature to investigate the brain source localization problem. We have divided them according to their principle of operation like Generative models, according to their data dimension in space and time, and finally, according to the modern techniques that we are concerned about here in this thesis -Compressed Sensing and Deep Learning.

2.2 Generative Models

Reconstruction of the neuronal source from the sensor level requires the inversion of a generative model. This inverse problem is fundamentally ill-posed. In the literature, two generative approaches were proposed for MEG inverse problems: equivalent current dipole models and distributed source models, which both could also be used for EEG inverse problems. The first current dipole models can be estimated via Bayesian particle filtering methods, as illustrated in [7] and proposed in [8] for brain source localization. Regarding the distributed source models, Commonly used techniques include the minimum-norm estimate (MNE) [9] (illustrated further below), the minimum-current estimate [10], the low resolution electromagnetic tomography (LORETA) [11], and their modified variants; namely, weighted-MNE [12], standardized-LORETA [13], variable-resolution-electrical-tomography (VARETA) [14]. Despite the generative models, different spatial scanning [15] and beamforming techniques (e.g. MUSIC [16]) were also developed.

Considered the establishing method for brain source localization, the Minimum Norm Estimate (MNE) is the most generalized technique to predict 3-dimensional source activity. It was introduced in 1984 by Hamalainen and Ilmoniemi. The algorithm solution does not require any *priori* information. It is based on the assumption that the solution is the current activity inside the brain with the slightest current intensity (i.e., minimum ℓ_2 -norm). The method defines the leadfield as the linear combination of the magnetometer leadfields, from which brain activity current distribution could be estimated. Therefore, the i th magnetometer measurement B_i corresponding

to the primarily current source distribution is given by:

$$B_i = \int L_i(r) \cdot J^P(r) dV \quad (2.1)$$

where L_i is the leadfield to the i th magnetometer sensor. Considering the overall system of all magnetometer sensors and brain current sources, the linear form of (2.1) is obtained by substituting the matrix form of the magnetometer leadfields L_i as follows:

$$B = LJ \quad (2.2)$$

The solution of the linear system is obtained by finding the pseudo-inverse of L , so that:

$$\hat{J} = L^+ B \quad (2.3)$$

this pseudo-inverse is the Moore–Penrose generalized inverse in which:

$$L^+ = L^T(LL^T)^+$$

Minimum Norm Estimate puts the first steps to solve the brain source localization as it provides good estimations depending on current source location and source space resolution. However, it faces a high localization error regarding deep brain current sources. That is due to the harmonic nature of the ℓ_2 -norm function, in which $\nabla^2 J = 0$. Thus, the harmonics have peak values at the domain boundaries.

2.3 Spatial and Spatio-Temporal Approaches

Various approaches to improve localization accuracy through spatial priors have been proposed, including spatially informed basis functions [17], [18], smoothness promoting spherical harmonic, spline basis functions [19], as well as various sparsity promoting priors [20], [21], [22]. Other approaches to improve localization accuracy include combined models for simultaneous noise suppression (Zumer et al., 2007), models for potential inaccuracies in the forward model [23], and models for potential inaccuracies in the forward model [24].

The foundational forward problem is commonly formulated in the absence of temporal information from the EEG data. The function between brain source activity and observed scalp EEG is commonly regarded as a linear model as in (1.1), rewritten here again:

$$y = Ax + n$$

in this approach, the spatial component is appropriately handled. But, the effect of time is ignored, and the propagation from neurons to electrodes is linear with a fixed time.

Temporal data is critical for reconstructed source activity. To begin with, the contiguous time samples contribute to the number of measurements available for the ill-posed problem, which narrows the solution space. Besides, EEG signals carry some sequential patterns that characterize the source and assist in restoration. The magnitude of the source signal diminishes with deeper sources, and the fluctuation of the EEG signal is associated with the position of sources. Furthermore, simulating the

temporal dynamics of the EEG signal reduces noise and smooths out source estimations across time.

Time involvement has recently been studied, and neural activity has been expressed via spatial basis and temporal dynamics. The mathematical expressions are formulated as follows:

$$y_t = Ax_t + n_t, \quad (2.4)$$

$$x_t = \Phi \cdot q_t, \quad (2.5)$$

$$q_t = f(q_{t-1}, \dots, q_{t-\tau}, w_t) + \eta_t, \quad (2.6)$$

$$w_t = g(w_{t-1}) + \epsilon_t, \quad (2.7)$$

thus, to represent source activity, $\Phi \in \mathbb{R}^{N \times d}$ contains the spatial basis, and $q_t \in \mathbb{R}^d$ contains the coefficients of each basis. Source activity q_t is modeled by $f(\cdot)$ to simulate the temporal impact of adjacent state change across time. w_t indicates f arguments approximated by the random walk process $g(\cdot)$.

Distributed source reconstruction techniques lose the temporal information that could be extracted from the temporal relations between neuronal currents. For this reason, many Spatio-temporal techniques have been proposed, including models based on different Spatio-temporal basis functions and carefully designed or empirically learned Gaussian Spatio-temporal priors [25], [26], [27]. Many approaches are specifically targeted to obtain spatially sparse but temporally informed, smooth, or oscillatory source estimates [28], [29].

High-quality Bayesian inverse solutions employ various Spatio-temporal regular-

ization forms by employing Gaussian processes (GPs) that use space-time separable covariance functions. These models also referred to as Gaussian random fields, have been successfully applied in various inverse problems [30], [31].

2.4 Compressed Sensing Techniques

Compressed sensing techniques are considered a promising solution for the inverse problem of brain source localization, and hence, recently applied to locate EEG/MEG sources as in [32], [33], [34], [35], [36]. These techniques include minimum current estimate [10], mixed-norm estimates (MxNE) [37], and the simplest and basic method is the FOCal Underdetermined System Solution (FOCUSS) [20].

FOCUSS is a tomographic reconstruction algorithm proposed in [38] to develop a sparse solution to the ill-posed EEG/MEG inverse problem. A high-resolution method utilizes the forward model to estimate a primary reconstruction region. It is a recursive method with mathematical calculations based on Weighted Minimum Norm (WMN) method. The location resolution is refined by recursively updating the weights of the solution from the previous step solution. For the magnetic field case, the current I is given by:

$$I = W(GW)^+ B = WW^T G^T (GWW^T G^T)^{-1} B \quad (2.8)$$

where $W \in \mathbb{R}^{N \times N}$ is a weight matrix that regularizes the elements of I by delivering constraints to the solution. FOCUSS' recursive stages update the weight matrix W . W is built from the elements of I in the previous iteration steps to improve some of the already significant elements. These W matrices are dimensionless scaling factors

that can be adjusted to generate an infinite number of recursive solutions. W can also be built by considering its diagonal elements to represent the previous iterative step solution, resulting in:

$$W_k = \begin{bmatrix} I_{1_{k-1}} & & 0 \\ & \ddots & \\ 0 & & I_{n_{k-1}} \end{bmatrix} \quad (2.9)$$

k represents the index of the iteration step, and $I_{i_{k-1}}$ is the i th element of the vector I at the $(k - 1)$ iteration.

In (2.8), $I \in \mathbb{R}^N$ is the unknown current dipole amplitudes vector to be computed by the inverse solution. $G \in \mathbb{R}^{M \times N}$ is the basis matrix whose elements are known from the model. These elements are called spatial weighting elements. They express the impact of one dipole component of unit magnitude on a specific magnetic measurement. This is equivalent to the leadfield matrix, which utilizes the head model to estimate the forward model for a particular conductivity. G entries are mere geometric factors and can be precalculated for a certain head and sensor layout. It is possible to compute for any forward models containing realistic geometry and significant compartments inside the head volume conductance. Additionally, $B \in \mathbb{R}^M$ is the vector representing M radial magnetic field measures at m sensors. The forward

model that describes the relation between I and B is given by:

$$\begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix} = \begin{bmatrix} G_1 & G_2 & G_3 & \cdots & G_N \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} \quad (2.10)$$

FOCUSS is a direct approach that assumes there is no prior knowledge about the source location. The solution is unstable, though. To address this constraint, regularization is used. FOCUSS is recognized as a basic method since it employs WMN to estimate the source. Its implementation is easy; however, it is computationally expensive due to its recursive nature. Compared to other algorithms, FOCUSS has a better localization performance and can address non-uniquely identified localized sources. Furthermore, it has a higher spatial resolution and is more consistent when exposed to changes.

2.5 Deep Learning Models

Neural Networks were early investigated to solve the brain source localization problem. To the extent of our knowledge, the oldest article on brain source localization using neural networks was published in 1991 [39]. The authors -Abeyratne et al., used a shallow neural network of two hidden layers. However, some papers refer to even an older article published in 1990 also by Abeyratne et al. [40]. We did our best to find an online version of this novel article, however, without success.

The generic structure of this oldest feed-forward neural network developed to localize human brain sources is shown in Fig. 2.1 (a). Circles and directed arrows symbolize neurons and connection, respectively. The input layer has the same number of neurons as electrodes on the scalp. Inside the spherical head model, the six dipole parameters are represented by six neurons in the output layer. This first model uses only two hidden layers, as demonstrated by hd1 and hd2 circles, making it a shallow model. Fig. 2.1 (b) shows the model's fundamental computational element that uses the tanh function as the non-linear activation.

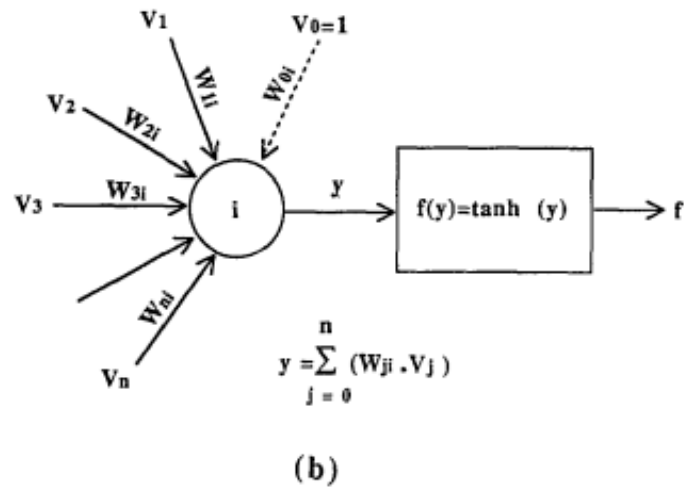
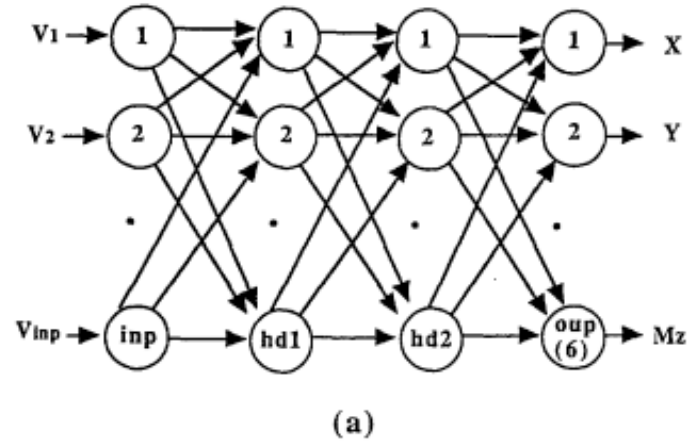


Figure 2.1: (a) Sort of feed-forward neural network used. (b) The network's fundamental computational element [39]

Later on, some early researchers followed Abeyratne et al. and used basic neural networks to reconstruct EEG and MEG sources [41], [42], [43], [44]. Back then, they were called Back-Propagation Neural Networks (BPNN) or Artificial Neural Networks (ANN). These basic neural networks are also considered shallow neural networks. The performance of those early artificial brain source localizers was not encouraging due to the poor computational power available back then and the lack of efficient optimization algorithms and deep learning strategies.

Regarding inverse problems in general, Neural Networks are considered a promising technique to map these complex problems that are extremely difficult to describe in a closed-form solution. Many articles were published emphasizing the superior of Deep Learning architectures in mapping inverse problems [45], [46], [47], [48], [49]. For example, Deep Learning models were utilized to address the inverse problem in Geoscience and Remote Sensing [50], computational heat transfer [51], and in medical imaging, like MRI [52].

Researchers have recently taken advantage of the current availability of high computational power and applied deep learning architectures to address the brain source reconstruction problem. Different methods have been proposed, like [53], which utilizes a Multi-Layer Perceptron (MLP) network. Also, computer vision’s most popular structure, Convolutional Neural Network (CNN), was applied to localize stationary EEG signals [53], [54], [55]. Nevertheless, U-Net and Generative Adversarial Networks (GANs) applications were investigated [56]. And few interesting works used Recurrent Neural Network (RNN) architecture to solve the EEG inverse problem using the temporal information of EEG signals [57], [58].

It is interesting to highlight the work of Cui et al. [58]. They took advantage of the dynamic temporal nature of brain signals to propose a Recurrent Neural Network (RNN) model since they are well suited for the mapping of dynamic sequence data. As shown in Fig. 2.2, their proposed module uses Long-Short Term Memory (LSTM) to encode the sEEG recordings and then localize the source by decoding the implicit location.

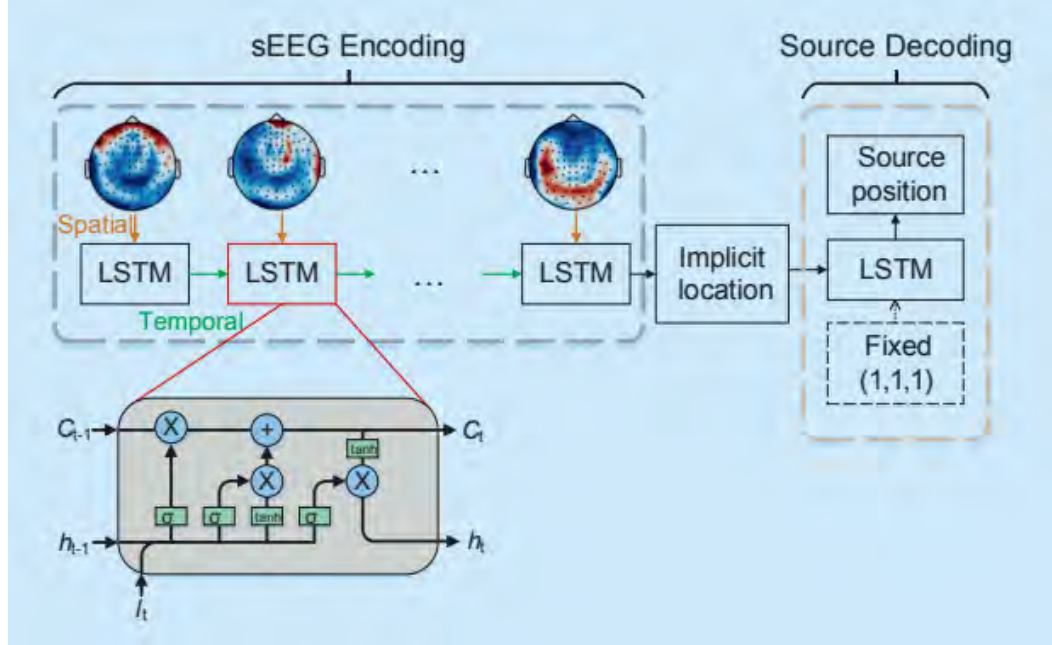


Figure 2.2: Spatio-temporal source decoding model using LSTM [58]

Following the huge research efforts to model the dynamic nature of brain signals, Cui et al. generated the stationary brain activity of (2.5) based on the model in [59], which is defined as:

$$x_k = A_1 x_{k-1} + A_2 x_{k-2} + A_3 x_{k-\tau} + A_4 x_{k-1}^2 + A_5 x_{k-1}^3 + \eta_k, \quad (2.11)$$

where x_k is the source signal at time k , and the values of (2.11) parameters are fixed

as follows: $\tau = 20$, $A_1 = 1.0628$, $A_2 = -0.42857$, $A_3 = 0.008$, $A_4 = 0.000143$,
 $A_5 = -0.000286$ and $||\eta_k|| \leq 0.05$.

CHAPTER 3

EEG AS A NEUROIMAGING TECHNIQUE

3.1 Introduction

This chapter discusses neuroimaging techniques for gaining a better understanding of the brain's physiological and cognitive functions.

Here we go on to briefly cover popular neuroimaging techniques such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG). However, as given in this thesis, we are to investigate EEG-based source localization; we will go thoroughly on EEG as a neuroimaging technique and its application in brain source localization in depth.

EEG and MEG are non-invasive electrophysiological measures commonly used to record brain activities. Both EEG and MEG have a high temporal resolution in milliseconds (~ 1 ms), making them much faster than any other physiological techniques, even functional MRI (fMRI). Therefore, EEG/MEG can capture Event-Related Po-

tential or Field (ERP/ERF), oscillatory activities, and cortical neuronal dynamics. However, since they are non-invasive techniques that acquire signals at the surface of the head, their components are overlapped and have poor spatial resolution compared to fMRI. Moreover, most EEG/MEG signals power belongs to the activities in the brain's outer cortex. Thus, they are poor at capturing activities deep inside the brain. That's why signals processing techniques are needed to distinguish and highlight deep brain signals components within EEG/MEG signals.

Therefore, the chapter begins with an overview of overall functional brain imaging and its applications. Recently, the neuroimaging field has achieved numerous advancements, especially in its technical techniques. In recent years, these techniques have significantly improved and are now used in various healthcare applications, including functional brain imaging. Applied mathematics, optimization, signaling, imaging, and photonics are part of this multidisciplinary field. Thus, non-invasive imaging techniques, including electrophysiological, hemodynamic, metabolic, and neurochemical processes that occur in the normal and disordered human brain, can be used to understand the human brain better. Each of these imaging modalities has unique temporal and spatial resolution characteristics. These sophisticated imaging modalities are used to understand the neuronal processes in a typical human brain.

3.2 Brain Mapping

In 1861, a patient arrived at Paris hospital that could only say the syllable “Tan,” although he could understand other people's words. After the patient died, doctor

Paul Broca, a French anatomist and physician, dissected his brain and found a Lesion (damage) on the left frontal lobe, later called “Broca’s area” [60]. Back then, Broca concluded that this part was responsible for speech generation. More experiments with animals convinced us that different parts of the brain are dedicated to different activities. Then, Broca’s lesion method was used to draw up a localized brain map and eventually proved after the invention of fMRI in the 1990s, which tracked moment-to-moment changes in brain activity.



Figure 3.1: The brain of leborgne, Broca’s patient, the first brain with lesion that had related a region to a function (here overt speech) [61]

Since then, researchers have conducted many experiments to prove that different parts of the brain are related to different jobs, which led to what was known as the localized brain map. However, in the 1980s, when fMRI was developed, the idea of a localized map of the brain vanished due to the realization that the brain can reroute the specified task of a damaged part into another healthy region. Still, the combination of the localized map of the brain and fMRI is utilized to generate a new map that shows which parts of the brain work together to perform tasks. This new

map is now called the functional map of the brain.

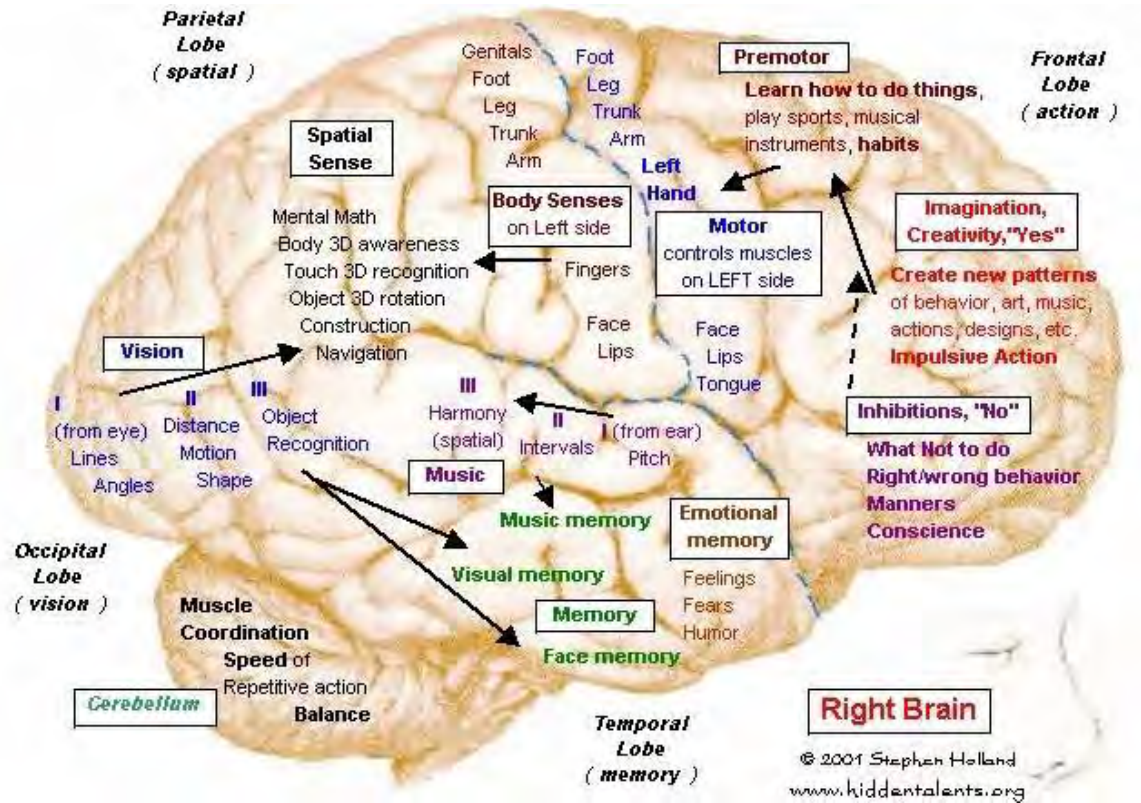


Figure 3.2: Functional map of the brain [62]

3.3 Source Modeling

Source localization is closed to recovering the spatial position of sources and figuring out the timecourse of that source. Time and frequency are the components of EEG/MEG signals that could be characterized naturally. However, the analysis of their spatial characteristics requires an interpretation of the head topography through which EEG/MEG propagates. Forward and inverse modeling of EEG/MEG signals is the way that helps in analyzing the spatial characteristics of brain activities by utilizing the interpretation of the head topography [63].

The low-frequency nature of EEG/MEG signals is referred to the sparse nature of neurons activation and unaligned orientation of neurons cells which results in cancellation or suppression of the overall generated signal. Therefore, brain activities are interpreted to be generated from low-frequency sources. Considering EEG signals, as shown in Fig. 3.3 propagate through the relatively conductive tissues of the brain, through the skull, until electrodes are sensing it at the head skin.

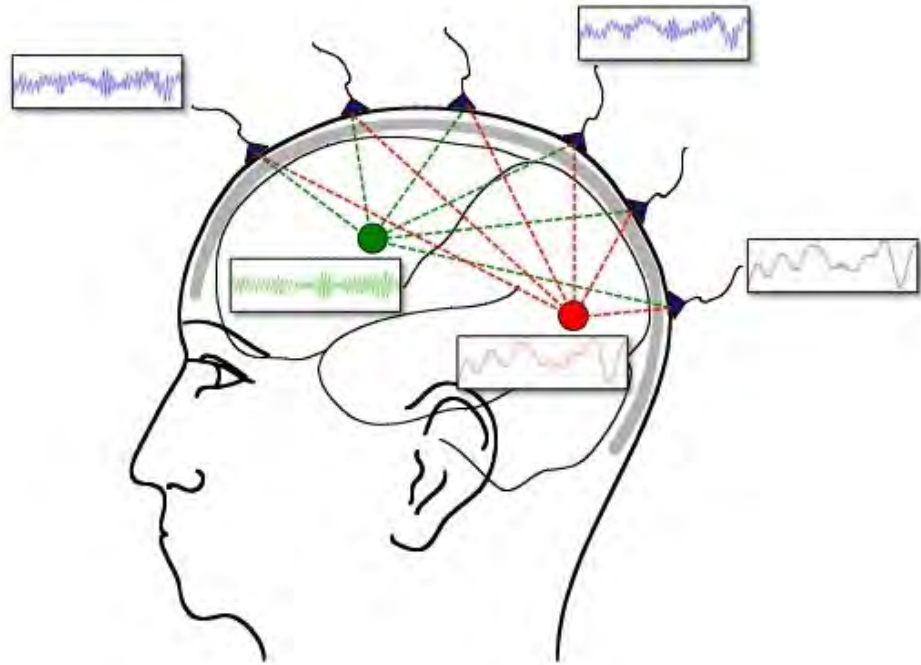


Figure 3.3: Superposition of source activity [64]

Each Source signal is assumed to be sensible by each sensor channel with varying visibility depending on the loss path and distance between the source position and the head electrode. The signal on each Electrode channel is a superposition of all sources that simultaneously activate within the brain.

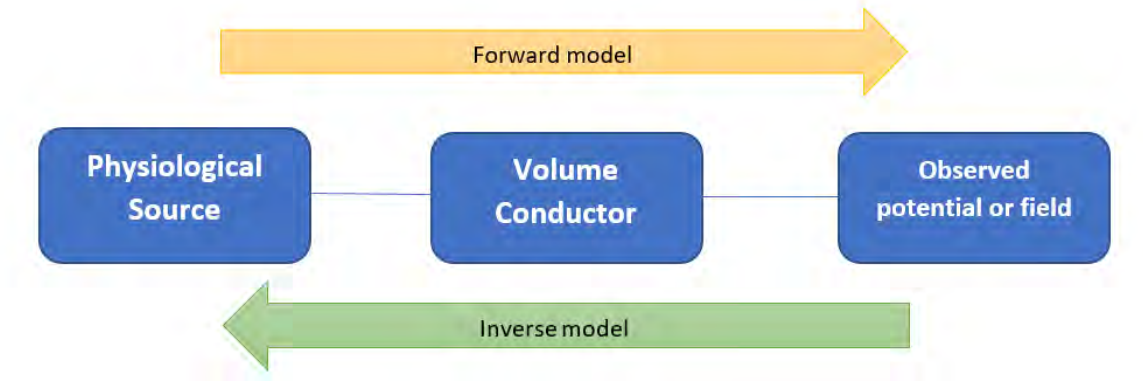


Figure 3.4: Source modelling

3.3.1 Volume Conductor Model

We are dealing with electromagnetic signals that get affected by the head topography. Therefore, to interpret our problem correctly, we need a solid description of the electrical properties of the head tissues. The volume conductor model is used to describe the geometrical model of the head through which the current flow and how those current flows, not from where the current originates.

- Spherical: the simplest but not accurate model.
- Realistic: relies on numerical techniques including:
 - Finite Difference Method: discretizes head geometry into squares.
 - Boundary Element Method (BEM): gives more attention to the surfaces of the geometry. It discretizes the tissues into a triangulated mesh.
 - Finite Element Method (FEM): discretizes volume into tetrahedra or hexahedra. It is considered the most accurate; however, it is computationally more expensive.

3.3.2 Inverse Model

In contrast to Forward modeling, which explains how the potential on the skull will look like given a model for the sources, the inverse model does the opposite, given the observation of how these sources should be interpreted. There are several inverse methods to locate the activity, including:

- Single and multiple dipole fitting: is the simplest method, and the idea is to minimize the error between the model and the measured potential/field.
- Distributed source models: are considered the perfect fit of the model to the measured potential/field. However, that requires additional constraints on the source, such as smoothness, power, or amplitude.
- Beamforming methods: is not concerned about fully explaining the measured topography. It scans the whole-brain with a single dipole and estimates the filter output independent of all other locations at every location.

CHAPTER 4

COMPRESSED SENSING

4.1 Introduction

Compressed sensing was presented early as a novel method to sample signals at rates much lower than the well-known Nyquist rate [65]. Compressed sensing was applied to image processing and computer vision not just to compress the size of data in storage but, more beneficially, to reduce the number of raw data at the early sampling stages. Therefore, the load of signals analog-to-digital conversion, processing, and compression could be eliminated with a trade-off of a bit of complexity in the acquisition stage. This Compressed sampling technique could also be used to acquire highspeed data, where the Nyquist sampling rate is unachievable or of high cost.

This field utilizes the sparse nature of natural signals and the incoherence representation of the same signal in two different transformable domains. The signal $x \in \mathbb{R}^N$ is K -sparse, if K elements are non-zero and $N - K$ elements are zero. Thus, x is a linear combination of K basis vectors. In general, any signal with few significant basis coefficients and many small coefficients is a compressible signal.

Compressed sensing is based on the non-adaptive linear projections of signals, then utilization of optimization techniques to reconstruct the original signal from these projections. To acquire a compressed representation of any signal, a linear measurement process could be formulated that measures $M \times 1$ vector y from $N \times 1$ vector x by computing the inner product between vector x and an $M \times N$ measurement matrix Φ . This Φ is an $M \times N$ basis matrix that consists of a collection of $M \times 1$ orthonormal basis vectors $\{\phi_j\}_{j=1}^M$, i.e. y elements are measured as in $y_j = \langle x, \phi_j \rangle$.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \dots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \phi_{23} & \dots & \phi_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \phi_{M3} & \dots & \phi_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix} \quad (4.1)$$

The efficient properties of compressed sensing attracted researchers to employ it in various applications. Now, it is not just used for compression of stored data but, furthermore, to reduce the traffic in vastly distributed networks with an enormous number of sensors. Compressed sensing is also utilized in the development of the Internet of Things (IoT) technology [66].

Researchers believe brain activity sources, i.e., regions of simultaneously active neuronal sources, are fired individually. Brain activity sources could be interpreted as elements distributed in 3D positions. And with the support of the high temporal

resolution of EEG/MEG measures, the captured EEG/MEG signal could be assumed to be the result of activation of a sparse number of those source elements.

This chapter summarizes compressed sensing reconstruction algorithms into related divisions: Convex Relaxation, Greedy Pursuit, and Bayesian algorithms. An illustration of each branch's main principle and mathematical model will be presented. The most common and popular state-of-art techniques will be briefly discussed accordingly.

4.2 Reconstruction Algorithms

Given that $M < N$, there are an infinite number of solutions to the linear system (1.1). Because the signal of interest is sparse in nature (has K non-zero elements), minimizing the ℓ_0 -norm of the sparse signal would be the most natural way to proceed. As a result, finding a solution may be accomplished using ℓ_0 -minimization.

$$\hat{x} = \arg \min \|x\|_0 \text{ subject to } \|y - Ax\|_2 \leq \epsilon, \quad (4.2)$$

which primarily runs down the number of non-zeros. The sparsest solution is achieved by minimizing the total number of non-zeros in x . However, as it involves an exhaustive search over all potential locations of the non-zero elements of x , this method can not be applied in practice because it is an NP-complete issue.

4.2.1 Convex Relaxation

Remarkably, it was shown in [67], [68] that the minimization problem in (4.2) can be loosened up. Therefore x in (1.1) might be rebuilt with high probability by addressing the following ℓ_1 -optimization problem.

$$\hat{x} = \arg \min \|x\|_1 \text{ subject to } \|y - Ax\|_2 \leq \epsilon. \quad (4.3)$$

This ℓ_1 -optimization is a convex optimization problem that conveniently lessens to a linear program called basis pursuit denoising (BPDN), which can be resolved using interior point techniques. These methods have polynomial computational complexity. Specifically, the computational intricacy is $O(N^3)$ which is not sensible, mainly if the signal size is enormous. The large computational complexity requirements sat by the convex leisure options do not make these approaches appropriate even for issues that have modest demands. As a result, other more efficient techniques have been suggested. This brings us to the following class of algorithms called greedy search formulas.

4.2.2 Greedy Pursuit Algorithms

The greedy pursuit techniques iteratively fine-tune the current estimate of the meager signal x by picking its coefficients which provide the most significant improvement. One of the most important algorithms of this group is the Orthogonal Matching Pursuit (OMP) [69]. Although the technique is quicker than convex relaxation, its options are less exact. Moreover, the algorithm needs to calculate inner products, which be-

comes costly as the length of x and the total amount of non-zeros elements increase. Various other greedy techniques that attempt to resolve the problems connected with OMP have been offered. Like: Subspace Pursuit (SP) [70], Regularized Orthogonal Matching Pursuit (ROMP) [71], Compressive SAMpling Matching Pursuit (CoSAMP) [72], the algorithm proposed by Haupt et al. [73] and Stagewise Orthogonal Matching Pursuit (StOMP) [74] among others.

All these formulas fall under the classification of greedy algorithms that are somewhat faster than basis search. Nonetheless, it should be kept in mind that a common feature among the greedy algorithms and convex relaxation ones is that the solely a priori knowledge they use is the sparsity information.

4.2.3 Bayesian Algorithms

Another class of methods based upon the Bayesian approach recognizes sparse signals' full a priori statistical information. The Fast Bayesian Matching Pursuit (FBMP) [75] adopts this method and presumes a Bernoulli-Gaussian prior to the unidentified vector x . This technique does sparse signal estimation using model selection and model averaging. The sparse signal is expressed as a composite of several elements, the choice of which is based upon successive model growth. FBMP acquires an approximate MMSE estimation of the unidentified vector with large precision and low complexity. It was reported to outmatch several sparse reconstruction algorithms, like a variational-Bayes implementation of BCS [76], Sparse Bayes [77], StOMP [74], GPSR-Basic [78], Bayesian Compressive Sensing (BCS) [79], and OMP [69]. Nevertheless, there are several drawbacks connected with this technique. The presumption

that the support distribution is Gaussian; is not practical since it is not Gaussian in many real circumstances or unidentified. On top of that, its performance relies on the awareness of the Gaussian parameters and Bernoulli priors, which are generally challenging to calculate. Although a parameter estimation procedure is introduced, it relies on information on the preliminary estimates of those signal parameters. Subsequently, the estimating procedure has an adverse consequence on the complexity of the technique.

Another prominent Bayesian technique suggested by Larsson and Selén [80] calculates the MMSE estimate of the undefined signal, x . Its strategy is similar to FBMP because the sparse signal is a combination of several parts picked based upon a successive model reduction. Also, it necessitates familiarity with noise and signal stats. Still, it was observed that the MMSE estimation is aloof to the a priori parameters. Thus, a result of an empirical-Bayesian variation that does not call for any a priori information of the data was designed.

CHAPTER 5

DATA GENERATION

5.1 Introduction

Throughout this thesis, we have used the Fieldtrip toolbox to simulate the forward model. We used its built-in functions to simulate the head model, which has two main models, the source model, and the volume conduction model. Also, we utilized Fieldtrip functions to project and align the EEG electrodes on the head surface.

This chapter shows the technical simulation steps we followed on Fieldtrip toolbox to simulate EEG signals and generate the datasets used throughout the thesis.

5.2 High Spatial Resolution MRI Image

To produce EEG signals close to real clinical ones, the head model should be as close as possible to a real head in shape, dimensions, and -most importantly- volume conductivity. A lot of work was done in that field, and one of the popular ways is to use MRI imaging to get a high spatial resolution image of the head volume as all, not

just the surface, as shown in Fig. 5.1, a subject MRI image was used.

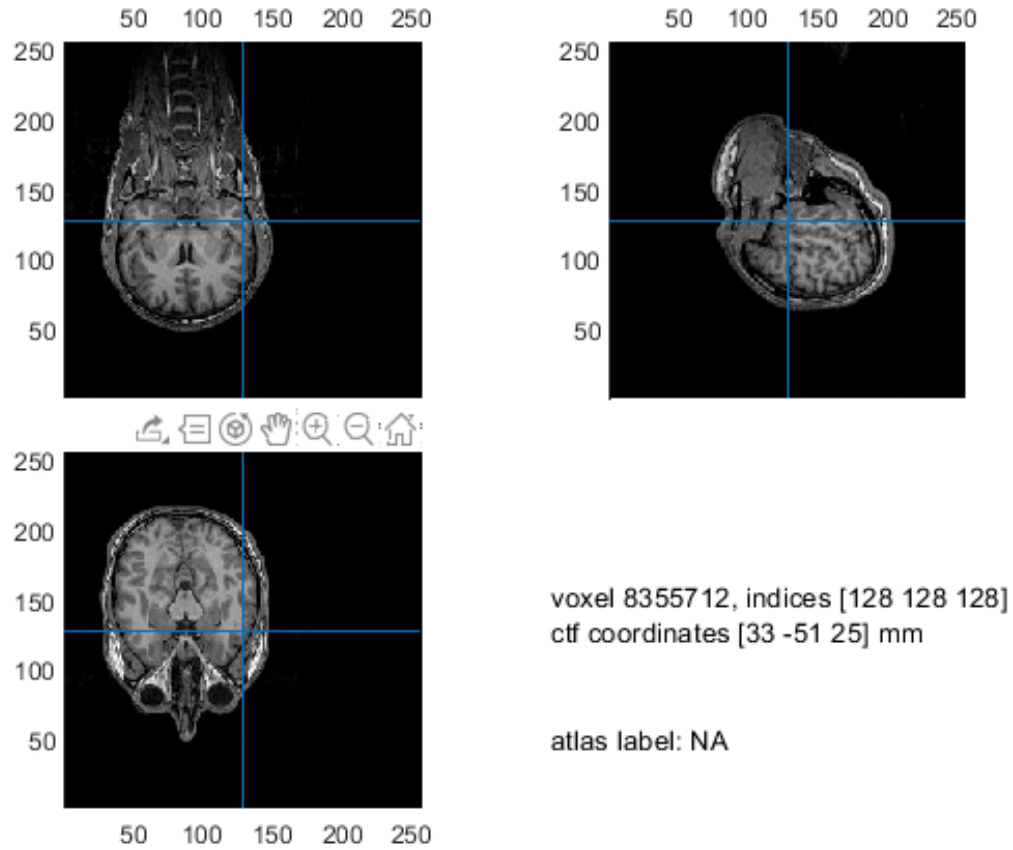


Figure 5.1: Raw MRI head image

Raw images need some preprocesses; therefore, Fieldtrip was used to realign the images with some standard coordinates system. In Fig. 5.2, the MRI image was realigned to match the head with the origin and axis.

5.3 Digitization

Some setups must be made to convert the problem into a computational one. Since computers prefer digital data, thus the problem must be converted from continuous to digital. Here, the contiguous volume of the head should be discretized into small

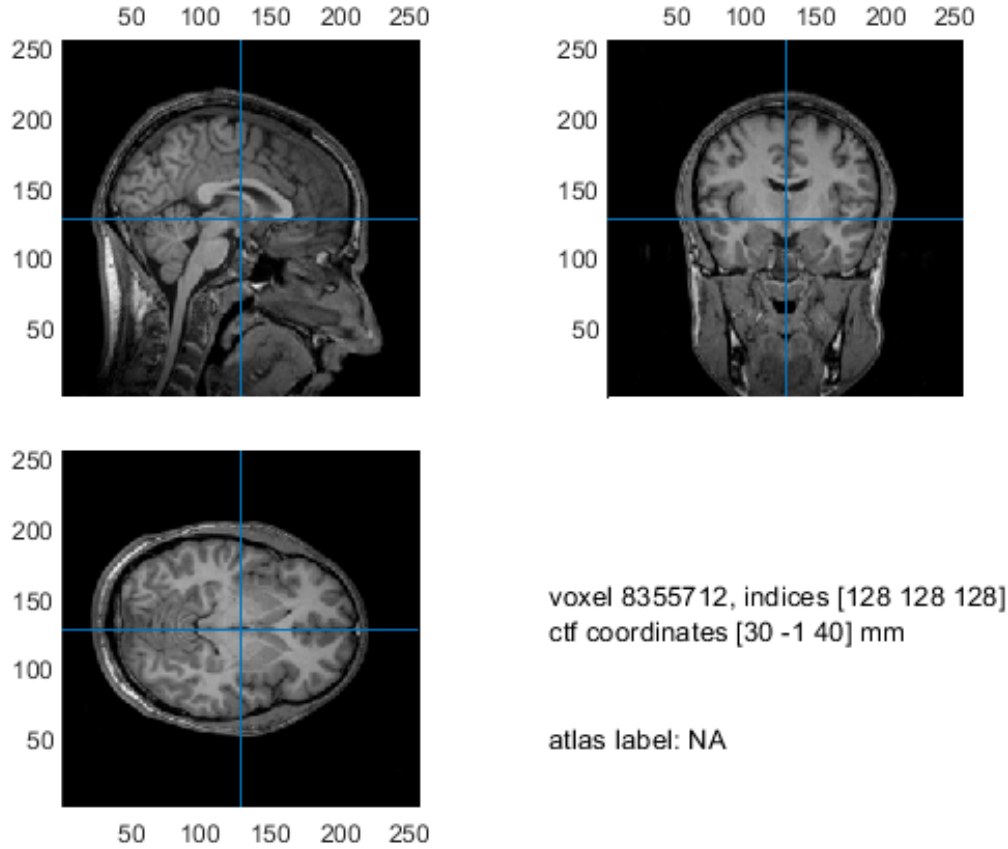


Figure 5.2: CTF coordinates aligned MRI images

cubes. Fieldtrip `ft_volumereslice` function was used to discretize the MRI image, as shown in Fig. 5.3.

5.4 Anatomical Segmentation

The most attractive feature of MRI in medical anatomical imaging is its high spatial resolution images on all three dimensions, which could be used to identify each physiological tissue inside the head. Therefore, with high 3-D resolution, we can differentiate between the main head tissues: scalp, skull, Cerebrospinal fluid (CSF), and gray and white matter. Then assign each their electrical properties. Fig. 5.4 shows a gray scale

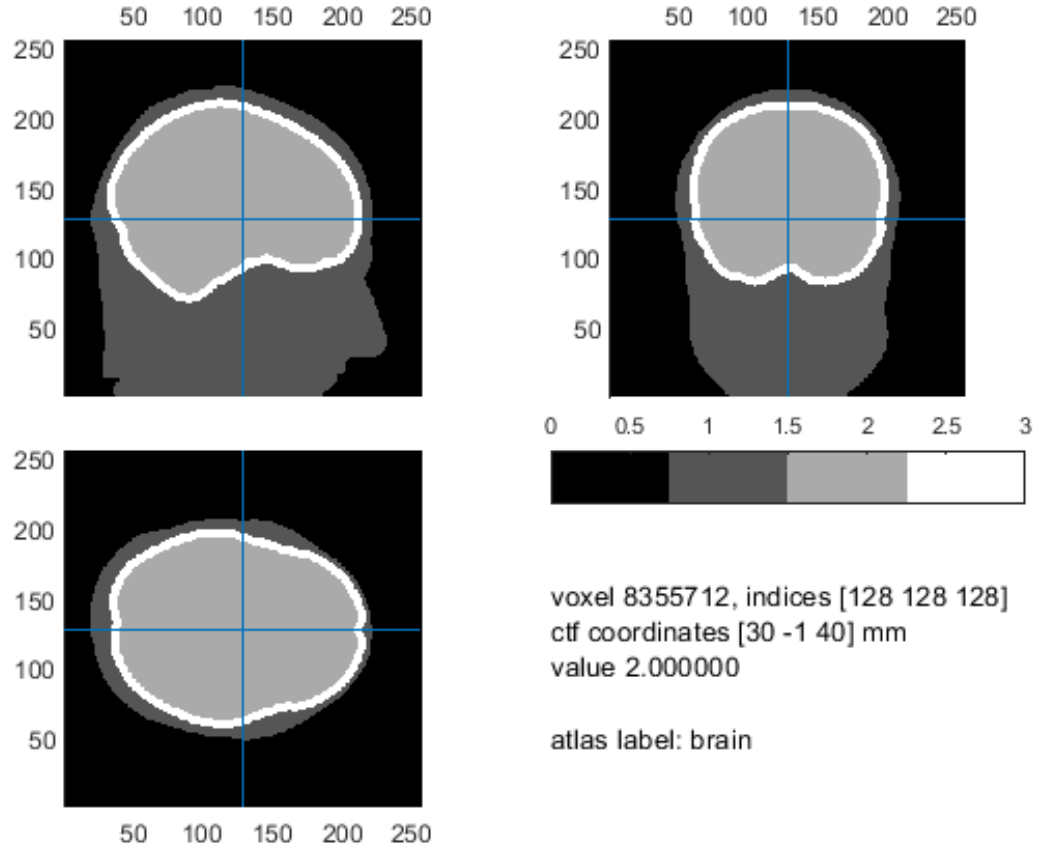


Figure 5.3: Head volume digitization

segmentation of an MRI image using Fieldtrip function: `ft_volumesegment`.

5.5 Electrodes Placement

There are many standards regarding EEG sensors. These standards mainly vary in the number of electrodes and their placement on the head. Since we are dealing with a highly underdetermined problem, we recommend using the standard with the highest number of electrodes. Until the day of writing this thesis, the 10-05 system (read: 10 to 5) is a proposed system -not produced as a product yet, with the highest and most dense head electrophysiological sensors. It defines the placement and distances

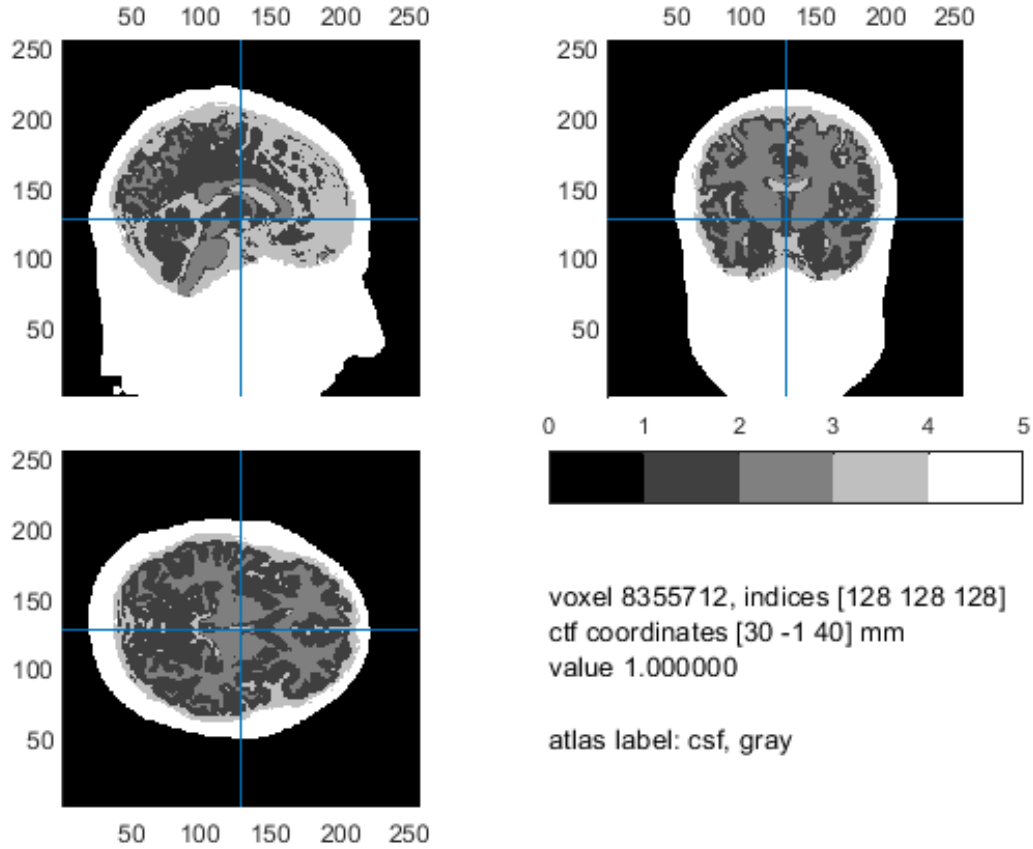


Figure 5.4: Head anatomy segmentation

between 346 sensors. It could be used to place both EEG and MEG sensors.

Since this problem is a spatial localization problem, the electrodes' exact fixed and aligned placement on the scalp is crucial. Thus, Fieldtrip `ft_electroderealign` function was used to project the electrodes on the scalp and align them with the facial fiducials, as shown in Fig. 5.5.

All forward problem components were simulated to the most realistic extent that current technology could provide. Fieldtrip has produced the function (`ft_prepare_leadfield`) that uses the computational Finite Element Method (FEM) to simulate the propagation of electromagnetic signals from the neuron sources inside the brain through the head volume until EEG electrodes are capturing them on the

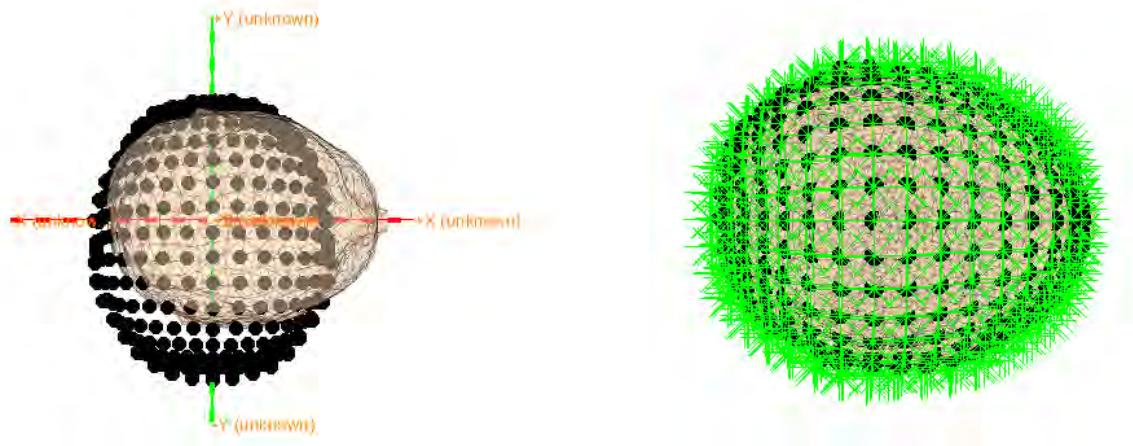


Figure 5.5: 10-05 system electrodes alignment with the head CTF coordinates system scalp.

CHAPTER 6

PROPOSED APPROACHES

6.1 Introduction

In our endeavor to solve the inverse problem, two main approaches were followed. We designated them as Single-stage approach and Multi-stage approach. These Stage divisions only concern the source space resolution, i.e., the resolution of the sources grid. Therefore, a stage means an increase in source grid resolution. The intuition behind that is to go around the high correlation in the EEG signals generated by very adjacent source points. More technically, we seek to formulate a leadfield matrix per stage whose basis vectors (columns) are nearly orthogonal, as illustrated more below.

This chapter introduces our stage-division approach into single-stage and multi-stage approaches. We illustrate how the leadfield matrix should be manipulated to simplify the implementation and turnaround difficulties.

6.2 Single-Stage Approach

The one-stage approach is just the direct way the problem is solved in one step using the whole-brain, thus, one source space resolution. The vector x contains all the Sources. This approach is straightforward; however, the complexity of the problem increases as the resolution of the source space increases. From a mathematical aspect, the size of the leadfield matrix increases as the number of sources within the brain increases, which makes the leadfield matrix more “fat.” Furthermore, the correlation between the leadfield matrix columns becomes much correlated since the increment in resolution makes source points closer to each other.

At first, when we started working on the problem, we were using a low source space resolution. The displacement between adjacent source points was 2 centimeters. The accuracy of reconstruction was very high for one active source. When the number of source points in the grid increases, reconstruction accuracy worsens. Of course, for the cases of more than one active source, the reconstruction accuracy is not that good in all grid formulations, As shown in the results.

6.3 Multi-Stage Approach

The multi-stage approach is a way that intends to simplify the solution by dividing the solution space into sub-spaces. The intuition behind that is to overcome the high correlation between adjacent sources by considering them as one source point. And to simplify the highly underdetermined system into a nearly determined or maybe overdetermined system. The source grid is divided into smaller regions, equal in

volume, dimensions, and shape. Therefore, we came up with two grids we will call coarse grid and fine grid. The coarse grid is a grid in which each region is represented by an only one-source point, and the fine grid is the grid within that region but with much fine resolution.

Fig. 6.1 reflects the idea of coarse grid, where each point virtually represents one region; those regions are similar to the one shown in Fig. 6.2. The number of regions depends on the selected dimensions and shapes of regions, which is subject to an optimization process regarding the resultant number of fine grid points in each region, and whether this number of source points is optimum for Deep Learning models.

6.4 Methodology

We have examined several techniques to solve the inverse problem of EEG signals. Two disciplines were applied to find the location of neurons responsible for the measured EEG signals: Compressed Sensing and Deep Learning.

6.4.1 Compressed Sensing Reconstruction Algorithms

Due to the sparse nature of brain neurons' activation, sparse reconstruction algorithms are considered promising techniques to solve the ill-posed EEG inverse problem. Within this thesis, compressed sensing techniques have been only applied to the EEG spatial data, which are one-time sample EEG signals. As described in the data generation chapter, these signals only consider the spatial information of EEG signals. On the other hand, advanced compressed sensing reconstruction algorithms may need

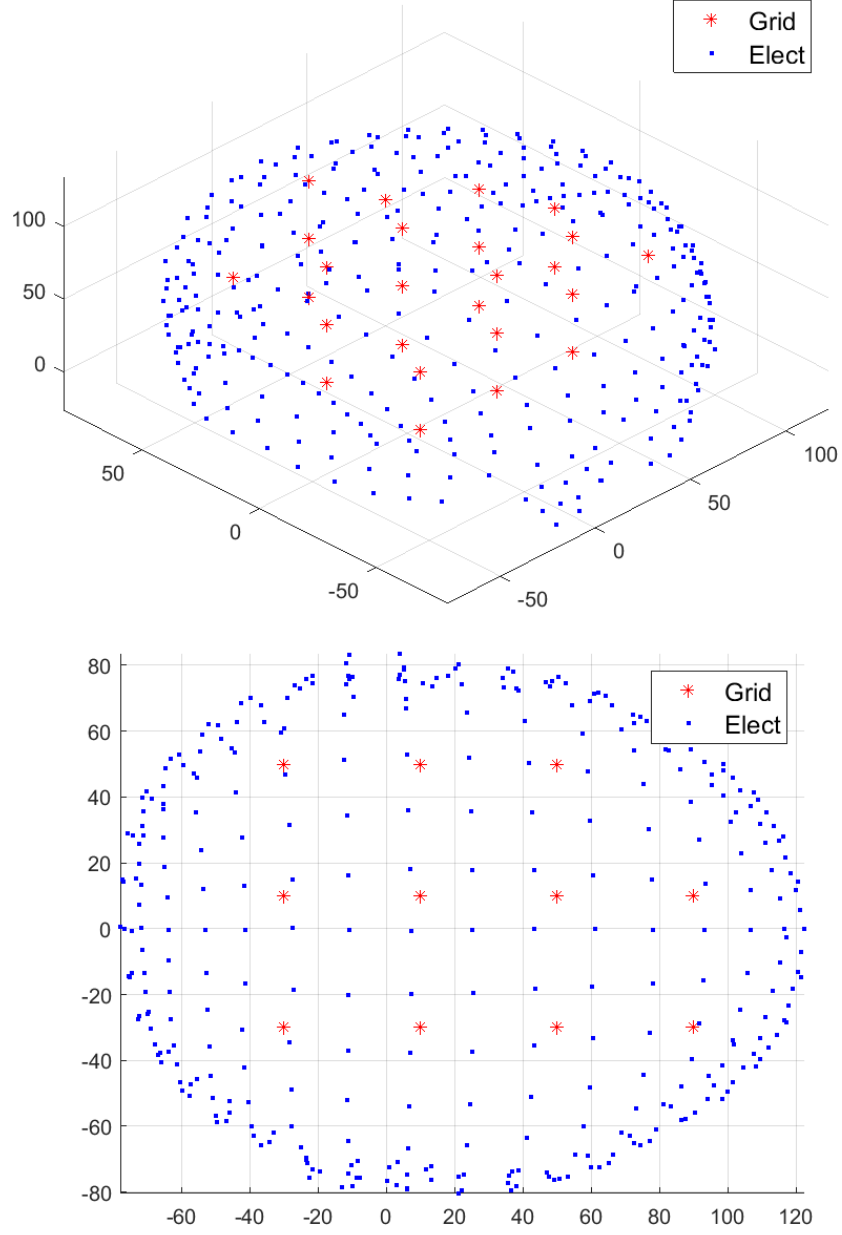


Figure 6.1: Center points of coarse grid regions

to be developed regarding the time-varying EEG signals, which contain both space and time information, as recommended in the future work chapter.

The applied sparse reconstruction algorithms are OMP, StOMP, and CoSaMP. As the iterative procedure of compressed sensing techniques, their reconstruction approach basically depends on finding the highest correlation between the measured

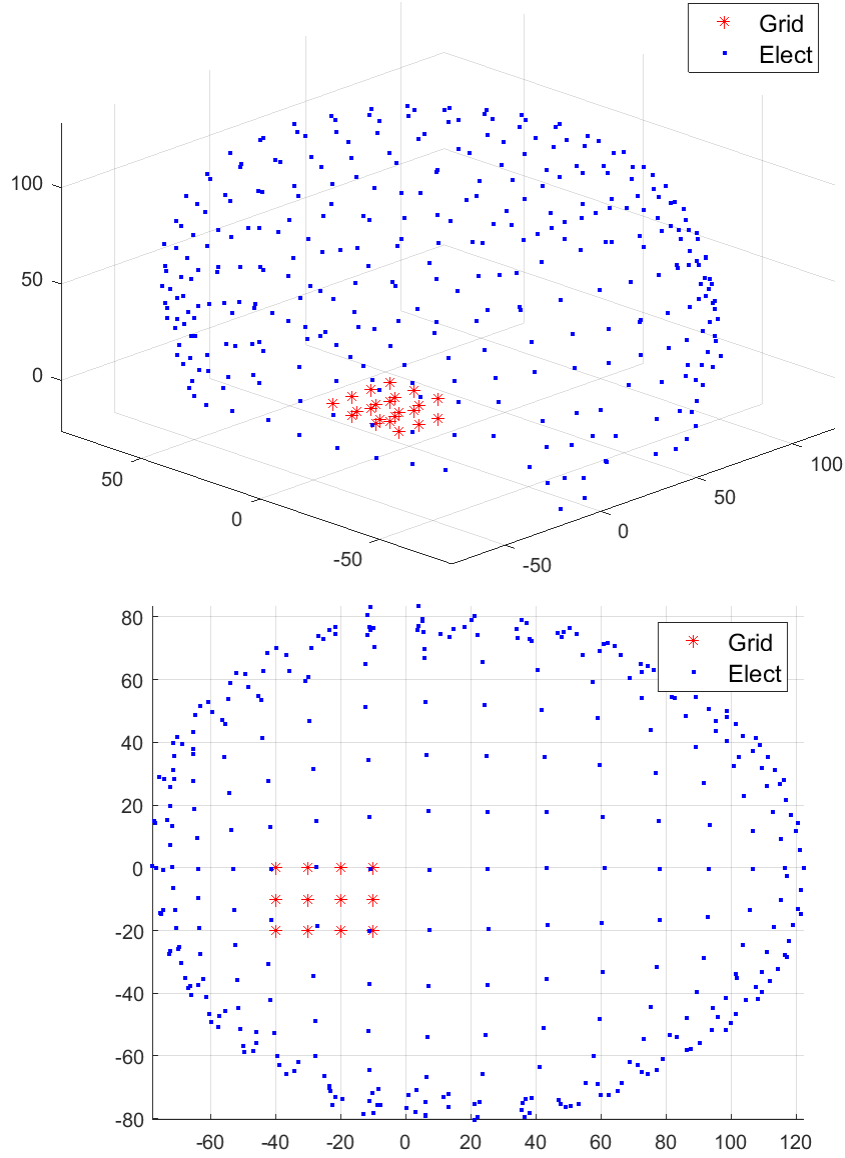


Figure 6.2: Fine grid points of one region

signal y and the corresponding bases vector within the lead field matrix. That is why it is evident that sparse reconstruction techniques perform very well in the cases of single active source and uncorrelated active multi-sources. For example, when more than one source is active at a time and they are far apart in space or very close to the cortex surface, thus very close to measurement sensors. Therefore, their basis vectors in the leadfield matrix are expected to be highly uncorrelated.

Therefore, our approach is to apply compressed sensing techniques to solve the two problems: the problem of single-active source and the problem of cortex-active sources. For the single-active source localization problem, the location of the source to be reconstructed does not matter, whether it is deep inside the brain or on the surface of the brain. Thereby, our approach is to utilize sparse reconstruction algorithms to solve single-active source localization problem for both cortex brain activity and deep brain activity.

For the multi-active source localization problem, our approach is to apply compressed sensing algorithms to localize only brain surface activity. The complexity of the inverse problem increases as the number of active sources increases. Furthermore, the possibility of correlation increases is higher. Therefore, iterative techniques are not encouraging in this case unless the active sources' signatures are very distinct from each other. For all of that, we limited the application of compressed sensing reconstruction algorithms to cortex activity localization only. However, the multi-source activity could be addressed using iterative techniques by utilizing the low dynamic nature of EEG signals, as recommended in the future work chapter.

6.4.2 Deep Learning Mapping Structures

The successful application of Deep Learning models in many highly complicated problems makes it an attractive technique to address the long-hanging brain source localization problem. Regarding this problem, the key task to solving the problem is to address the inverse problem. Unlike the forward problem, there is no closed-form solution to the inverse problem yet. In other terms, the inverse problem is not modeled

in a closed-form mathematical function.

On the other hand, nowadays, the brain forward problem is being modeled to a high degree of accuracy after the enormous efforts exerted by researchers, especially in the fields of computational electromagnetic and neuroscience. But, the inverse problem is still an issue; since there is no direct mathematical way or derivation to put it in a close form solution. Thereby comes the benefit of Deep Learning models. Despite that, since neural networks are superior in mapping the function governing the relation between any two sets, they could also be utilized to imitate a closed-form solution to the brain inverse problem.

This thesis proposes to map the scalp EEG signals to their corresponding neuron sources using Deep Learning architectures. Our work is to train a Deep Learning model to map the EEG signals dataset to the labeled active-sources vectors dataset. The active-sources dataset could easily be labeled using the available forward model computational techniques. Also, keeping in mind the data-driven model's constraints, we propose to apply the Deep Learning architectures to the multi-stage approach only. Data-driven solutions depend more on the data rather than the algorithm or model.

Consequently, much of the effort should be exerted on the data. On that account, our proposed model emphasizes providing a large enough amount of well-distributed data. Since, in the forward model, we are modeling the EEG signals as the response of the activation of current dipoles inside the brain, the dataset for the model should be generated to include all possible magnitude and direction activation of each current dipole. Nevertheless, the number of dipoles is large (here in thousands) as the source space resolution is increased. Therefore, the amount of data that needs to be generated

becomes enormous (in this approach, in terms of Gigabytes).

For all the above, our Deep Learning structures proposal is to be applied on a multi-stage approach only. Thus, the expected substantial overall dataset size shall be divided into smaller datasets per region. As the multi-stage approach divides the whole-brain volume into smaller regions, the proposed Deep Learning architecture shall develop a neural network module for each region rather than one extensive neural network for the whole-brain volume. Finally, one Deep Learning model shall be built and trained over those smaller individual modules. This governing model's job is to act as a link between the captured EEG signals on the scalp and the regions' modules.

We have many small Deep Learning models activated by a governing module according to the expected active dipole region. The individual module within that region pin-points the active dipole location.

CHAPTER 7

IMPLEMENTATION, RESULTS, AND DISCUSSION

7.1 Introduction

After we have introduced our proposed approaches to address the problem under study, as well as all the techniques related to the formulation of the problem, like the neuroimaging brain mapping and modeling technology, and the electrophysiological EEG signals, through the presentation of compressed sensing and Deep Learning disciplines whose performance to be investigated as they are applied to the problem, and the simulation of the signals and datasets essential to the problem; we put our work into the tests.

Using MATLAB, we studied the application of OMP, StOMP, and CoSAMP algorithms on the ℓ_0 -minimization problem in (4.2). To simplify the readability of the thesis while keeping the flow of contents intact, we have included the implementations and the results consequently in one place rather than splitting them up into different

chapters.

This chapter directly includes all of our implementations alongside their results and discussion. We start with the applied compressed sensing reconstruction algorithms on both low-resolution and high-resolution source space models. Then, we present the implementation of source space division stages before showing the developed Deep Learning models. Afterward, the spatial and Spatio-temporal learning structures, Residual network, and LSTM modules are explained thoroughly. We finalize with the overall whole-brain structure that accumulates the results of all modules together.

7.2 Compressed Sensing

To investigate the appliance of sparse reconstruction techniques, we started by applying OMP, StOMP, and CoSaMP algorithms on a straightforward setup; without noise, with one active source and one-component source. By simple, we mainly mean a meager number of source points per cubic centimeter, considering a few source points and keeping the sparse vector x short. Thus, reducing the effect of correlation between nearby sources on source reconstruction accuracy. Then, we increase resolution and set up complexity step-by-step.

7.2.1 Low-Resolution Source Model

Fig. 7.1 shows that the low-resolution is one source point per 2 cubic centimeters (i.e., $0.125 \text{ source/cm}^3$). Thus, the resultant total number of low-resolution grid points is 186 source points inside the brain. This sparse grid makes the inverse problem an

overdetermined rather than an underdetermined since the number of measurement vector elements – EEG sensors- is 345.

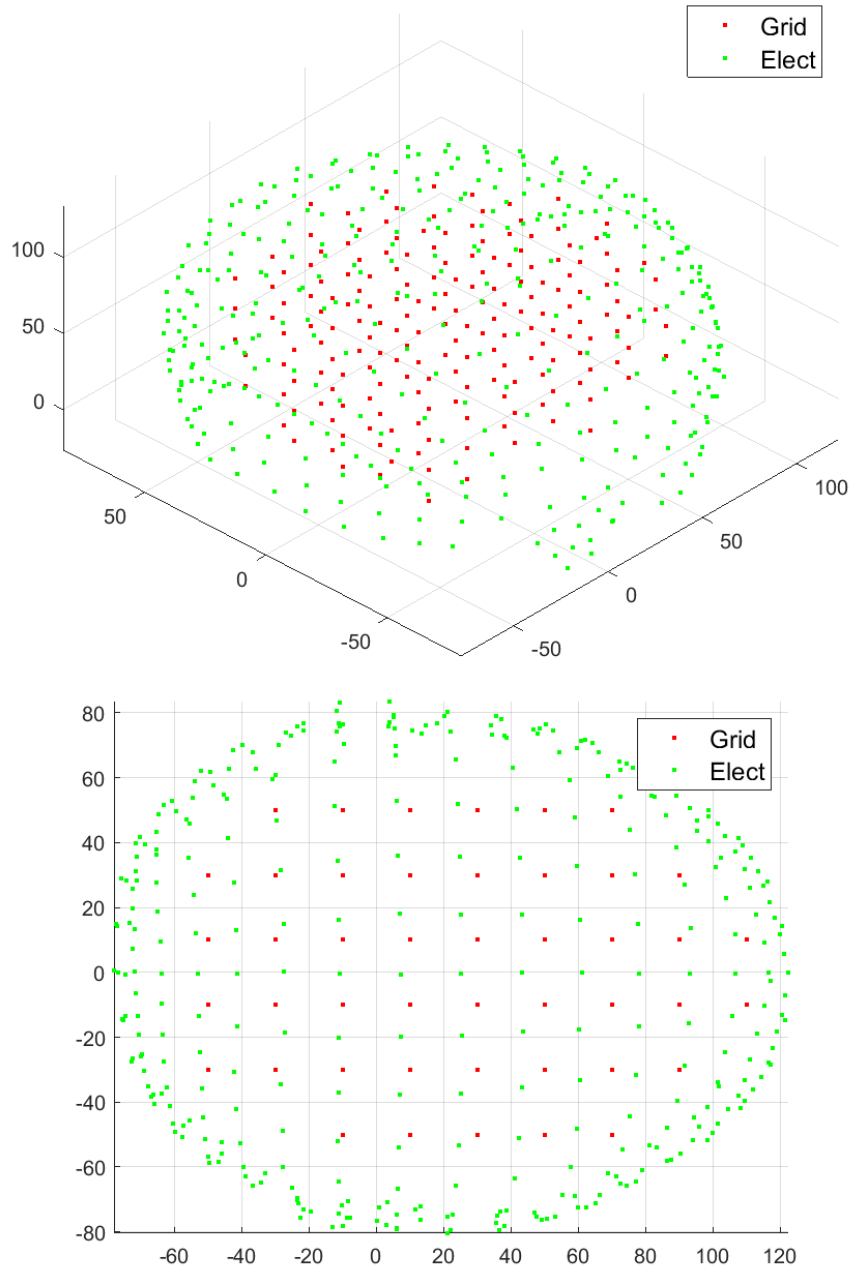


Figure 7.1: Source points of low-resolution source space

Here along with low resolution, we consider simulating source dipoles that are all oriented in one direction. More specifically, since neuron sources are simulated

as current dipoles, only one component of these current vectors is considered rather than the 3-component vector. That could be implemented simply by extracting the undesired components from the leadfield matrix. The extracted leadfield matrix will be one-third of the full leadfield matrix generated using Fieldtrip toolbox.

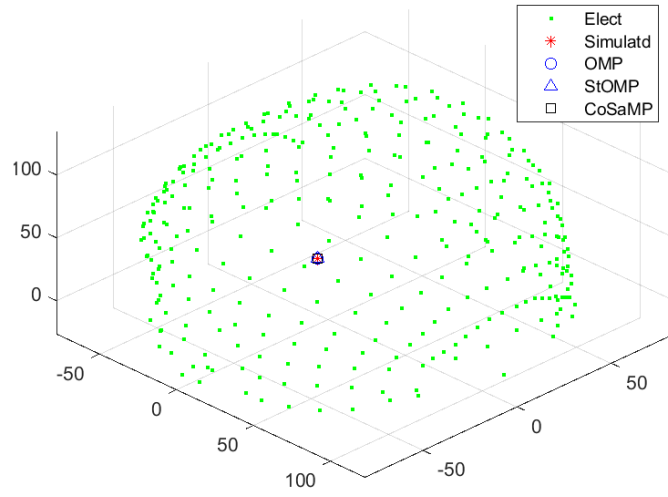
Table 7.1: Low-resolution source model parameters

Parameter	value
Displacement between source points	2 cm
Source points resolution	0.125 source/cm ³
Total source space points	186
number of EEG electrodes	345

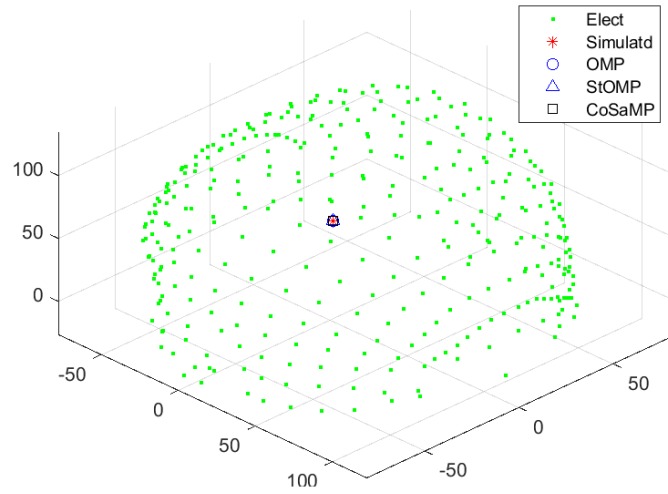
With the setups mentioned above, summarized in Table 7.1, , trials of one active source were simulated on the cortex and deep inside the brain. Then in each trail, compressed sensing reconstruction algorithms OMP, StOMP, and CoSaMP were applied. The algorithms inputs are the measured EEG signals, leadfield matrix, and the number of active elements. As a result, the algorithms were able to locate the stimulated source precisely, as Fig. 7.2 shows the results of 3 trial examples. Fig. 7.2 (a) shows the reconstruction of a deep brain source, while Fig. 7.2 (b) and (c) show the reconstruction of cortex brain sources.

Then, the complexity was increased by simultaneously simulating two active sources. Fig. 7.3 shows the localization of OMP, StOMP, and CoSaMP algorithms. OMP sometimes fails to even calculate a solution due to high correlation. And, CoSaMP sometimes fails to predict the exact location, as Fig. 7.3 (c) shows an example of one case. Fortunately, StOMP seems very good and promising.

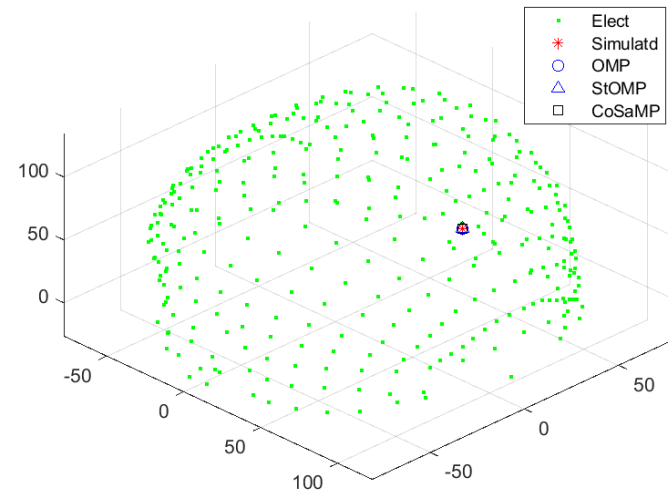
This failure of both OMP AND CoSaMP is not encouraging since we are only



(a)

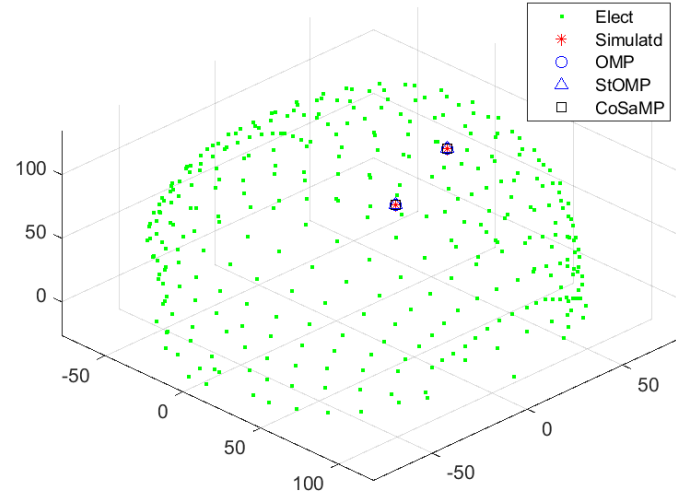


(b)

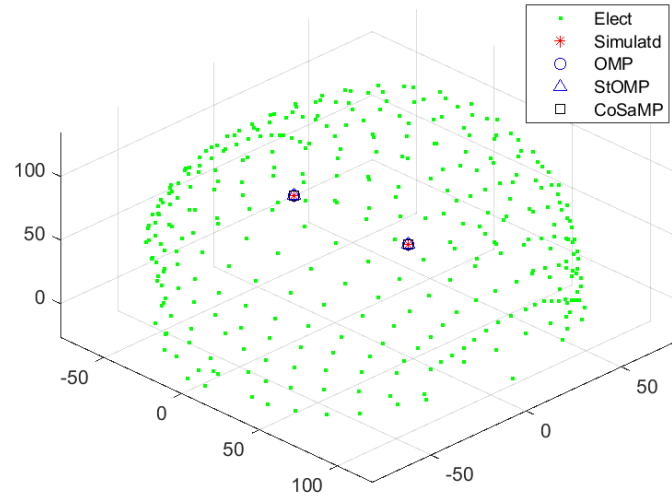


(c)

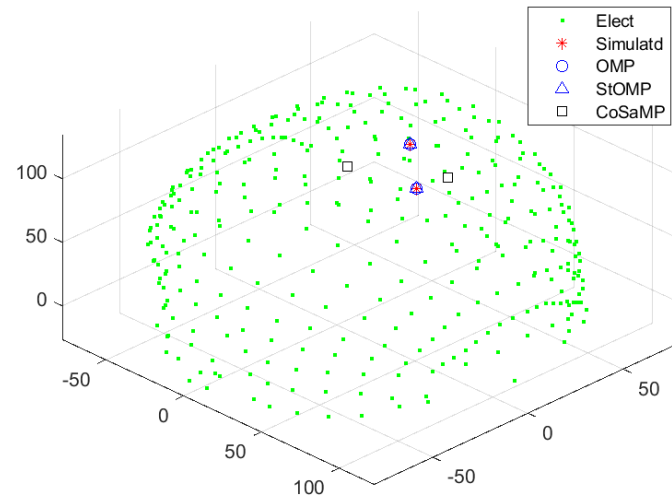
Figure 7.2: Low-resolution one-component source space: One active source reconstruction using compressed sensing algorithms



(a)



(b)



(c)

Figure 7.3: Low-resolution one-component source space: two active sources reconstruction using compressed sensing algorithms

using an effortless setup, and the problem is overdetermined, which is considered a simple inverse imaging problem.

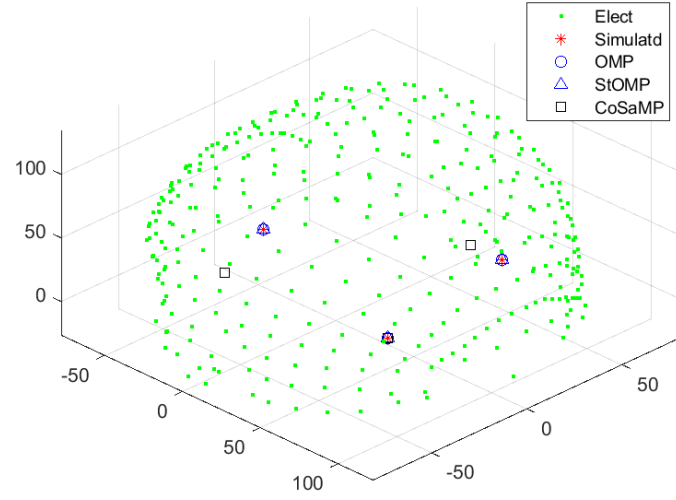
In the next step, we added one more stimulated active source. So now, three sources are active at a time. Fig. 7.4 shows the localization results of reconstruction algorithms. Although this figure shows exact results for OMP, it sometimes fails, and no prediction is even computed in that case. Also, it is observable that CoSaMP fails a lot, and StOMP is still promising.

Finally, 5, 7, and 10 active sources were simulated, as shown in Fig. 7.5 (a), (b), and (c), respectively. As well as before, OMP and CoSaMP fail a lot, and there is hope in StOMP.

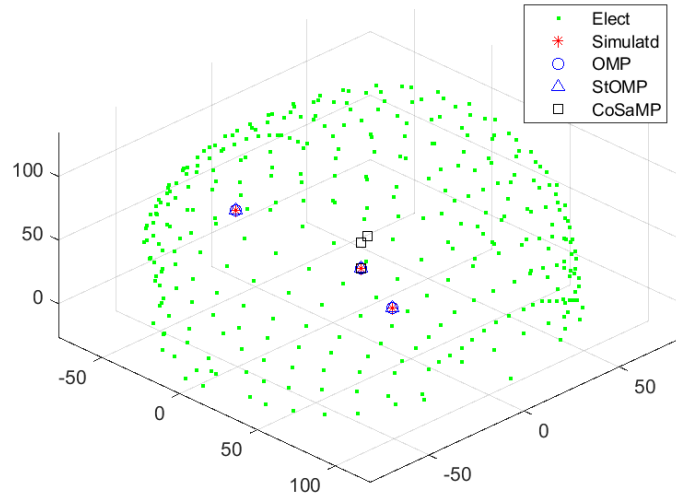
Due to the failure of both OMP and CoSaMP in reconstructing even the noiseless data, we will skip the investigation of their appliance on noisy data now and go ahead to check whether the promising algorithm, StOMP, will work on a high-resolution source model.

7.2.2 High-Resolution Source Model

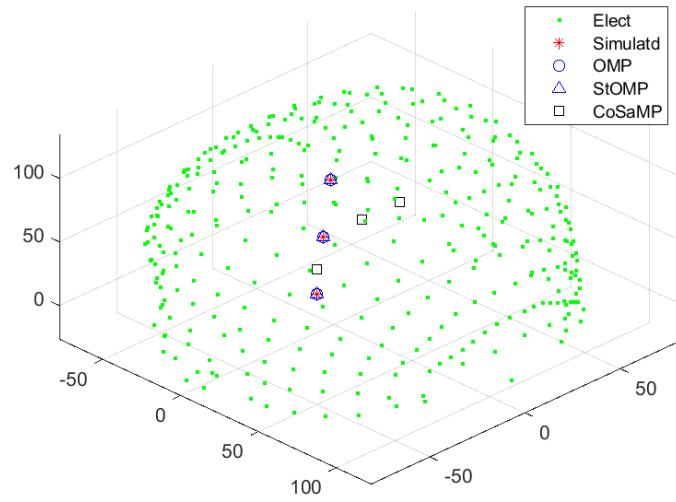
Despite the discouraging results of OMP and CoSaMP, they will be tested again, this time in a higher resolution source space. Nevertheless, mainly, the promising performance of StOMP is to be verified on a much rough setup. As summarized in Table 7.2, the source grid resolution was increased eight folds for the following steps. Thus, the new resolution is one source point per cubic centimeter (i.e., 1 source/cm³). The number of points in the whole grid now becomes 1,475 points. The vector x to be reconstructed is much longer, therefore, much sparser. The inverse problem returns to



(a)



(b)



(c)

Figure 7.4: Low-resolution one-component source space: three active sources reconstruction using compressed sensing algorithms

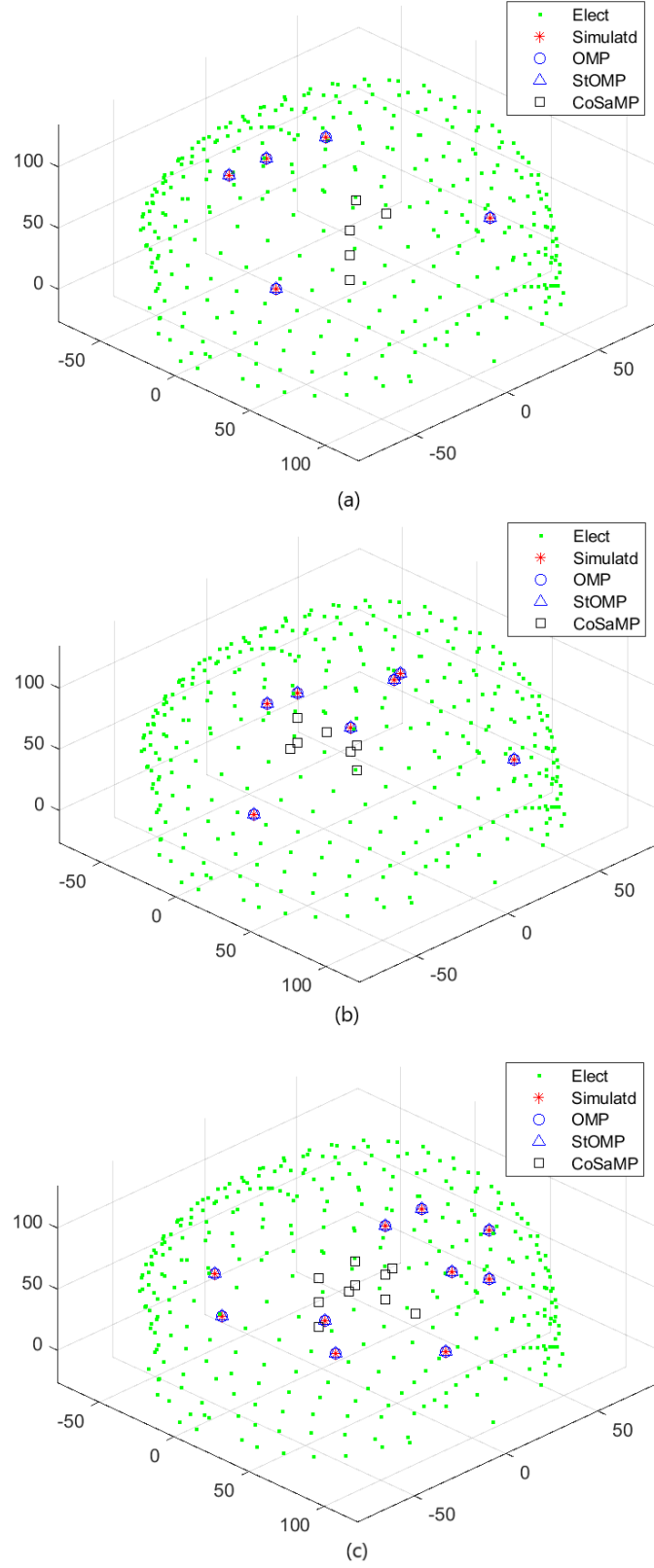


Figure 7.5: Low-resolution one-component source space: multi-active sources reconstruction using compressed sensing algorithms. (a) five active sources, (b) seven active sources and (c) ten active sources

be complex since it is now ill-posed; However, we are using the densest EEG electrodes placement standard -the proposed 10-05 standard, which contains 346 EEG electrodes.

The problem is underdetermined; the leadfield matrix is fat.

Table 7.2: high-resolution source model parameters

Parameter	value
Displacement between source points	1 cm
Source points resolution	1 source/cm ³
Total source space points	1475
number of EEG electrodes	345

Again, we started with a simple setup, one active source at a time, then the number of active sources increased one at a time. Fig. 7.6 shows the localization accuracy of compressed sensing algorithms altogether in the same figure. OMP and CoSaMP performed very well, with exact reconstruction. However, and unfortunately, StOMP, which we believed would be a promising algorithm, fails. It tries to locate the source but with a colossal localization error.

Next, the number of active sources was increased by one. As Fig. 7.7 shows, CoSaMP is still very good; it can reconstruct the two sources precisely. However, OMP and StOMP sometimes fail in reconstructing one of the sources, although Fig. 7.7 does not show OMP failures, as mentioned earlier, due to OMP failure of even computing the solution.

Then, three active sources were simulated. The localization results are shown in Fig. 7.8. Unfortunately, now all three algorithms failed. Therefore, we can notice that CoSaMP reconstructions follow only one source. It is like it is only trying to reconstruct one source.

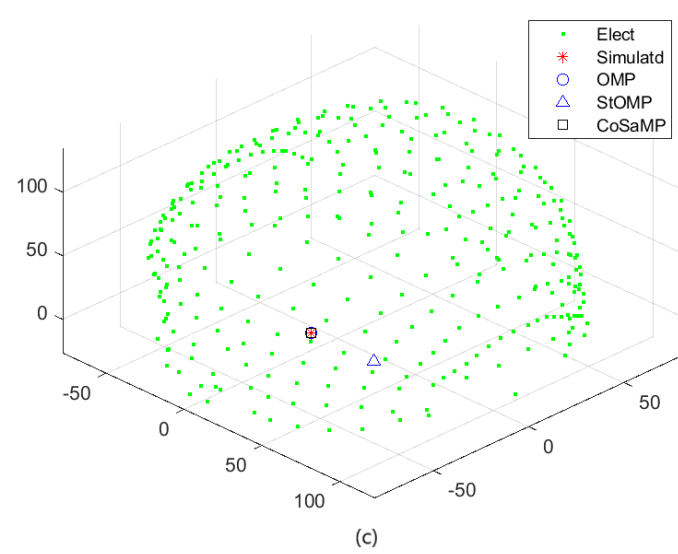
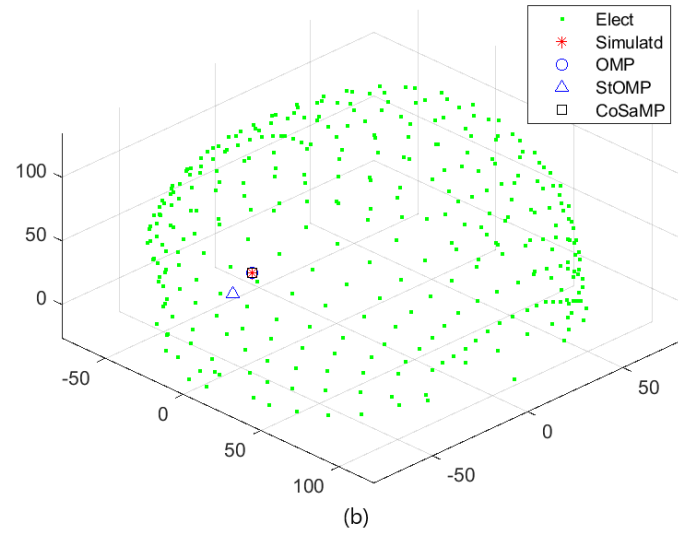
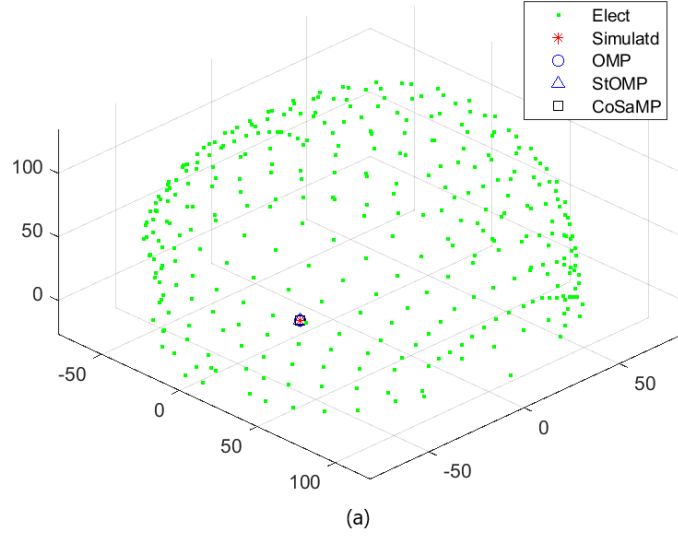


Figure 7.6: High-resolution one-component source space: One active source reconstruction using compressed sensing algorithms

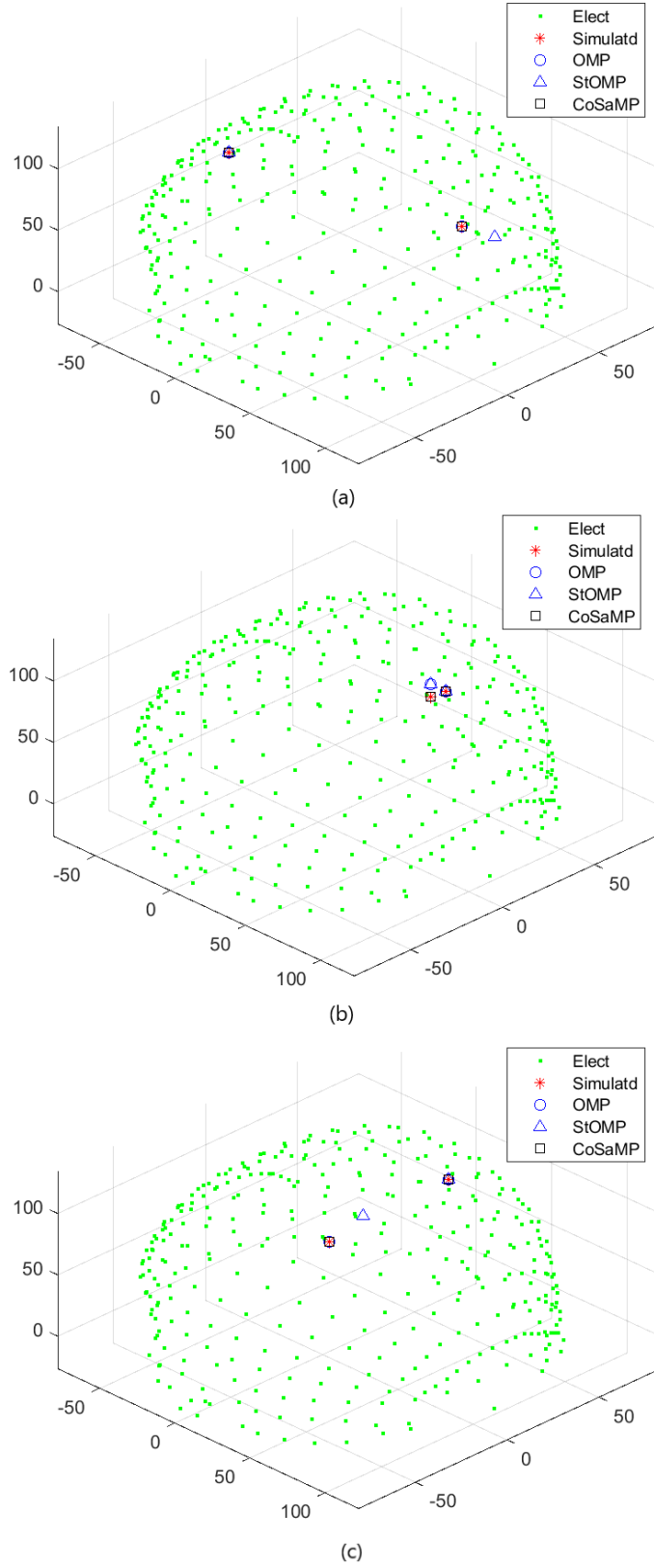


Figure 7.7: High-resolution one-component source space: two active sources reconstruction using compressed sensing algorithms

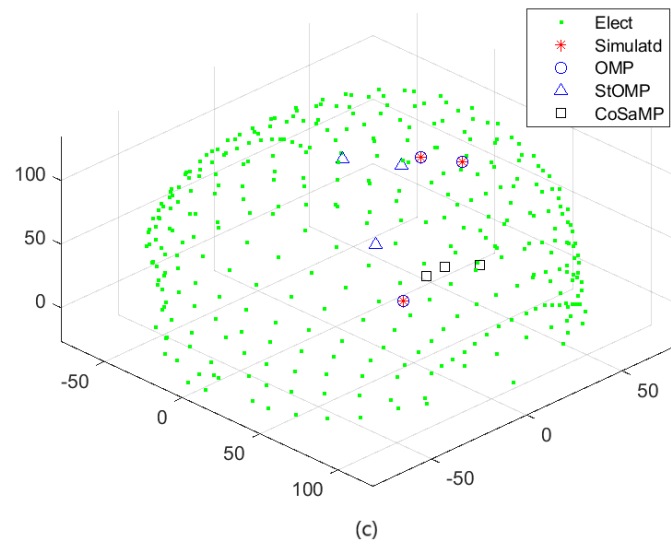
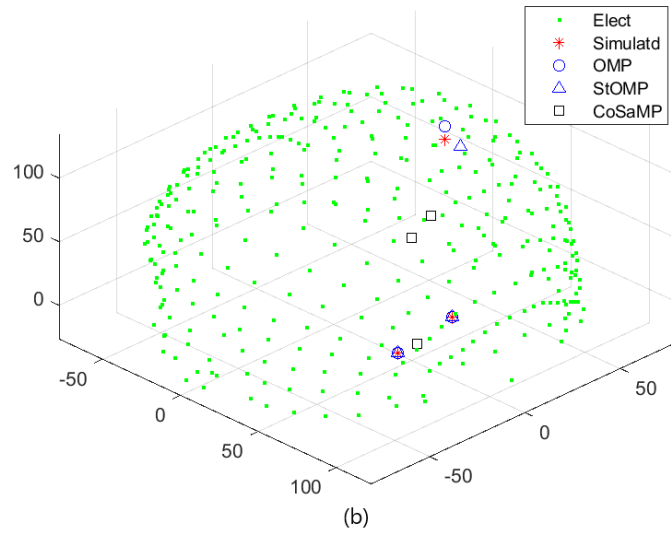
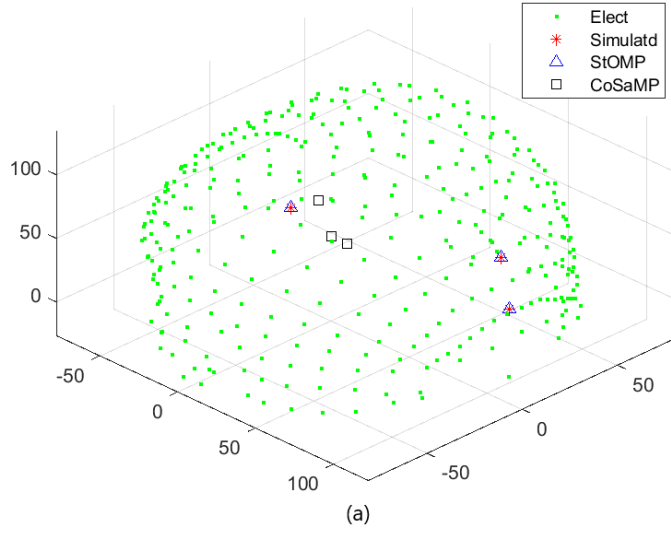


Figure 7.8: High-resolution one-component source space: three active sources reconstruction using compressed sensing algorithms

Finally, five active sources were simulated, and their localization results are shown in Fig. 7.9. The figure shows that all algorithms failed to reconstruct, usually, even one source.

7.3 Deep Learning

7.3.1 Source Space Division

Besides, our main aim of proposing the multi-stage resolution approach is to solve the brain source inverse problem; another side purpose is to simplify the computation headage of the Softmax activation layer. Although most deep learning modules use Softmax activation, it is well-known that its computation increases exponentially as the number of elements in the output layer increase. So, despite the many ways followed in literature to simplify this computational power consumption, like hierarchy classifiers, we prefer to divide the outputs into smaller groups. Therefore, Softmax computation will depend on the number of elements within that group, not the whole problem output (the total brain source points).

This staging process could be implemented by dividing the source space grid into coarse and fine grids. Unlike the fine grid, the coarse grid is not real; it is a virtual grid. A coarse grid could be implemented as a binary output vector with elements equal to the number of regions. Each element indicates whether its corresponding region is active or not. The fine grid resolution is better to be as fine as possible, which depends on the resolution of the source model generated by solving the forward problem using the simulation software. As the number of simulated source model points increases, the

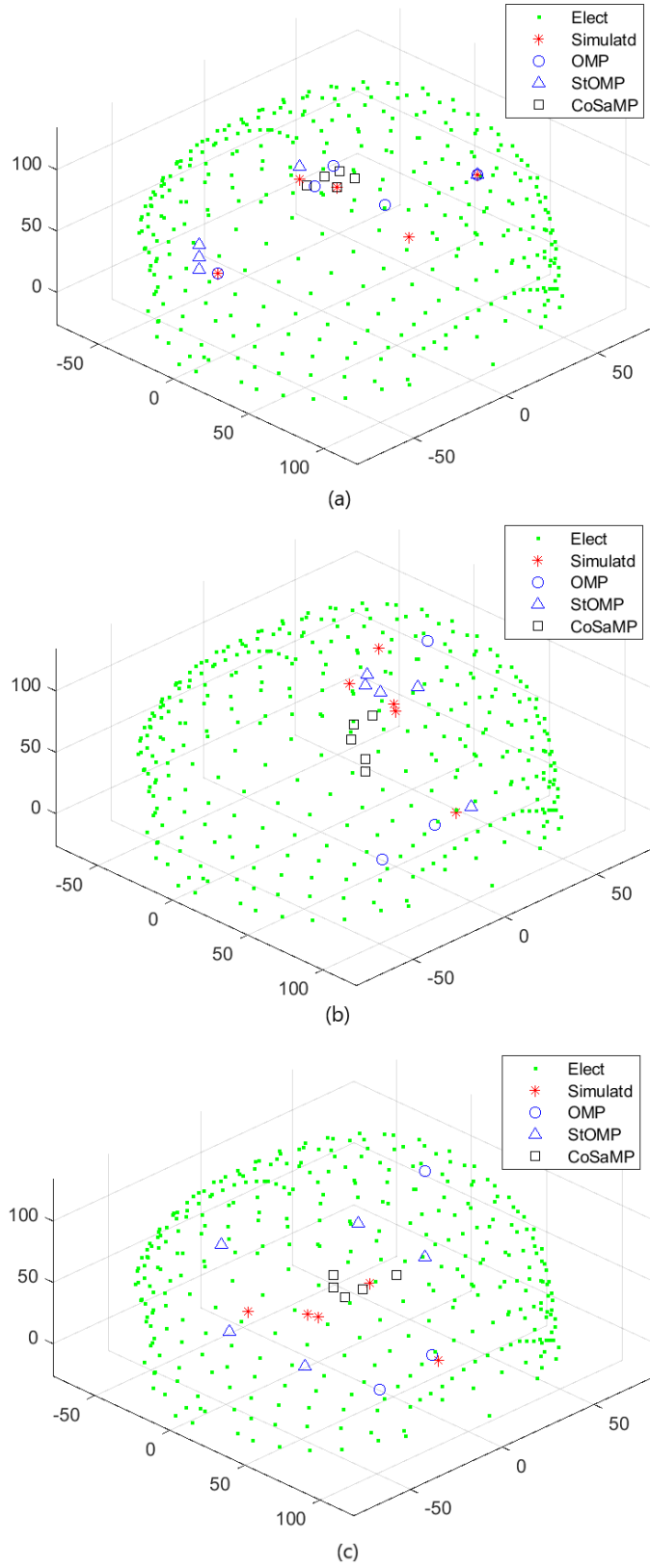


Figure 7.9: High-resolution one-component source space: five active sources reconstruction using compressed sensing algorithms

forward problem simulation's required computational power increases exponentially. Then, the leadfield matrix will become huge; thus, its computation will take a long time and memory. Fig. 7.10 shows the fine grid resolution, which, as stated earlier, is the same as the most satisfactory resolution, which is one source point per cubic centimeter (1 source/cm³).

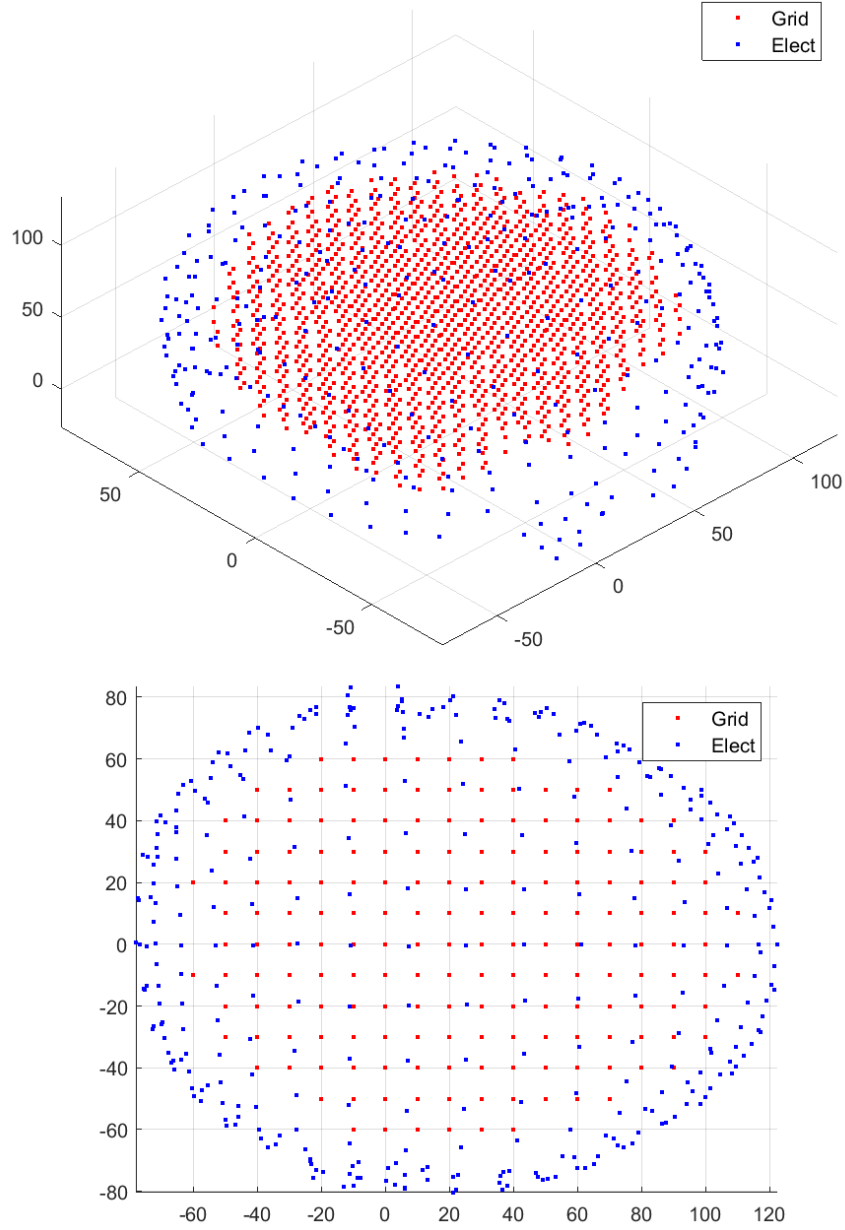


Figure 7.10: Source points of high-resolution source space

Initially, only three regions were constructed to simplify implementation and reduce the amount of data generated, thus reducing the size of datasets. Fig. 7.11 shows those implemented regions' positions and shapes. The main idea is to create regions with the same shape, dimensions, and number of source points. However, due to the irregular shape of the regions adjacent to the brain's surface, we intended to simulate regions that have a slightly different shape from each other. However, the main simulation parameters must be kept general to represent the rest of the non-constructed regions' behavior. For example, the number of source points within each region is kept equal, 22 source points per region. Two regions were placed adjacent to each other, we called them Region1 and Region2, and they were designed with the exact shape. Their dimensions are 30x20x20 mm. While the third region, Region3, was positioned on the other side of the brain with different dimensions of 30x30x20 mm. The idea behind the adjacent and apart regions is to consider the effect of correlation in EEG signals generated by close sources.

After experimenting with the simple initial setup, we can go to the final setup, where regions division will be expanded to cover the complete brain volume. In this step, the whole-brain volume was divided into 56 regions; each region is 40x40x40 mm, as shown in Fig. 7.12. All regions' dimensions are kept the same wherever their location is to keep the implementation as simple as possible. The choice of these dimensions could be different according to designer selection. However, we have chosen these values since the spread of regions over the whole-brain, and the distribution of source points within each region are consistent with these cube dimensions. The maximum number of source points within these new regions is 64 sources, and it goes down for

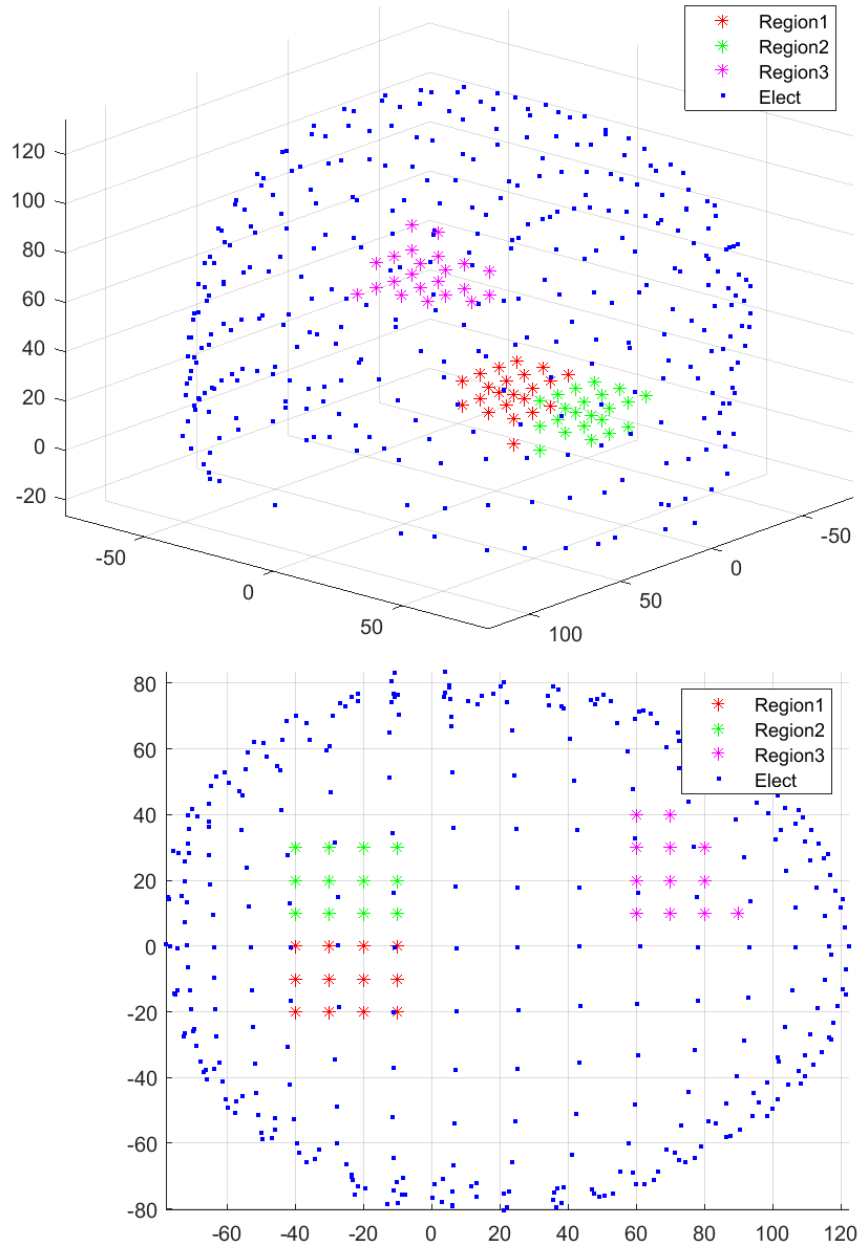


Figure 7.11: Simulated regions, Region1 and Region2 are adjacent while Region3 is placed on the further side

regions on the edges of the brain, like some regions contain only one source point.

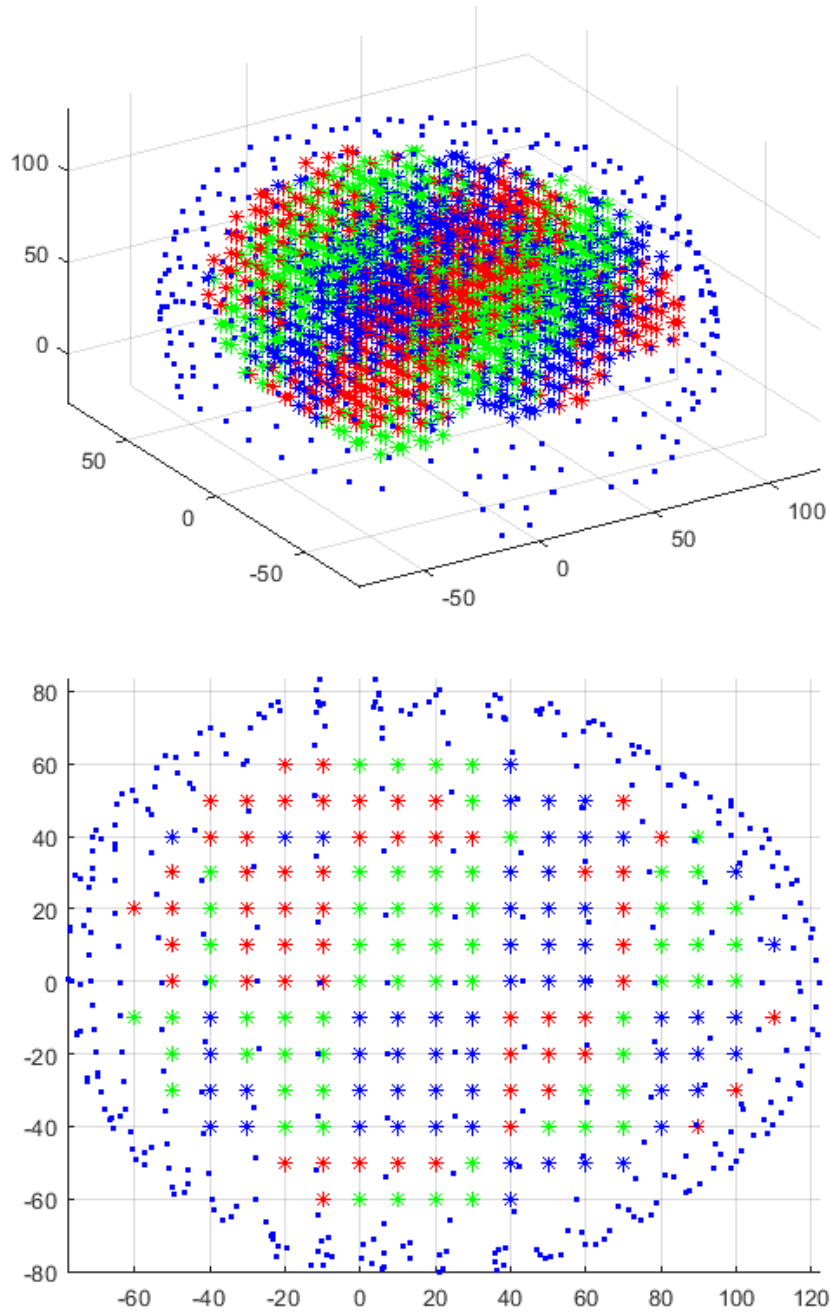


Figure 7.12: Full brain volume division into regions colored in red, green and blue

7.3.2 Learning Setups

All datasets generated from each region were divided into training, validation, and testing datasets with approximate percentage divisions of 70%, 20%, and 10%, re-

spectively.

For all Deep Learning structures, Adam optimizer was chosen as the optimization algorithm, and the loss function was calculated using Categorical Cross-Entropy.

7.3.3 Simplified Residual Networks Model

As discussed earlier in the methodology chapter, ResNets were selected as the Convolutional Neural Networks to detect the features that differentiate between neuron sources responsible for the propagated EEG signals. ResNets were applied within this experiment to localize the active source points in coarse and fine grids. The same ResNet structure, layers, filters, and inputs' shape were used for both grids. The only difference is the number of output elements. Even though that may be just for the current case since just three regions are currently implemented to verify our approach. Thus, only 3-elements output is required for the coarse grid module. But, of course, that will not be the case when the whole-brain volume is considered.

A mother ResNet module was implemented with five convolutional blocks, one skip connection, a fully connected layer, and a Softmax prediction output layer. Fig. 7.13 shows the structure of this mother ResNet with the functional blocks that it consists of.

All blocks and the skip connection consist of a convolutional layer, activation layer, and MaxPooling layer. All convolution filters have the same 3x3 kernel size. A Stride of one and Valid padding are used in all convolution operations except for the third and fourth blocks, where the Same padding is used. The number of convolution filters increases gradually: 16, 32, 64, 128, 256, and 512, as blocks go deeper from

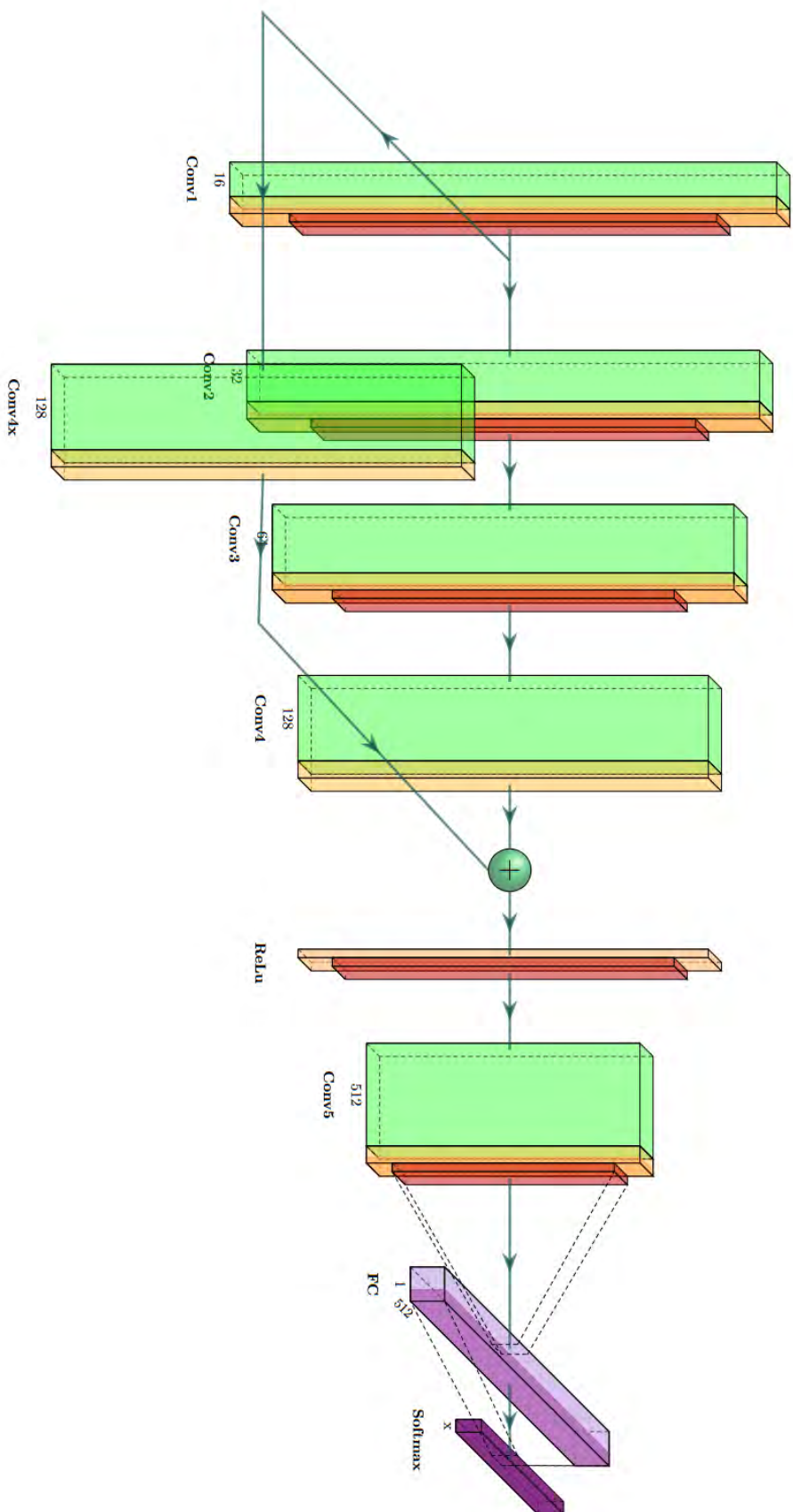


Figure 7.13: The implemented ResNet structure

inputs to outputs, respectively. Regarding the skip connection block, 128 convolution filters are also used to match the activations of the fourth block; thus, they can be added together. For units' activations inside the blocks, ReLU non-linear activation function is used. After convolution and before ReLU activation, a batch normalization operation is performed to keep the learning process easy, fast, and robust. At the end of each block, MaxPooling of pool size of 2 is performed to regularize the feature extraction operation performed by the convolutional filters. After all those blocks, a fully connected layer with 512 units and ReLU activation is added before predicting the final output using a Softmax layer.

Fine Grid Module

As stated above, from this mother ResNet, two modules were inherited. Those two child modules are almost the same except for the output layer, where one has three output units, which we called `coarse_grid_spatial_localizer` for coarse grid modeling, and the other has 22 output units we called `fine_grid_spatial_localizer` for regions fine grid modeling. Three instances from the `fine_grid_spatial_localizer` module were compiled and assigned to separate regions. Each was trained on the EEG spatial dataset generated from their corresponding region source points. As mentioned earlier, within this thesis, for all Deep Learning models, the Categorical Cross-Entropy function was used to penalize the prediction of the instances, and the Adam algorithm was used to pursue the optimum value of the network parameters. After compilation, as summarized in Table 7.3, each `fine_grid_spatial_localizer` instance has a total number of 1,333,366 parameters. 1,332,118 of them are trainable, and the rest 1,248 are

not. To fit the model instances, they were trained and validated on their corresponding regions' spatial training and validation datasets. The fitting operation was set to an early stop if training accuracy would not improve for 10 epochs. Table 7.3 sketches the learning performance curves of the three regions' models. The trained instances' predictions were tested on a completely separate testing dataset from the validation dataset. Region1, Region2 and Region3 instances' test accuracy were about 94.5%, 94.7%, and 99.9%, respectively. If early stopping was set based on the validation accuracy baseline, better performance could be achieved. It is observable that Region3 spatial data was straightforward to map. Even from early epochs, validation accuracy was approximately stable and above 95%. That could be due to many reasons; for example, 1) Region3 is placed apart from the other regions, so there will be no correlation between EEG signals of adjacent source points located in different regions, and 2) region3 was implemented in a different shape that is more parallel to the head surface and positioned closer to the surface. Thus, EEG electrodes' sensing of their signals would be clearer.

Coarse Grid Module

Regarding the coarse grid modeling, as mentioned above, the same mother ResNet module was used; despite that, the output layer only consists of a 3-elements vector. The module instance was compiled for the same Adam algorithm and Categorical Cross-Entropy. The instance had a slightly fewer number of parameters than the fine ones. As summarized in Table 7.3, the total number of parameters is 1,323,619. The trainable parameters are 1,322,371, and the rest of 1,248 are non-trainable. The module instance

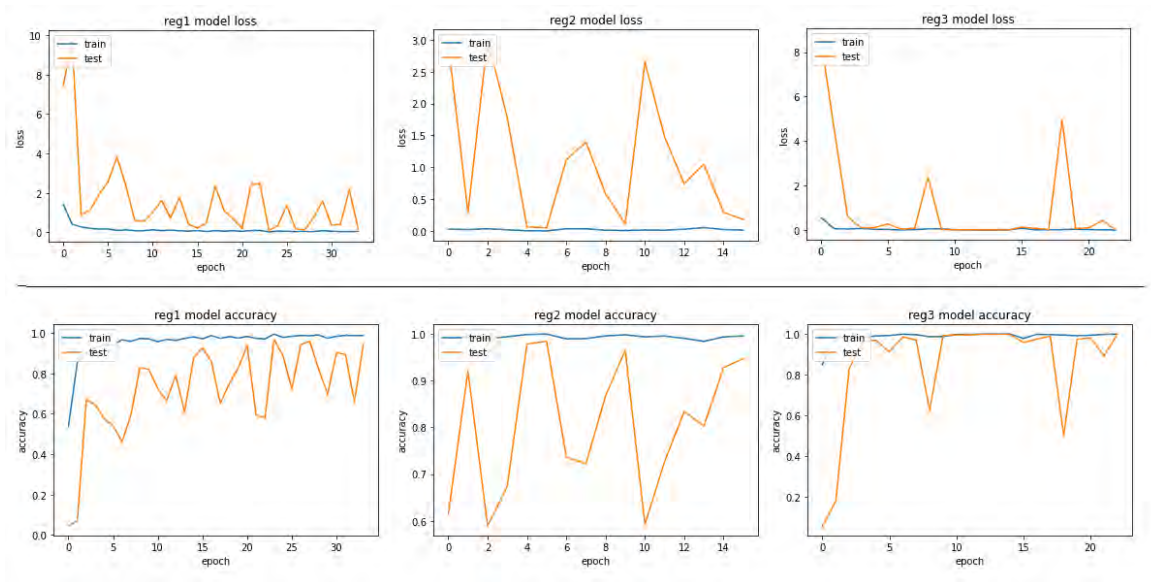
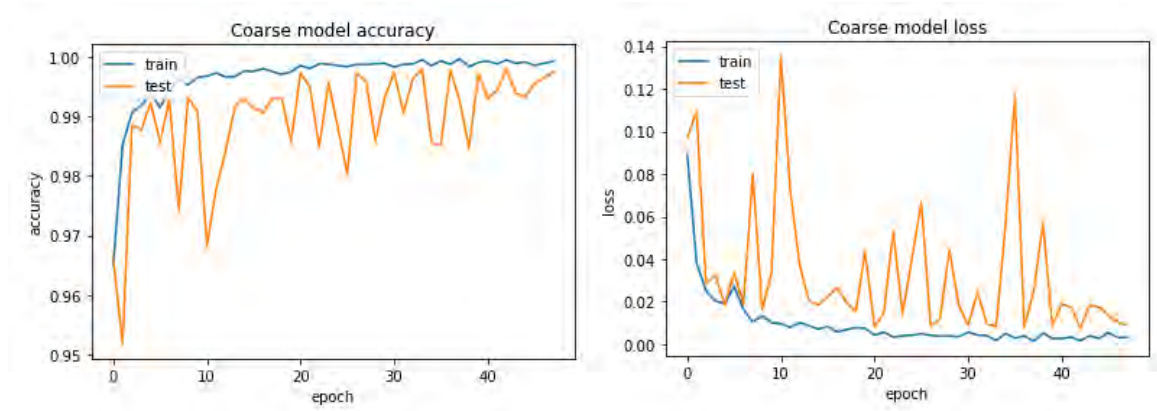


Figure 7.14: Training performance curves of individual regions ResNets

was trained on the same EEG data of the regions; however, the target outputs are only three labels corresponding to the source region. As well as before, the module was set to train until there was no improvement in training accuracy for 10 epochs. Fig. 7.15 shows the module's instant performance during training. It is observable that fitting the coarse grid was easy and fast; however, that is not to be expected when the whole-brain volume is considered and divided into regions. Finally, the instant was tested on a separate testing dataset, and its prediction accuracy was about 99.7%. The confusion matrix in Fig. 7.15 interprets the mutual interference between Region1 and Region2 predictions and ultimately no interference with the far placed Region3.

Table 7.3: Residual networks models summary

ResNet module	fine_grid_ spatial_ localizer	coarse_grid_ spatial_ localizer
#units in output layer	22	3
#trainable parameters	1,332,118	1,322,371
#non-trainable parameters	1,248	1,248
Total #parameters	1,333,366	1,323,619



(a)

```
[[2187  13   0]
 [  9 2191   0]
 [  0   0 2200]]
```

(b)

Figure 7.15: (a) Training performance curves of coarse grid ResNet, and (b) Confusion matrix

Overall Module

To connect all the trained modules, a relay function was implemented. This governing function takes the scalp EEG signals as an input. Then use the `coarse_grid_spatial_localizer` instant to predict the region that is expected to be active. Then predicts the source within that region corresponding to the same scalp EEG signals using the `fine_grid_spatial_localizer` module instant of the region that was predicted to be active. Thus, the function has two output vectors, one indicates the active regions, and the other determines the expected active source points within that region. There is no need to train this function; however, its prediction was tested

on the test datasets of the three regions. The prediction accuracy of the overall system was about 96%. Table 7.5 summarizes the performance of both the residual model and the upcoming sequence model.

7.3.4 Simplified Sequence Models

First, some performance measures should be taken care of. As for all Deep Learning models, LSTMs' performance depends on some hyperparameters. Thus, few initial experiments were executed to hyper-tune those hyperparameters.

The first and usually the only important hyperparameter of LSTM models is the number of hidden states. In Keras framework, the argument “units” of LSTM class is usually misunderstood as the number of outputs ($\hat{y}$) of the whole LSTM cell. However, it means the number of hidden states inside the cell. Fig. 7.16 illustrates one of the basic structures of sequence models, many-to-one structure, using LSTM cells, which we will use to address the Spatio-temporal EEG inverse problem. Upon this number of hidden state units depends the fitting performance of LSTMs and the number of parameters within the cell, therefore, the size of the whole model. We have tested three LSTM models: one contains 128 units, another has 256 units, and the last includes 512 units. We concluded that the best choice is 128 hidden state units; since they fit the problem without overfitting, unlike 512 units, and are fast to train. The 256 units model achieves high accuracy much earlier or faster than the 128 units one, but they consume much more memory. Although training time is an important variable, we will trade it for memory since the same fitting performance is achieved after all.

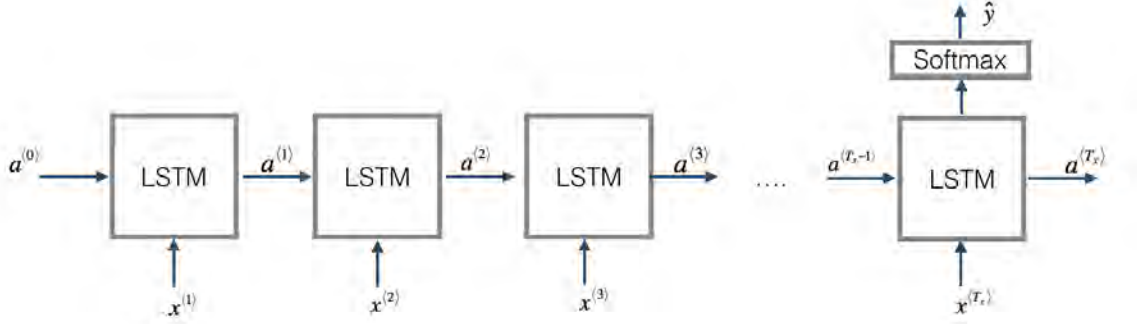


Figure 7.16: Many-to-one Sequence model structure with LSTM cells

Fine Grid Module

So, the chosen LSTM model for the final fine grid modeling has 128 hidden state units. The inputs are the EEG signals of shape (345,125); 345 is the number of EEG electrodes, and 125 is the number of samples per half-second within one EEG channel. As mentioned in Chapter 5: Data Generation, we settled with half a second just to reduce the size of the dataset; for sure, learning time will also decrease accordingly, which is also good. The LSTM outputs, the 128 hidden states, are fully connected to 22 units corresponding to each region's unified number of source points. A Softmax layer was used to activate the final prediction, $\hat{y}$. The total number of parameters of the whole `fine_grid_spatio_temporal_localizer` is 132,886, with no Non-trainable parameters, as summarized in Table 7.4.

As stated previously, to verify our approach, three regions were simulated. Therefore, three instances of `fine_grid_spatio_temporal_localizer` were compiled, an instance for each region. Then, all three were trained on their corresponding region Spatio-temporal generated data. Due to the inconsistency in sequence model learning, the training process was set to stop whenever validation accuracy crossed the 98% baseline.

After all, training did take a long time even though a High-Performance Computer was utilized. Usually, the training accuracy is the metric used for early stopping; however, we have adopted validation accuracy since we have recognized from pre-implementation that the model is almost perfectly fitting. i.e., training and validation performance curves are very close, and there is no overfitting. Fig. 7.17 below shows the performance curves of the three models' instants during training. The trained models' accuracy was then measured on separate testing datasets that were not introduced to the models before.

As summarized in Table 7.5, the test accuracy of Region1, Region2 and Region3 models' instants were 98.4%, 98.1%, and 93.2%, respectively. Fig. 7.17 shows only the training performance curves of each region model's instant. The accuracy of the region3 model's instant is much lower than other regions' models. Region3 instant was trained for many sessions, though. Each one of them is 200 epochs. The curves shown in Fig. 7.17 are for the last session, after which we decided to stop the training. It is noticeable that was not the case with the spatial model (ResNet). This issue could be due to many reasons: 1) The asymmetrical shape of Region3; if that is the case, then the regions' shape and dimensions could be considered a hyperparameter that needs to be optimized. Alternatively, 2) The EEG signals generated by these region sources are very correlated. This region is closer and more parallel to the surface, though. After all, the `fine_grid_spatio_temporal_localizer` model's instant was not as good as the other regions' instants in filtering the appropriate distinguishing features. However, all of that may be solved by fitting the data more by increasing the number of hidden states to 256 units.

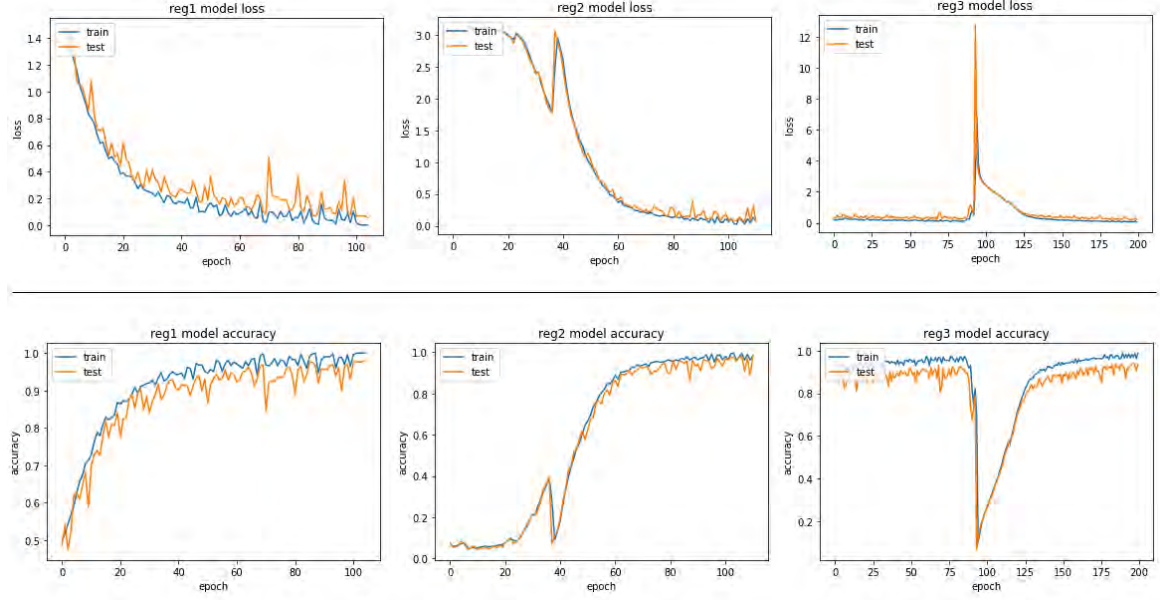


Figure 7.17: Training performance curves of individual regions LSTMs

Comparing regions' models, regarding the spatial and the Spatio-temporal modules, all instants performance is considered good, however, swapped in excellence. Moreover, as we mentioned earlier, the separated region fitting using spatial modules was better than adjacent regions, in contrast to Spatio-temporal modules fitting, which is better for neighboring regions. This is strange and could be concluded as follows: different deep learning architectures may have different performances due to the difference in concept between spatial and Spatio-temporal simulated data.

Coarse Grid Module

Same as the fine grid module, the `coarse_grid_spatio_temporal_localizer` module consists of an LSTM cell, dense layer, and Softmax activation layer. Since we had implemented only three regions, the number of points in the coarse grid is only three. Thus, the dataset generated for the coarse grid has only three elements in the target

labels vector. Therefore, to map the coarse grid EEG inputs with the 3-elements output, not many hidden state units are needed inside the LSTM. So, the LSTM cell was chosen to have 64 hidden states, and of course, the dense layer parameters will be much less since we have only three outputs to predict. Although, when the coarse dataset was generated, we had to consider the case of correlation between adjacent source points, however, in two different regions, as illustrated in Chapter 5: Data Generation. This correlation to be sufficiently mapped may require a large number of parameters (hidden states, dense and activation layers' parameters). So, an instant of the `coarse_grid_spatio_temporal_localizer` model was compiled. Categorical Cross-Entropy and Adam were used as loss function and optimization algorithms, respectively. The total number of parameters of this instant is 48,835 parameters, and all are trainable. Then, it was trained on the same Spatio-temporal EEG signals data that `fine_grid_spatio_temporal_localizer` instants were trained on; however, the labels are only 3-elements vector corresponding to the regions. The training terminated whenever validation accuracy crossed the 98% baseline. Fig. 7.18 (a) shows the training performance curves. Fitting was easy. This is not to be expected when considering the whole-brain regions, though. Fig. 7.18 (b) shows the confusion matrix report. Despite the high accuracy, It is clear that there is misprediction between the adjacent regions, while nearly exact prediction for far ones' data. The prediction accuracy of the coarse grid module on the testing data is about 98.2%, as summarized in Table 7.5.

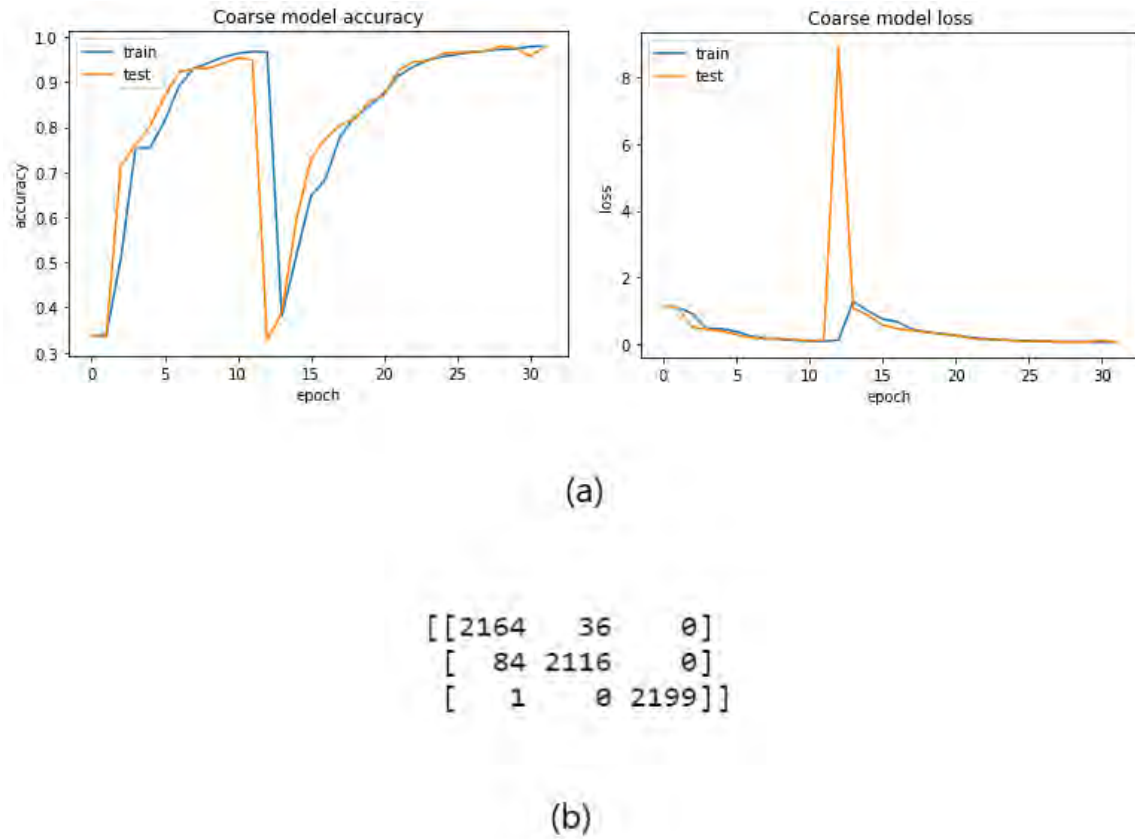


Figure 7.18: (a) Training performance curves of coarse grid LSTM, and (b) Confusion matrix

Table 7.4: Sequence models summary

LSTM module	fine_grid_spatio_temporal_localizer	coarse_grid_spatio_temporal_localizer
#units in output layer	22	3
#trainable parameters	132,886	48,835
#non-trainable parameters	0	0
Total #parameters	132,886	48,835

Overall Module

To connect the two stages, the function, `Spatio_Temporal_Localizer`, was defined to relay the desired modules. It detects the prediction of `coarse_grid_spatio_temporal_localizer`, then accordingly, activates the desired

fine_grid_spatio_temporal_localizer instant which represents that region. This function inputs only the measured Spatio-temporal EEG signals on the scalp (the same in all Spatio-temporal datasets). It returns an output of two tensors: One determines the active region, and the other predicts the location inside that region. This function does not need training. Testing performance of this function represents the performance of the whole localization system. The Spatio_Temporal_Localizer performance was measured on the same regions' test datasets for this simplified implementation. The overall performance of the function, thus the whole localization system, was 95%. Table 7.5 summarizes all Deep learning models, Residual and sequence models, and accuracies.

Table 7.5: Deep Learning models accuracy summary

Model	Residual model	Sequence model
Region1 accuracy	94.5%	98.4%
Region2 accuracy	94.7%	98.1%
Region3 accuracy	99.9%	93.2%
Coarse grid accuracy	99.7%	98.2%
Overall accuracy	96.0%	95.0%

7.3.5 Complete Whole-Brain Structure

After experiencing our stage-wise approach to the simplified problem space and observing the high accuracy results the approach has provided, we can conclude that our stage-wise approach is a promising solution to the brain source localization problem. As the overall accuracy of the stage-wise models on simple problem setup is greater than 95%, they could be considered excellent models to solve the complete problem.

Therefore, we can go ahead with the expansion of the model to the whole-brain problem space. The stage-wise approach will be developed for the whole-brain volume. As mentioned in the approach chapter, the whole-brain volume is divided into regions. Thus, the whole-brain source space was divided into 56 regions. Each region contains as max as 64 source points. The on-surface regions could have source points as low as a single source point.

In a simplified implementation, the whole-brain source space grid is now divided into coarse and fine grids. The whole-brain coarse grid is a virtual grid that has 56 grid points that represent the regions. And the fine grid is the actual source space grid, however, divided and clustered inside the whole-brain coarse grid. So, the fine grids have the exact resolution (source/cm³) as the source model resolution. The same Residual Networks models of Fig. 7.13, which were used to map the simplified three regions problem, are also to be used to map the complete whole-brain problem. The same ResNet is used to module the whole-brain coarse and fine grids, however, with a different number of units in the output layer depending on the number of the corresponding grid points.

Whole-Brain Fine Grid Module

For the fine grids, modules will be developed to map the scalp EEG signals with their corresponding sources inside the regions. A ResNet is assigned for each of the 56 regions we have created virtually. Each ResNet is like the mother ResNet in Fig. 7.13; however, the number of units in the output layer is 64 units. Each region contains a different number of source points, though.

This module, which we called “Full Brain fine grid spatial localizer,” has a total number of 1,354,912 parameters. 1,353,664 of them are trainable parameters, and the rest 1,248 are non-trainable ones, as summarized in Table 7.6.

This structure is fixed to all regions, and 56 instances of this module were assigned, compiled, and trained for each region out of all 56 regions. The training performance curves of all these FullBrain_fine_grid_spatial_localizer instances are available in Appendix A. Some training performance figures do not contain curves, like in Fig. A.7, which should not be mistaken as there is an error in this region, on the contrary, that means this region module instant fitted the data to more than 99% in only one training epoch. Due to the high number of figures and the many pages they consume, we prefer to present them at the end of the thesis.

Unlike the simplified three regions implementation, the complete full brain implementation module instances training was not that easy due to the complexity of the increased number of outputs to be predicted. Some whole-brain fine grid spatial localizer instances took a very long time to reach validation accuracy of 99%, while others were swift learners. Nevertheless, some regions’ instances could not reach 99% after 100 training epochs, like regions 17, 22, and 39, whose validation accuracy is 96.4%, 96.1%, and 95.6%, respectively, as summarized in Table 7.7. Special care should be paid into region 22, whose validation accuracy was just 92.5%. Those regions’ instances may need special treatment to increase their prediction accuracy, like increasing training epochs more or providing more training data.

Whole-Brain Coarse Grid Module

As well as the whole-brain fine grid module, the whole-brain coarse grid module uses the same residual network represented in Fig. 7.13. The only difference is the number of units in the output layer. Here, the whole-brain coarse grid module has 56 labels in the output layer corresponding to the total number of regions that virtually interpret the whole-brain source space. As listed in Table 7.6, the coarse module has 1,349,560 trainable parameters and 1,248 non-trainable ones. Thus, the module contains a total of 1,350,808 parameters.

Unlike the fine grid, the coarse grid only requires one instant of the module. The instant was compiled, then trained on the surface EEG data labeled to the active source region. Fig. 7.19 shows the learning performance curves of the instant. The trained coarse grid instant validation accuracy is 98.9%, as listed in Table 7.7.

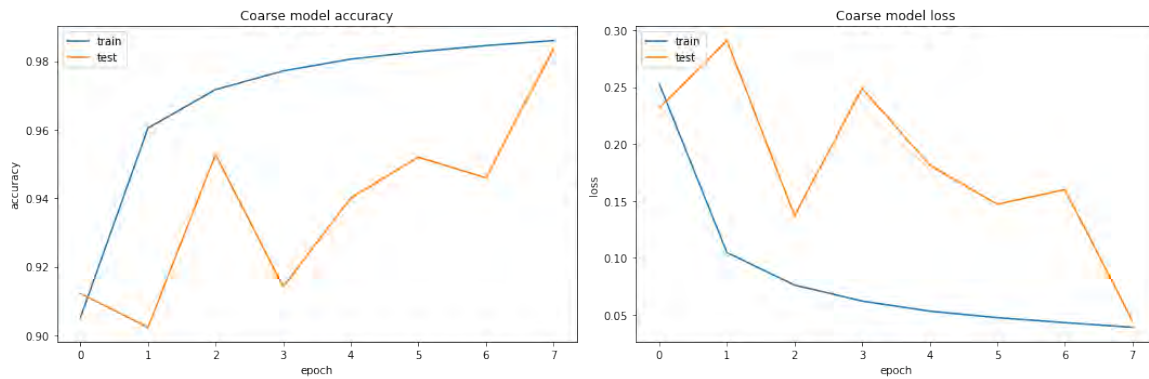


Figure 7.19: Training performance curves of full brain coarse grid ResNet

Overall Whole-Brain Module

Finally, putting all pieces together, an overall controlling module must be defined. This overall full brain module connects the captured EEG signals on the scalp to

its corresponding neuron source inside the brain via the whole-brain coarse and fine grid modules. The module takes the scalp EEG signals as an input, then uses the prediction of the whole-brain coarse grid module instant to highlight the expected region to contain the active source. Then, the exact active source inside that region is estimated via activating the whole-brain fine grid module instant of the highlighted region. This overall module does not need to be trained since it is just a switching, non-trainable function.

The performance of the module, hence the performance of the whole proposed stage-wise approach, was tested on a separate testing dataset. The prediction accuracy of the system was found to be 92.7%, as summarized in Table 7.7.

Table 7.6: Full Brain Residual networks models summary

Full Brain ResNet module	FullBrain_ fine_grid_ spatial_ localizer	FullBrain_ coarse_grid_ spatial_ localizer
#units in output layer	64	56
#trainable parameters	1,353,664	1,349,560
#non-trainable parameters	1,248	1,248
Total #parameters	1,354,912	1,350,808

Table 7.7: Full Brain Deep Learning models accuracy summary

Region	ResNet model accuracy	Region	ResNet model accuracy
Region1	99.3%	Region2	99.0%
Region3	99.6%	Region4	98.0%
Region5	98.0%	Region6	98.9%
Region7	100.0%	Region8	99.4%
Region9	99.0%	Region10	99.0%
Region11	98.5%	Region12	99.7%
Region13	99.2%	Region14	100.0%
Region15	99.0%	Region16	97.7%
Region17	96.4%	Region18	99.6%
Region19	99.8%	Region20	97.5%
Region21	92.5%	Region22	96.1%
Region23	99.2%	Region24	100.0%
Region25	98.9%	Region26	99.5%
Region27	99.5%	Region28	99.6%
Region29	99.3%	Region30	99.3%
Region31	99.9%	Region32	99.9%
Region33	99.1%	Region34	98.1%
Region35	99.4%	Region36	100.0%
Region37	99.5%	Region38	98.0%
Region39	95.6%	Region40	98.6%
Region41	99.3%	Region42	100.0%
Region43	99.9%	Region44	99.4%
Region45	99.1%	Region46	100.0%
Region47	100.0%	Region48	99.9%
Region49	99.7%	Region50	99.2%
Region51	99.9%	Region52	99.7%
Region53	99.7%	Region54	100.0%
Region55	100.0%	Region56	100.0%
Coarse grid ac- curacy	98.9%	Overall accuracy	92.7%

CHAPTER 8

CONCLUSION AND FUTURE WORK

8.1 Conclusion

Compressed sensing reconstruction techniques are, in general, successful in localizing brain activity in one-dimensional space. These algorithms could reconstruct the active source perfectly wherever it was, in the brain's cortex or deep inside the brain for single-source activity. And for multi-source activity, their performance was quite well on reconstructing cortex sources. Unfortunately, compressed sensing algorithms' performance was not good; due to the high correlation between scalp EEG signals generated from close neurons. This correlation increases as sources go deep inside the brain; thus, iterative techniques can no longer differentiate between the propagated electromagnetic signals.

On the other hand, deep learning proves its success in solving ill-posed inverse problems. Deep Learning architectures were capable of localizing brain source ac-

tivity with a high degree of accuracy. Deep learning architectures were applied to single-active source problem. Residual Networks were used for the spatial data, while Recurrent Neural Networks (RNNs) were utilized for Spatio-temporal data. A multi-stage approach was proposed to simplify the problem by staging the source space and reducing the number of computations needed. Two grids were formulated: coarse grid and fine grid. The proposed approach is to solve the inverse problem to find the active regions of the coarse grid corresponding to the captured EEG signals. Then, within those active regions, the inverse problem of the same EEG signals is to be solved again to identify the exact active fine grid sources.

The approach was investigated on a simplified setup of three regions, two adjacent to each other and the third on the other side of the brain. Whether ResNet or LSTM, each region had been assigned a model, and that model's parameters were trained on the data generated from that region. On top of those region models, another model is assigned to the coarse grid and trained on labels of that coarse grid. Deep learning models gave high prediction accuracy. For ResNets, the least region model test accuracy is 94.5%. The coarse grid ResNet test accuracy is 99.7%. Thus, the overall spatial model test accuracy is 96%. For LSTM models, the least subregion model test accuracy is about 93.2%, and the coarse grid LSTM test accuracy is about 98.2%. Thus, the overall Spatio-temporal model test accuracy is 95%.

Due to this astonishing accuracy in the simple testing problem, Residual Networks were extended and applied to the real problem of whole-brain volume. The overall full brain spatial localization accuracy of EEG signals is 92.7%. Therefore, our conclusion is: Deep learning techniques are a promising solution to the brain source localization

problem. Thus, we recommend the application of deep learning models on multi-active sources scenarios.

8.2 Future Work

We recommend more investigation of the application of compressed sensing to the brain source localization problem by considering the problem as a multi-measurement problem of common sparsity structure. Already there is a research trend in that direction, and some papers have been published in that regard. Some enhancements and adjustments are being made to some well-performing compressed sensing algorithms to adapt to the multi-measurement vectors problem.

Regarding data-driven models, we recommend investigating the development of structures that work on EEG signals as 2-dimensional data. Within this thesis, we had simulated EEG data as a vector of channels, such that electrodes measurements are flattened into a one-dimensional vector. However, their placement on the head is instead a surface plane. Although the head surface is not an exact 2-D surface, it could be projected into a 2-D surface without losing information. For sure, adjacent electrode channels capture signals that have correlated signatures. Thus, when electrodes are flattened into a vector, the adjacent channels will be separated, making extracting features from these channels using, for example, convolutional filters much more difficult.

8.2.1 Block Sparsity

Often the sparse signals under study have added structure. Specifically, in our case study here, we are concerned about the natural block structure and slowly changing source activity (in which brain activity could be described as jointly sparse signals that could be represented systematically in multiple measurement vectors structure. An all-natural block structure could exist in the sparse signal wherever a few non-zero components arise in groups. For instance, the neural activity might be described in a block sparse structure [20]. It has been recognized that the block structure can be used for improved recovery. Moreover, it was attested that if the information of the block structure (expected positions where blocks could occur and the block dimensions) is readily achievable under specific conditions, it enables us to lower the variety of observations needed for recovery [81].

Various algorithms have been suggested considering the information of the block structure. The fundamental work in this regard was [82], which introduced the group-LASSO algorithm. Though it has restricted applicability compared to various other techniques, it makes some presumptions about the dictionary. Block-OMP [83], proposed by Eldar et al., is an expansion of the standard orthogonal matching pursuit algorithm (OMP) [84]. One more technique called the mixed ℓ_2/ℓ_1 -norm recovery algorithm suggested in [85] expanded the BP algorithm to take on block sparsity. Likewise, expansions of the CoSAMP technique [72] and IHT [86] were made use of to propose a technique named Block-CoSAMP [81], which has deductible recovery guarantees and also robustness properties. The LaMP technique introduced in [87]

utilized a Markov Random Fields model to catch the formation of the sparse signals.

Every one of the techniques mentioned previously comes from the group of greedy algorithms. Their attraction is that they are agnostic to support distribution, thus demonstrating robust performance. Nevertheless, a lot of them function exclusively in a noiseless environment. Despite numerous Bayesian methods (e.g., [88]) might be expanded to employ block structure, those are not as popular. A significant deviation is the cluster-sparse Bayesian learning algorithm (cluster-SBL) introduced in [89], which presumes that the block supports obey a multivariate Gaussian distribution. In addition, cluster-SBL considers the intra-block correlation to enhance the recovery performance.

Mainly, there are two possible schemes for block sparsity:

- Known-block partition: in this scheme, the size of blocks and their probable positions are known a priori; however, the precise position of the active block is unidentified.
- Unknown-block partition: this scheme has no details about the block positions or their sizes. There could be blocks of differing sizes taking place out of any particular pattern. The only information is that in this scheme is the likely optimum size that any block might have.

Instances of block sparse signals from both schemes are displayed in Fig. 8.1.

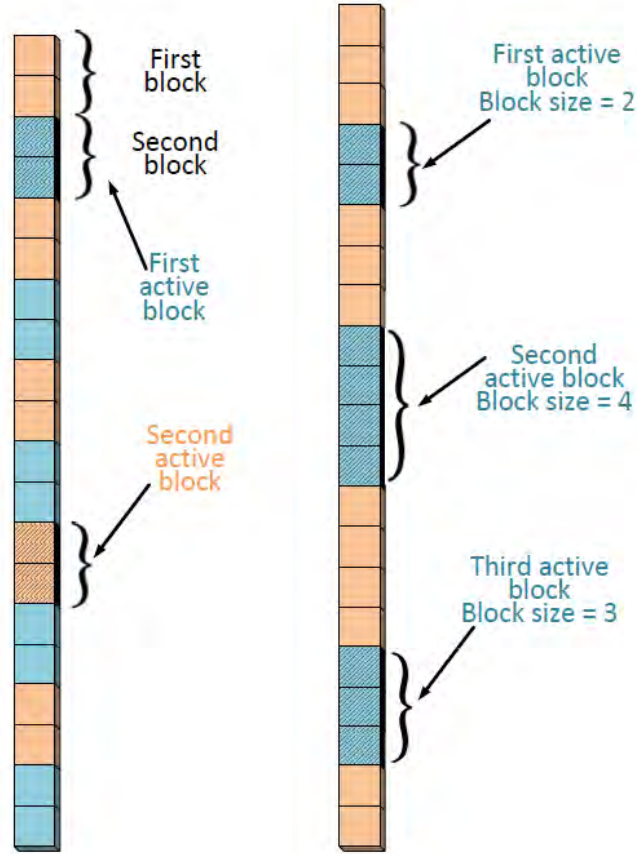


Figure 8.1: Demonstration of block-sparse signals. (left) Known-block partitions and (right) Unknown-block partitions block partitions [90]

Both sorts of block-sparsity occur in real-world problems and are equally important. However, in our problem, neural activity should be considered an unknown block partition; since neither the location of active regions is known nor the 3-D volume space of these simultaneous active neurons.

8.2.2 Joint Sparsity

A carefully associated problem of jointly recovering multiple sparse vectors from Multiple Measurement Vectors (MMV) has intensified attention. Fig. 8.2 shows a mathematical demonstration of the MMV structure.

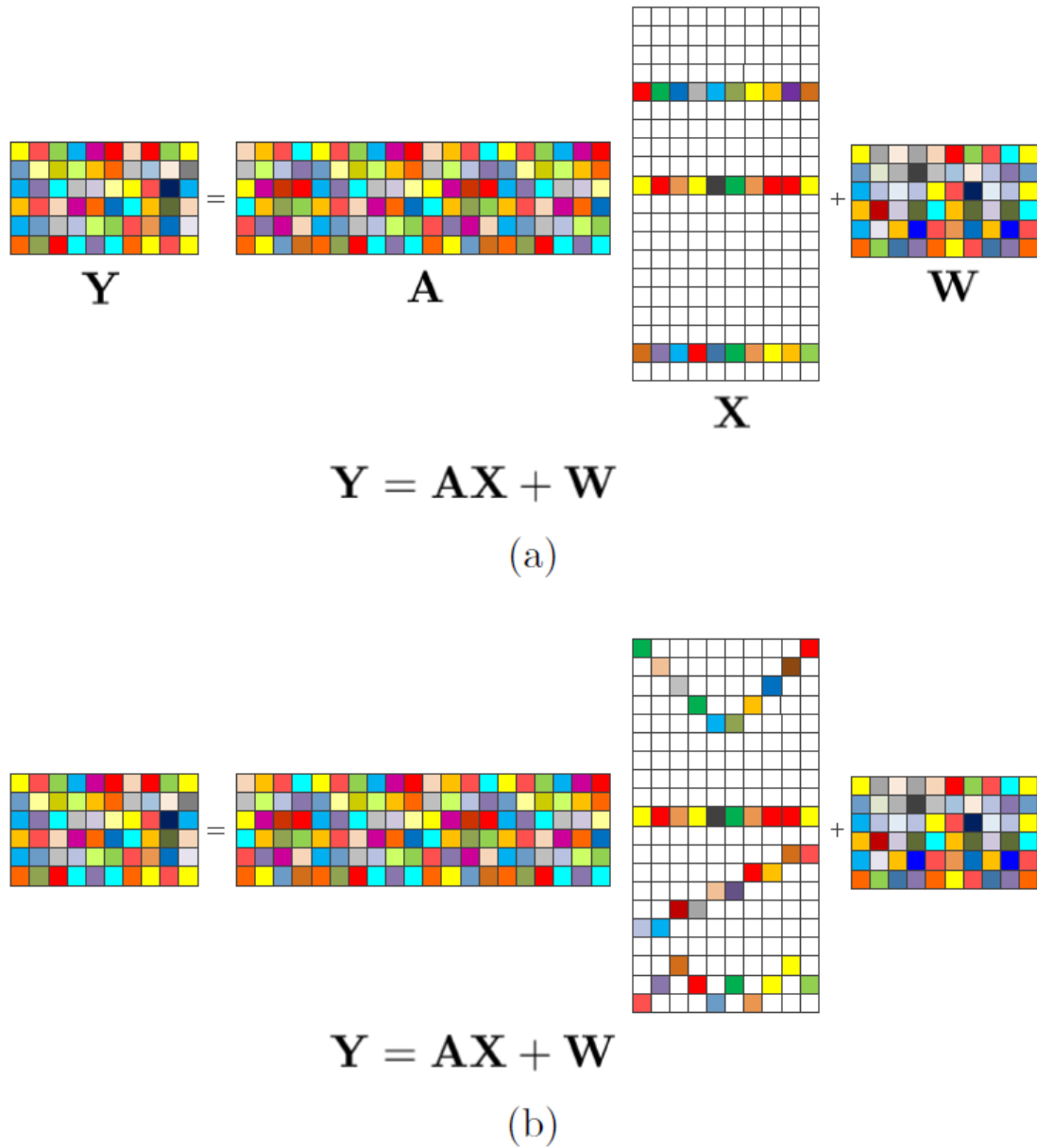


Figure 8.2: Matrices representation of Multiple Measurement Vectors [91]

A few of the applications where multiple measurements could be employed involve imaging of the brain utilizing Electroencephalography (EEG) and Magnetoencephalography (MEG) [20], [92], [93], Multivariate Regression [94] and Blind Source Separation [95]. Within these and many other analogous applications, the multiple sparse vectors might either share a common or a nearly common sparsity pattern. The joint recovery algorithms use these common sparsity patterns for the cost-effective restoration of sparse signals with considerable precision. Several joint recovery algorithms have been suggested. Most of the foundational and essential work in this regard is the expansion of already available Single Measurement Vector (SMV) techniques. For instance, MMV-FOCal Underdetermined System Solver (M-FOCUSS) [96], MMV-Orthogonal Matching Pursuit (M-OMP), Simultaneous Orthogonal Matching Pursuit (S-OMP) [97], multivariate group LASSO [94], and Multiple Sparse Bayesian Learning (M-SBL) [98]. A different technique named the ReMBo algorithm [99] obeys a different approach. It suggests modeling the MMV problem as an SMV problem. The technique linearly joins the measurement vectors and then determines the resulting SMV problem using a pre-set algorithm. One more course of algorithms exploits the characteristics of the unknown sparse signals, such as structure and correlation. For instance, Sparse Bayesian Learning (SBL)-based algorithms [100], [101], Orthogonal Subspace Matching Pursuit (OSMP), as well as Subspace-Augmented Multiple Signal Classification (SA-MUSIC) algorithms. Every one of the previously mentioned algorithms is centralized in nature which is the case for our neural activity signals. All measurement vectors and matrices are processed in a single processing unit. On the other hand, for situations such as sensor networks, where each node/sensor has its

very own measurement vector and matrix, the employment of centralized algorithms would need to exchange a massive volume of data to a centralized location. But, in ordinarily energy-constrained sensor networks, such centralized approaches are not feasible.

APPENDIX A

FULL BRAIN 56 REGIONS’ RESNET MODELS’ PERFORMANCE CURVES

This appendix contains the training performance curves of all FullBrain_ fine_grid_ spatial_ localizer 56 instants. Due to the high number of figures and the many pages they consume, we prefer to present them at the end of the thesis. Some training performance figures do not contain curves, like in Fig. A.7, which should not be mistaken as there is an error in this region. On the contrary, this region module instant fitted the data to more than 99% in only one training epoch.

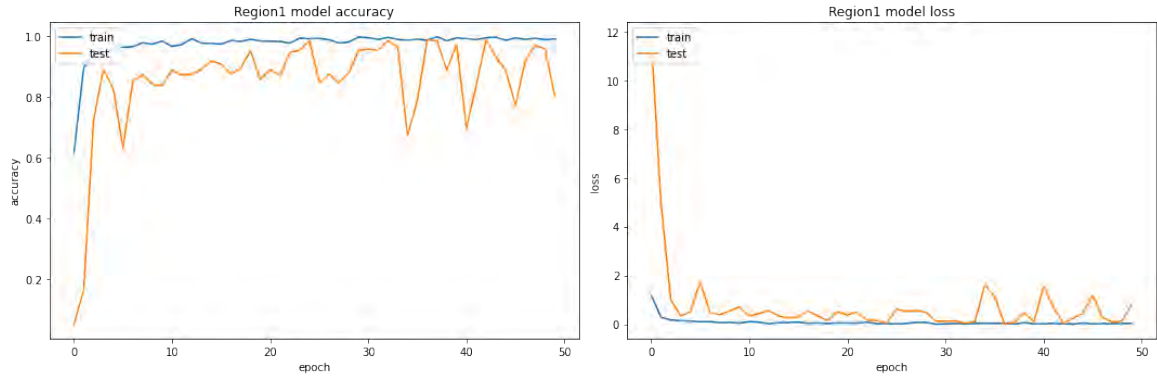


Figure A.1: ResNet model's Training performance curves of Region1

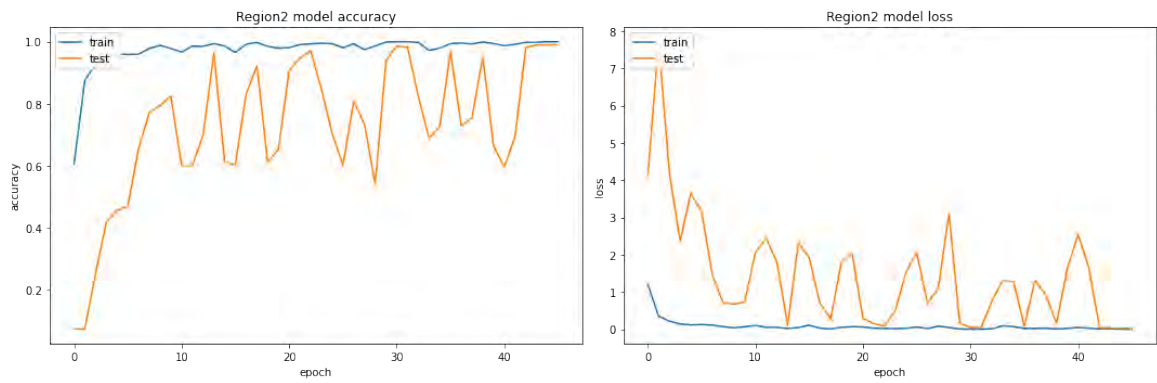


Figure A.2: ResNet model's Training performance curves of Region2

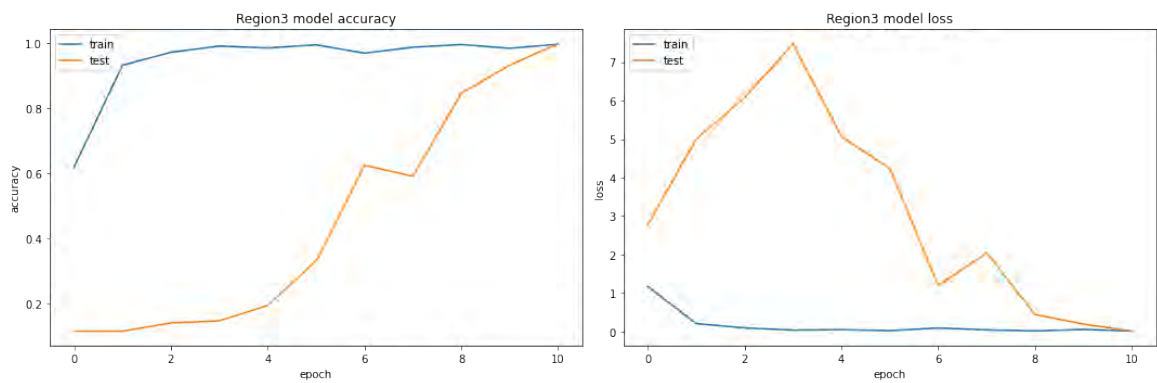


Figure A.3: ResNet model's Training performance curves of Region3

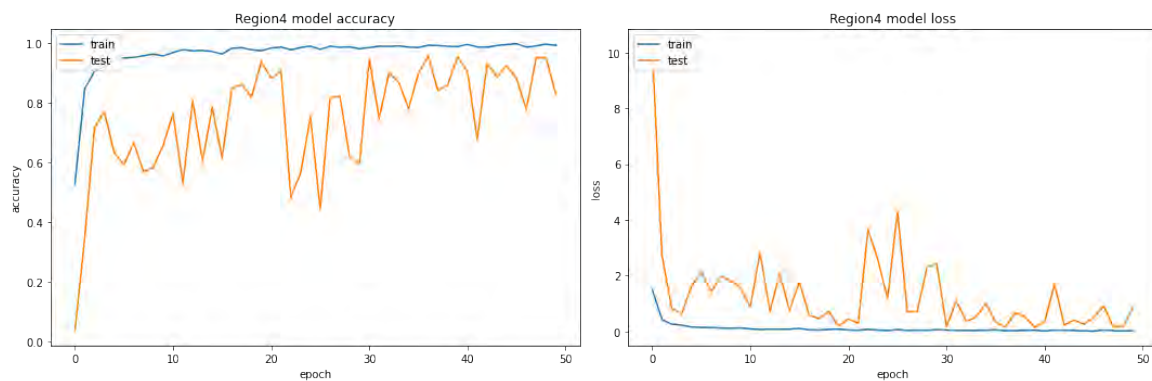


Figure A.4: ResNet model's Training performance curves of Region4

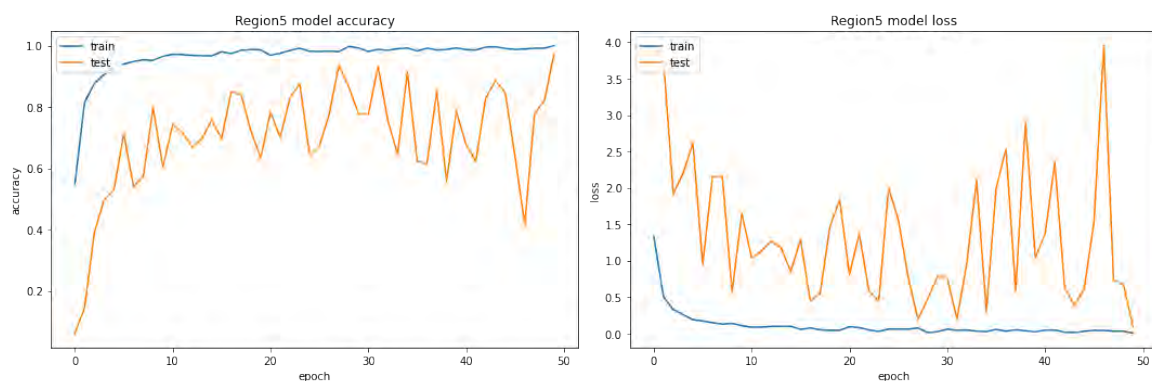


Figure A.5: ResNet model's Training performance curves of Region5

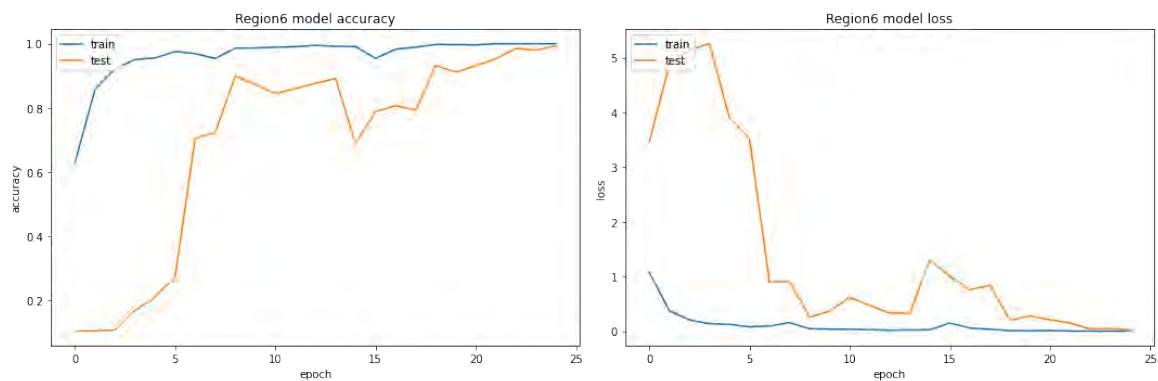


Figure A.6: ResNet model's Training performance curves of Region6

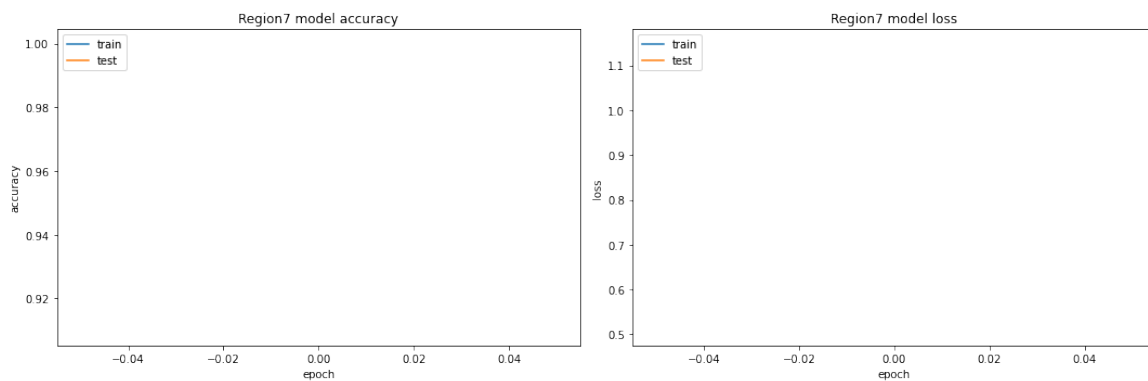


Figure A.7: ResNet model's Training performance curves of Region7

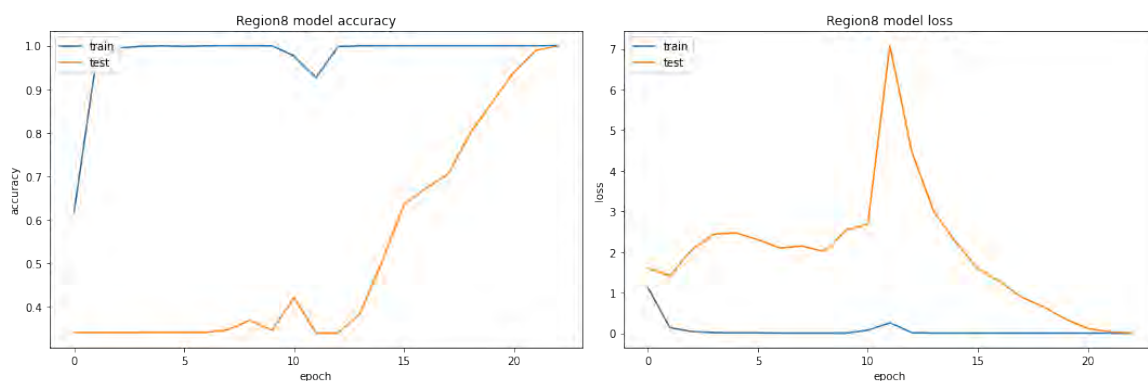


Figure A.8: ResNet model's Training performance curves of Region8

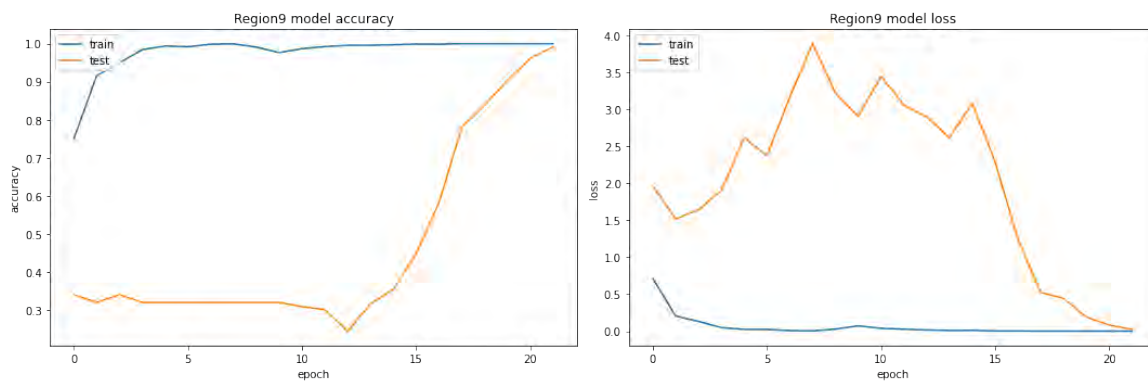


Figure A.9: ResNet model's Training performance curves of Region9

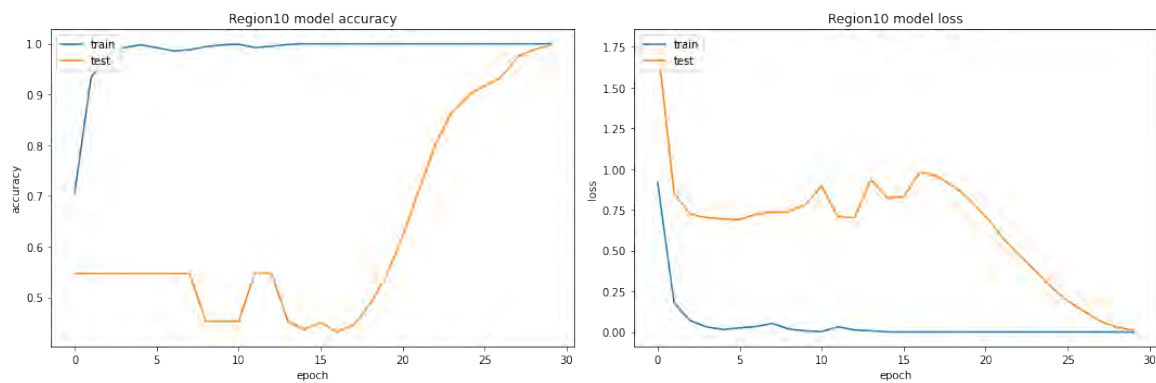


Figure A.10: ResNet model's Training performance curves of Region10

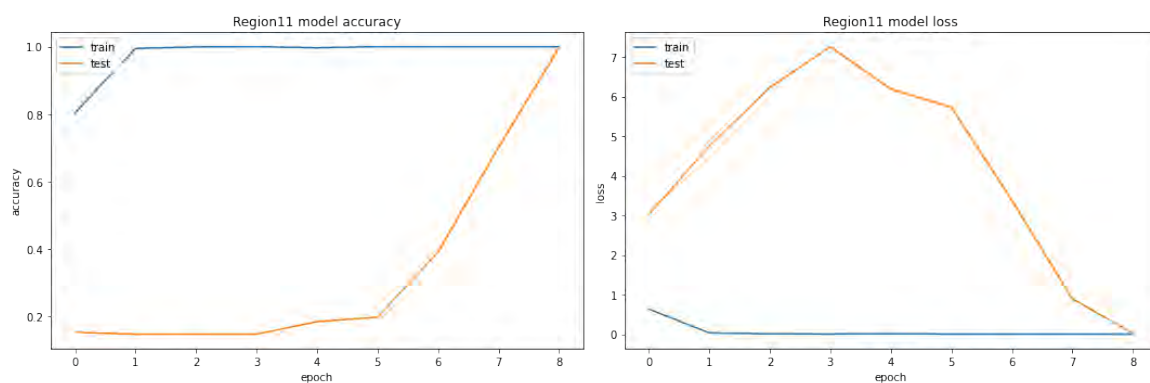


Figure A.11: ResNet model's Training performance curves of Region11

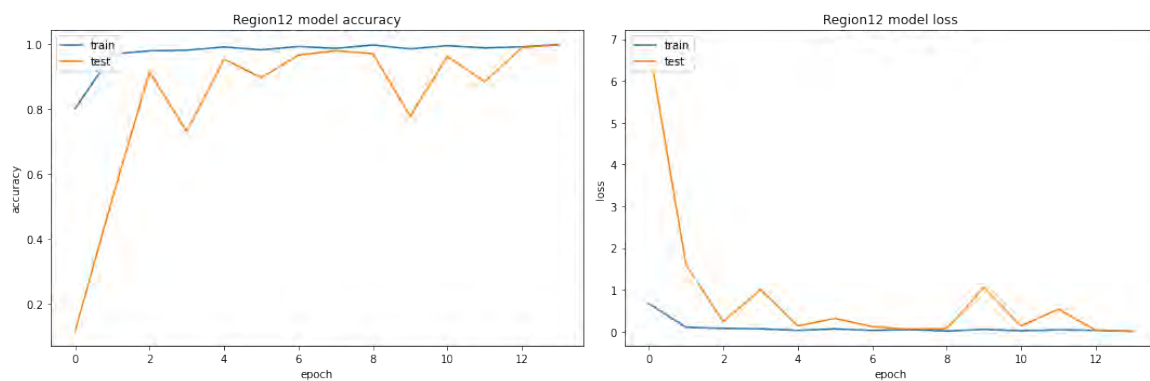


Figure A.12: ResNet model's Training performance curves of Region12

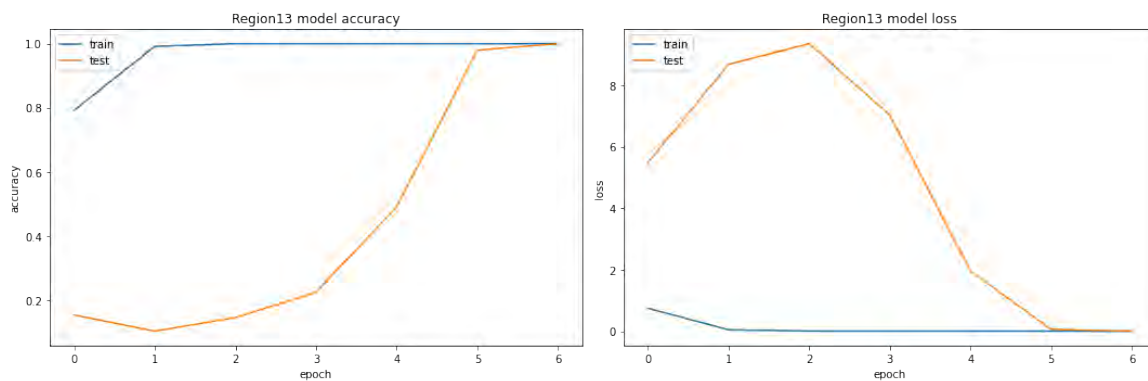


Figure A.13: ResNet model's Training performance curves of Region13

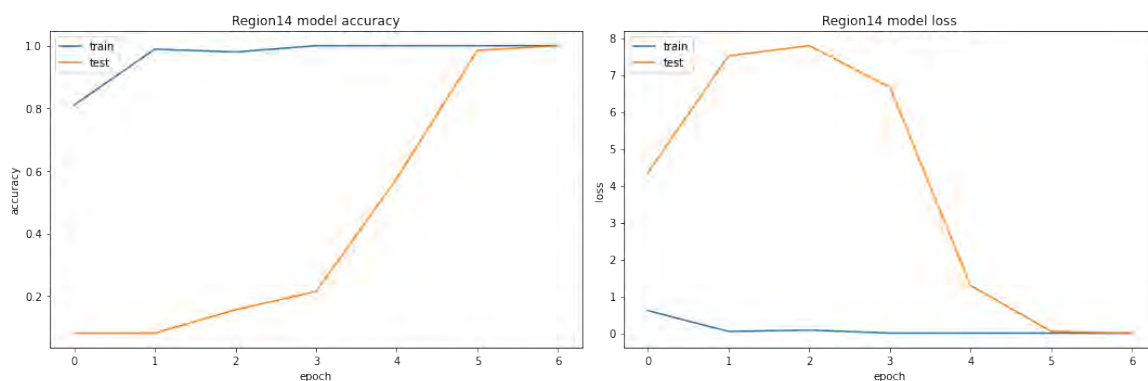


Figure A.14: ResNet model's Training performance curves of Region14

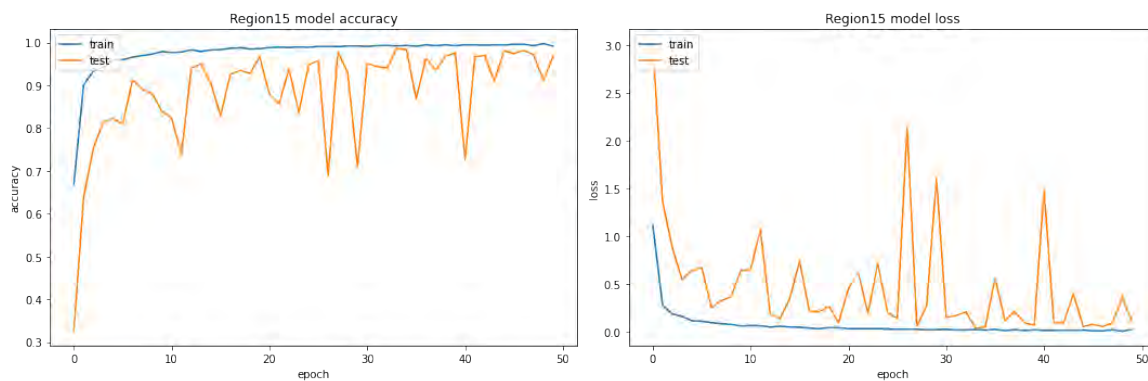


Figure A.15: ResNet model's Training performance curves of Region15

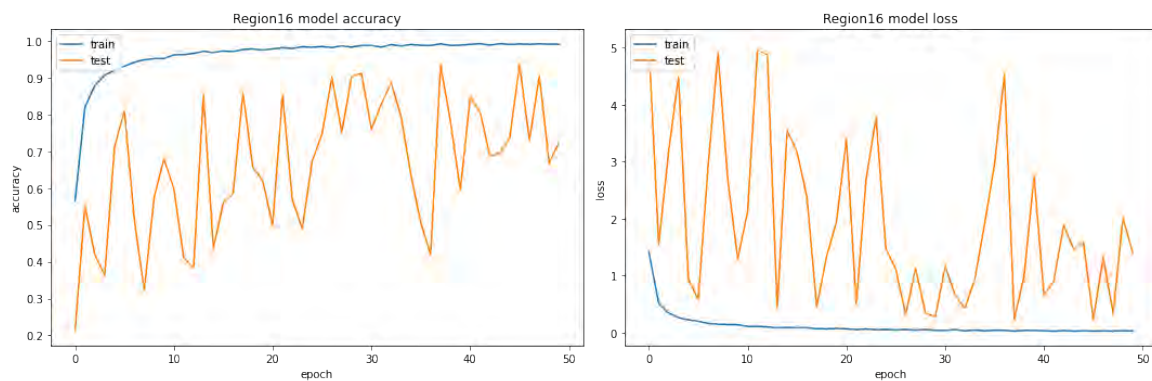


Figure A.16: ResNet model's Training performance curves of Region16

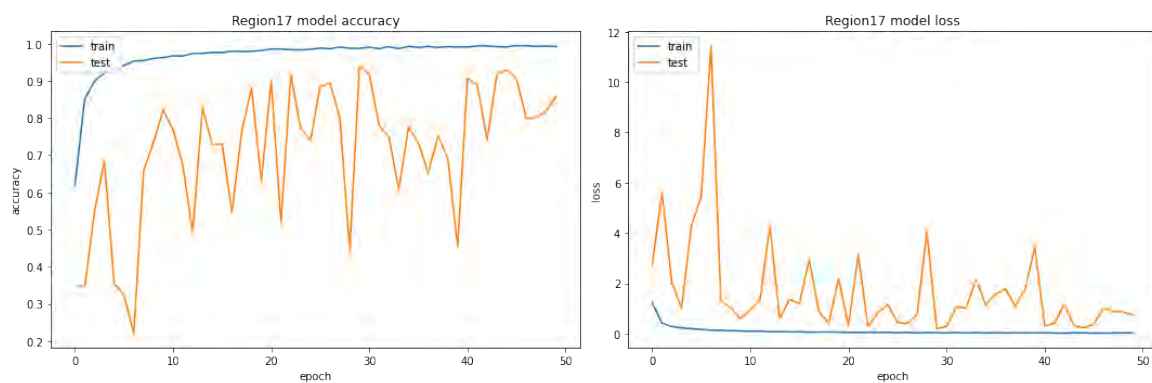


Figure A.17: ResNet model's Training performance curves of Region17

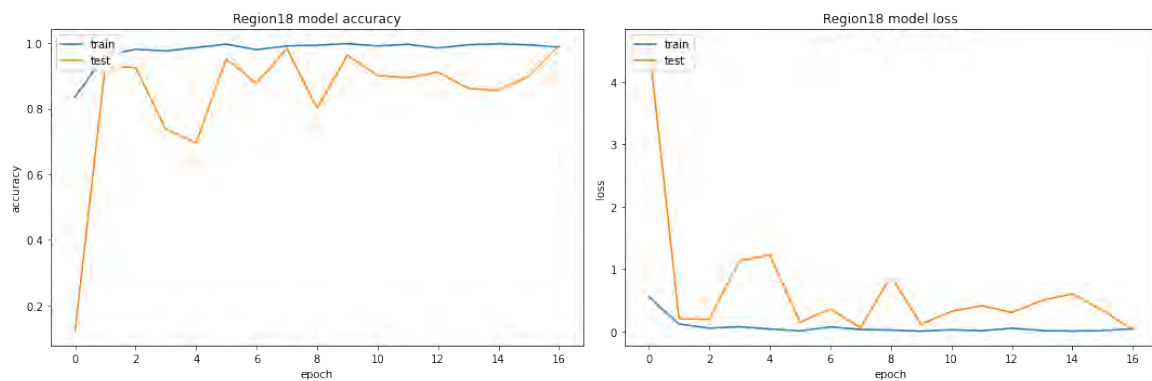


Figure A.18: ResNet model's Training performance curves of Region18

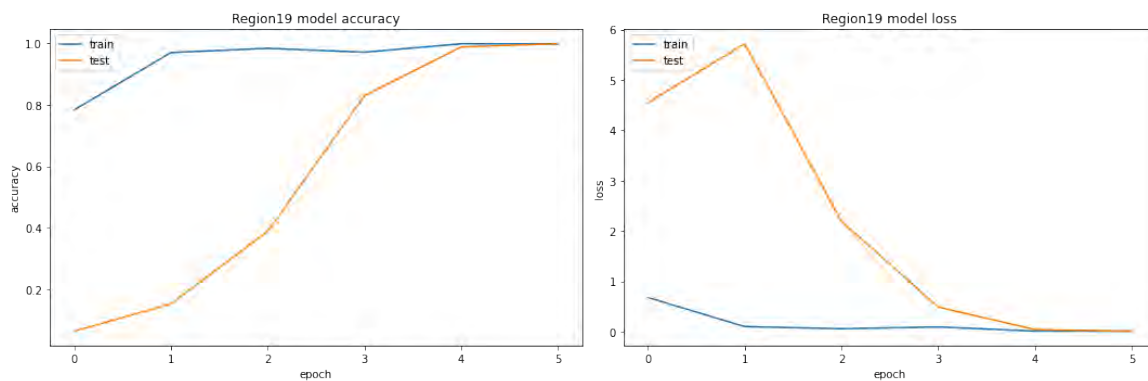


Figure A.19: ResNet model's Training performance curves of Region19

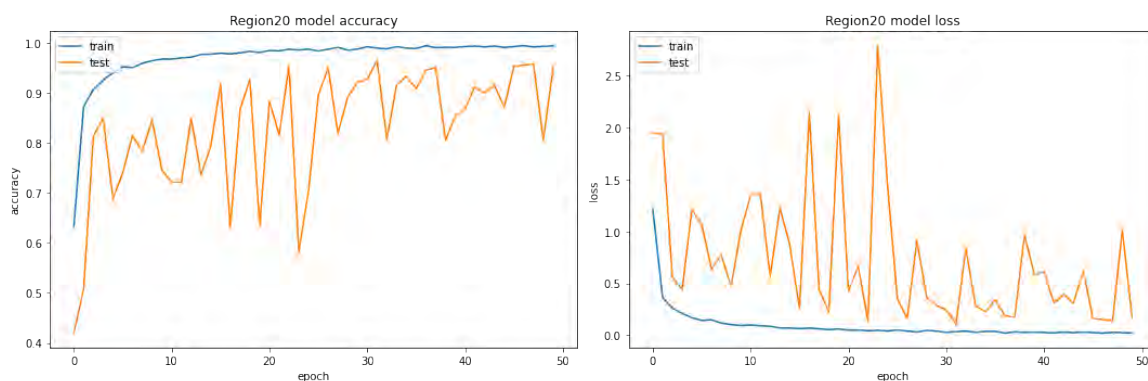


Figure A.20: ResNet model's Training performance curves of Region20

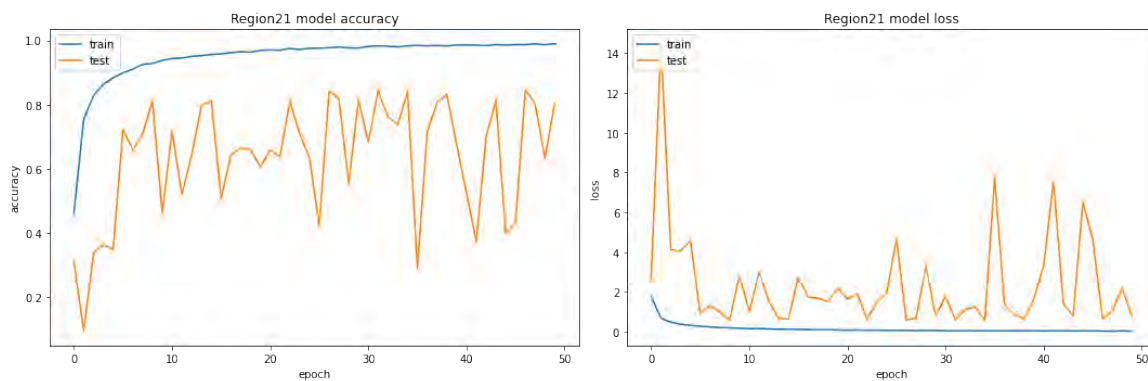


Figure A.21: ResNet model's Training performance curves of Region21

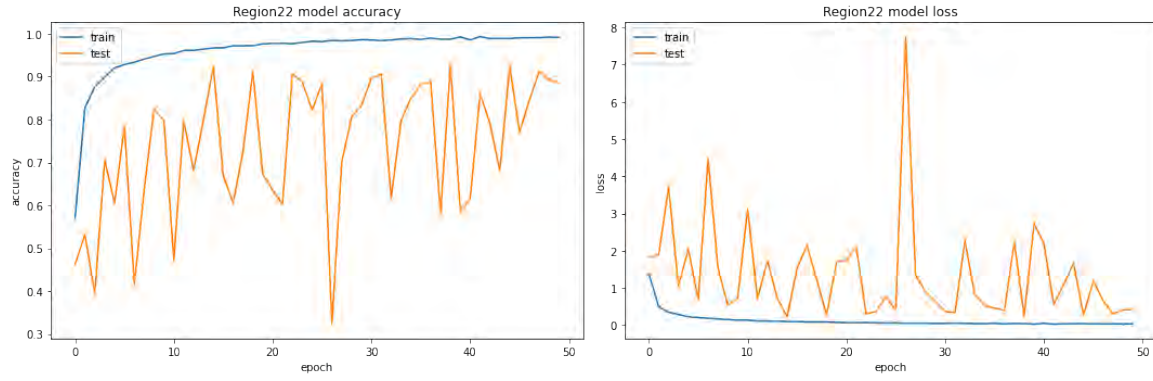


Figure A.22: ResNet model's Training performance curves of Region22

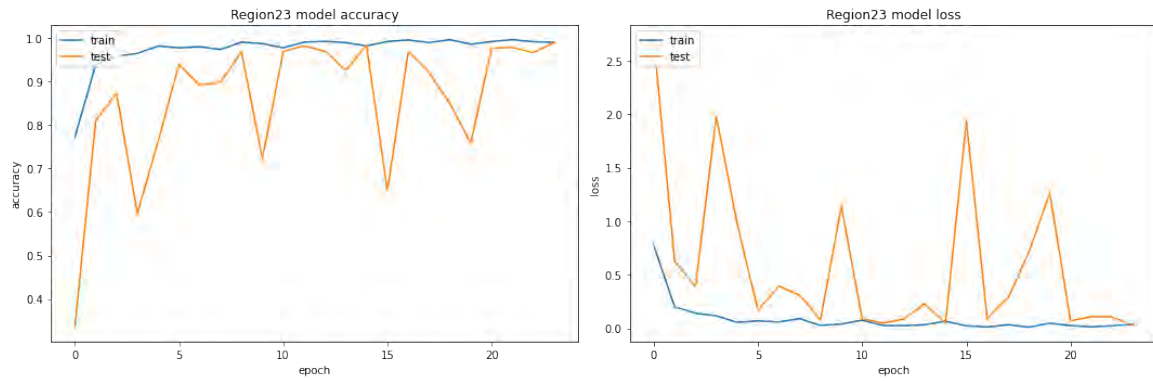


Figure A.23: ResNet model's Training performance curves of Region23

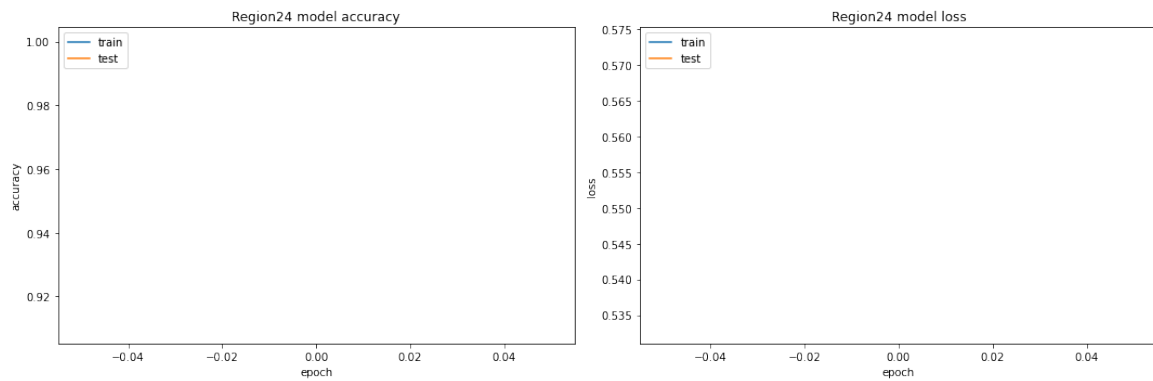


Figure A.24: ResNet model's Training performance curves of Region24

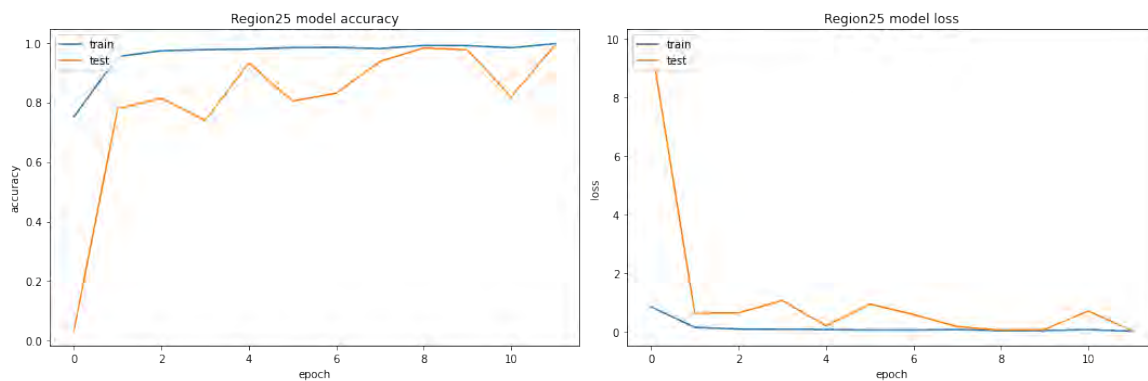


Figure A.25: ResNet model's Training performance curves of Region25

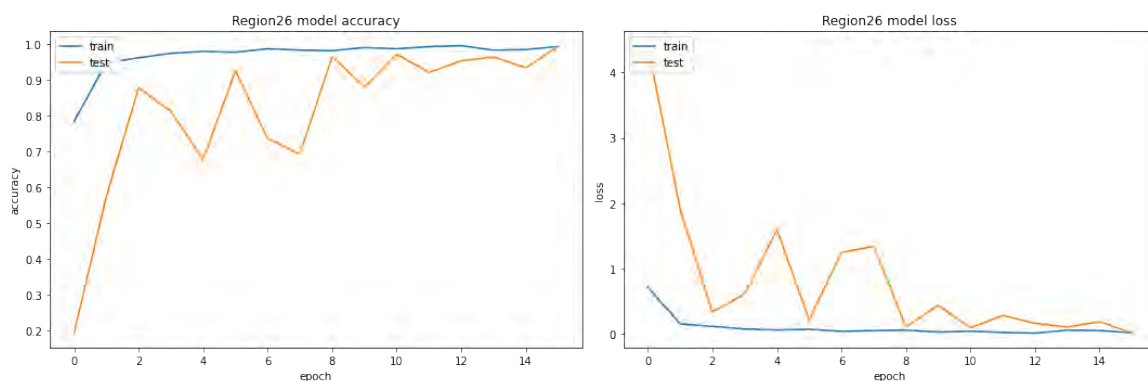


Figure A.26: ResNet model's Training performance curves of Region26

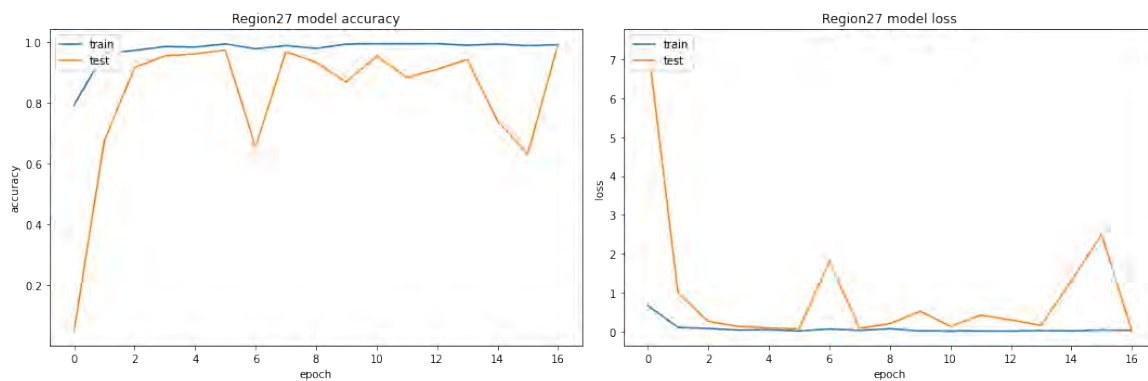


Figure A.27: ResNet model's Training performance curves of Region27

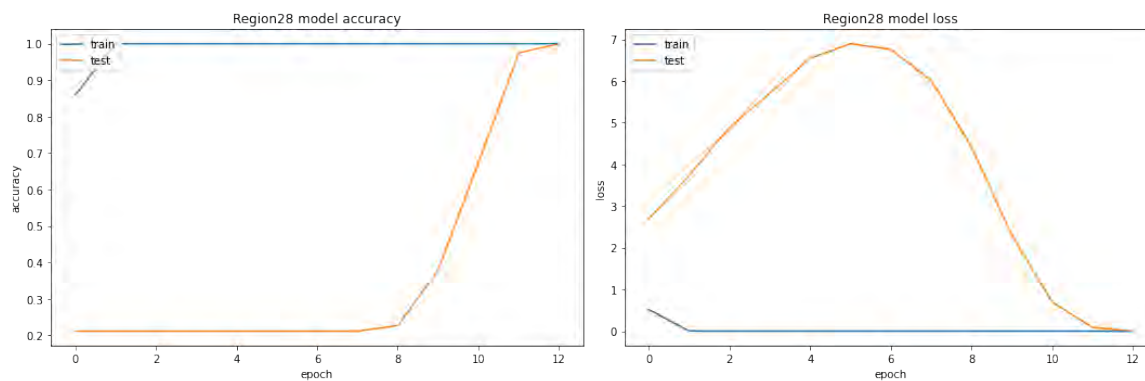


Figure A.28: ResNet model's Training performance curves of Region28

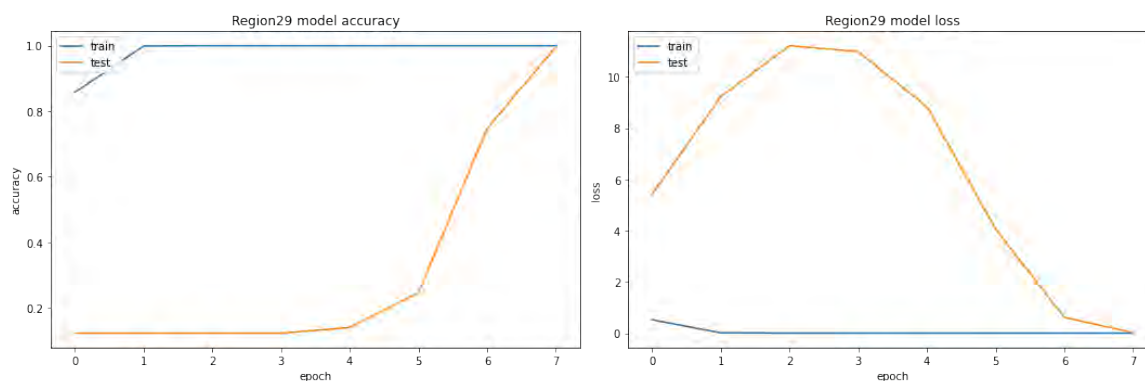


Figure A.29: ResNet model's Training performance curves of Region29

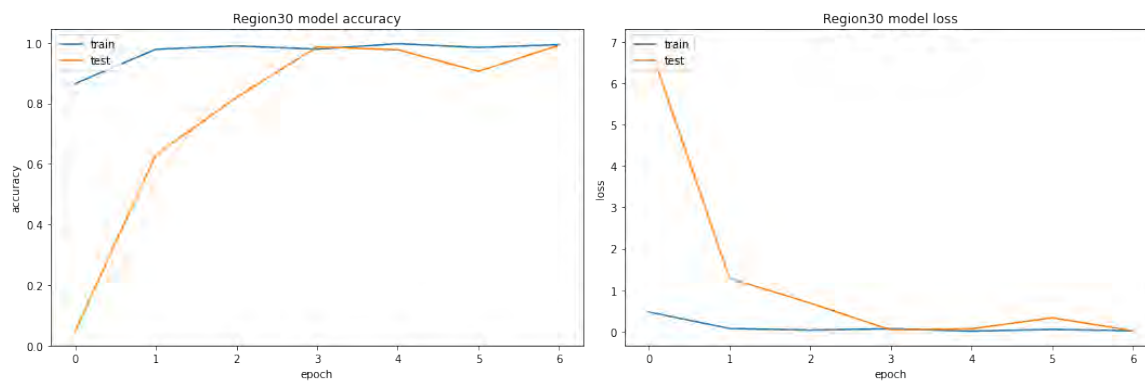


Figure A.30: ResNet model's Training performance curves of Region30

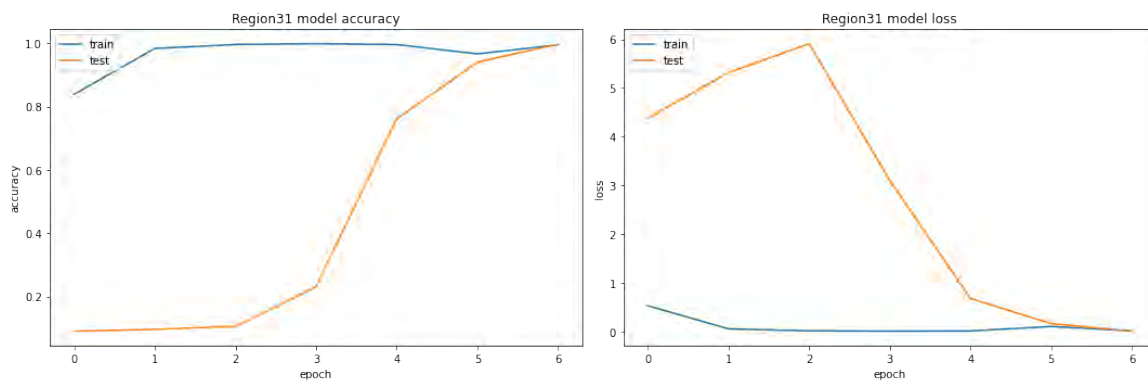


Figure A.31: ResNet model's Training performance curves of Region31

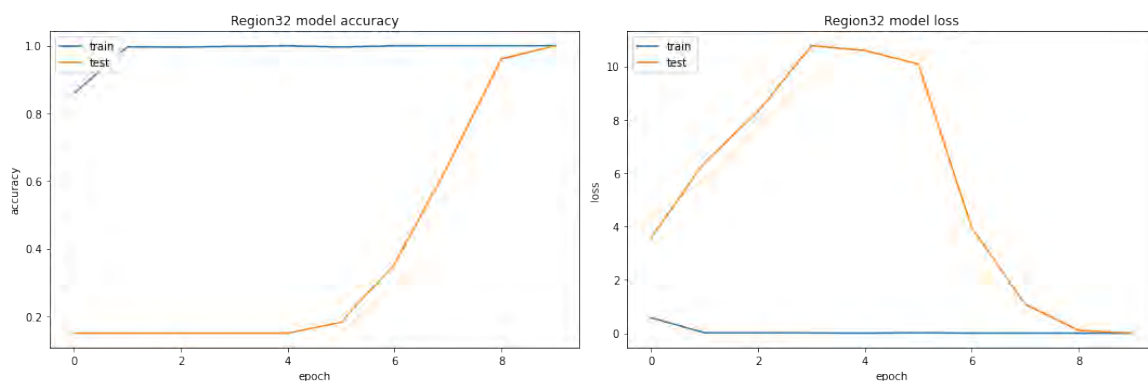


Figure A.32: ResNet model's Training performance curves of Region32

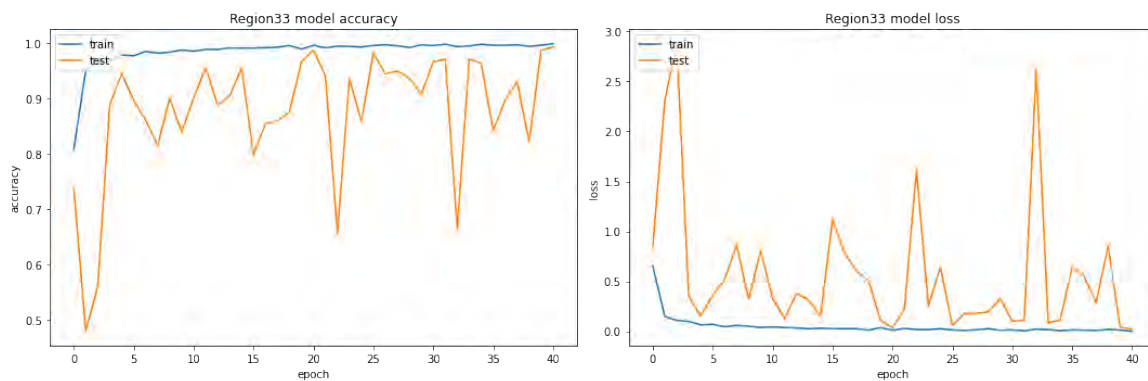


Figure A.33: ResNet model's Training performance curves of Region33

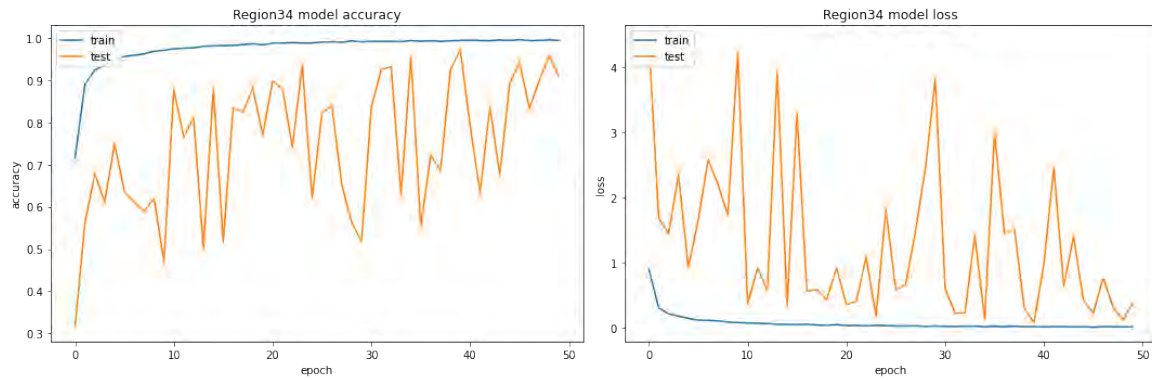


Figure A.34: ResNet model's Training performance curves of Region34

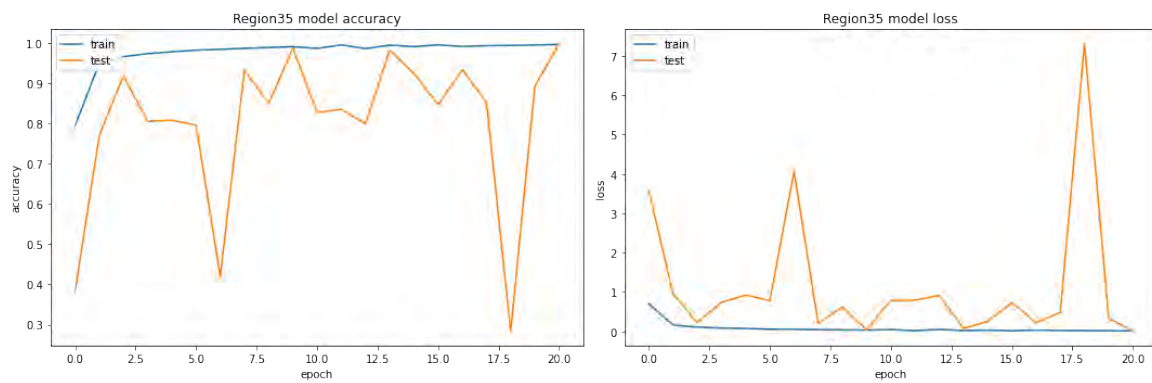


Figure A.35: ResNet model's Training performance curves of Region35

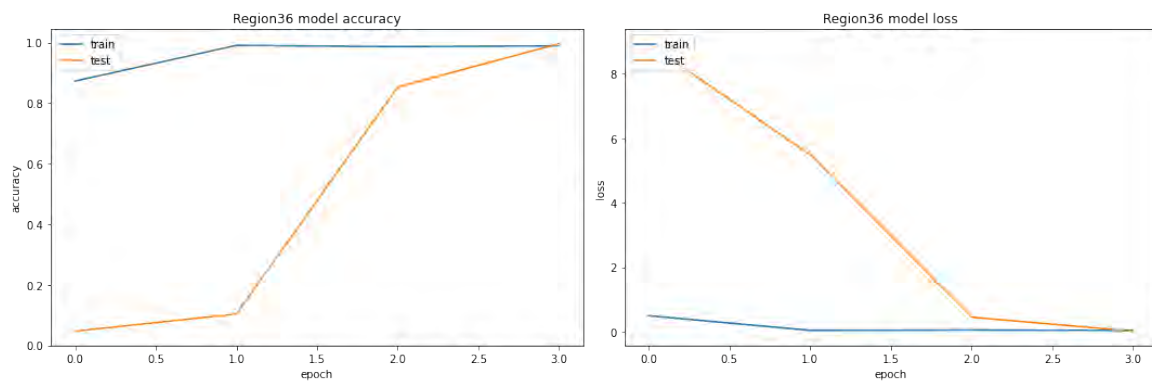


Figure A.36: ResNet model's Training performance curves of Region36

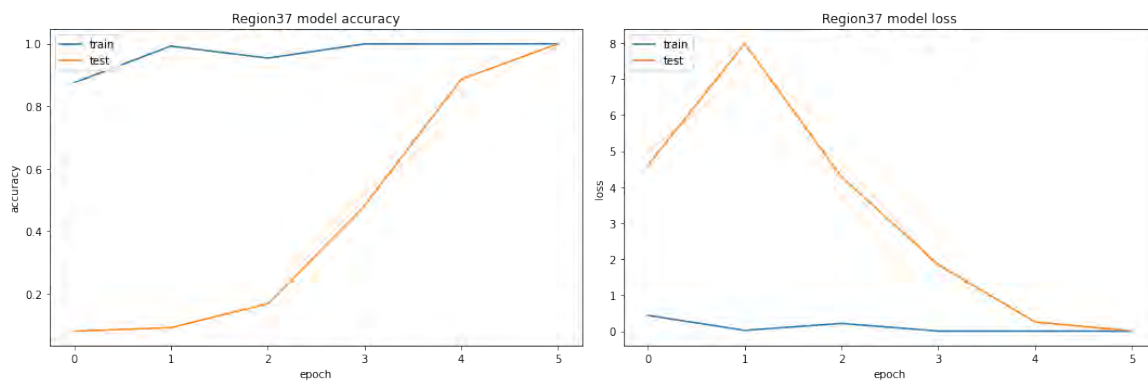


Figure A.37: ResNet model's Training performance curves of Region37

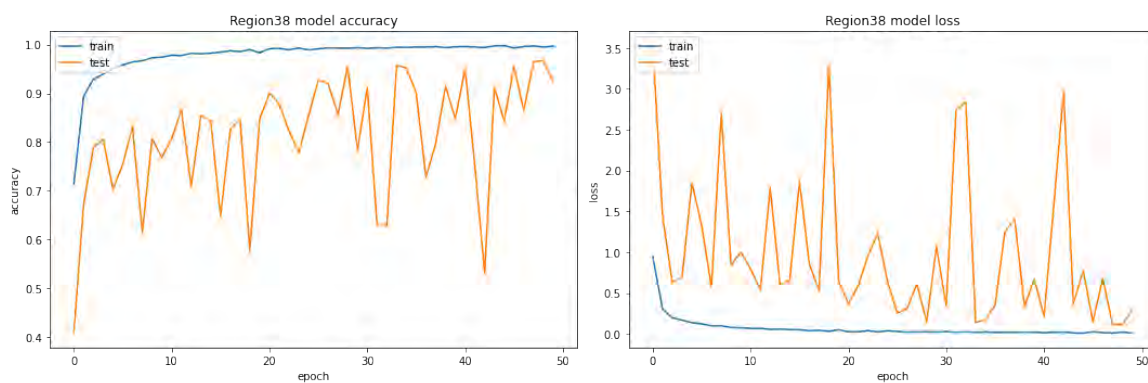


Figure A.38: ResNet model's Training performance curves of Region38

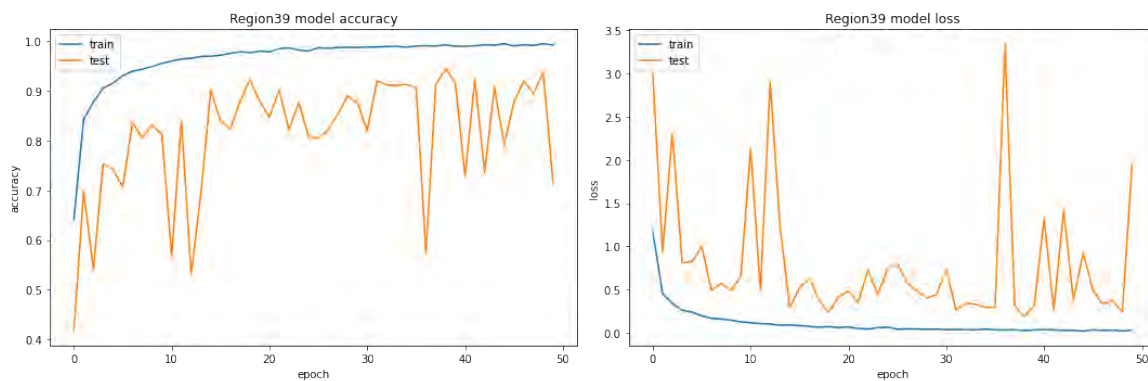


Figure A.39: ResNet model's Training performance curves of Region39

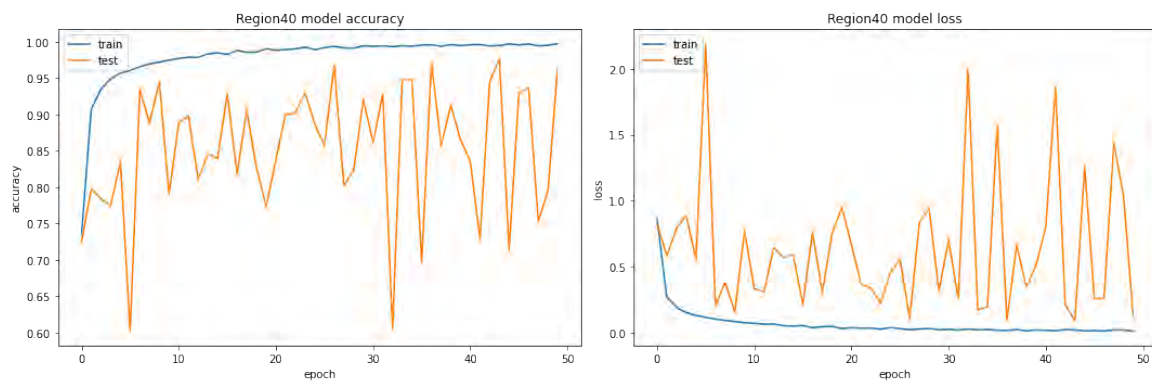


Figure A.40: ResNet model's Training performance curves of Region40

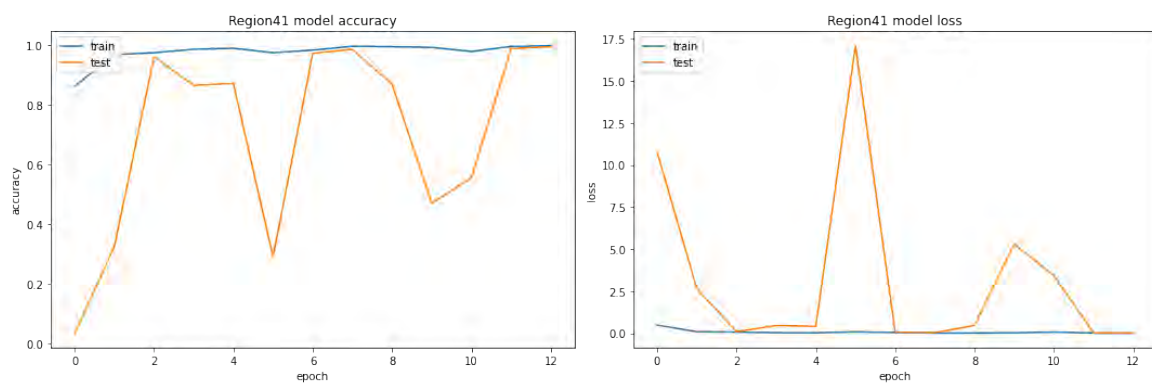


Figure A.41: ResNet model's Training performance curves of Region41

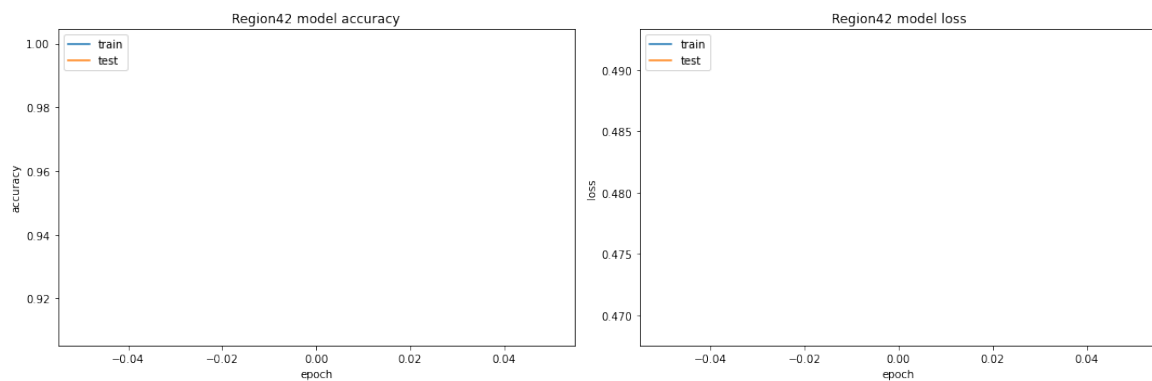


Figure A.42: ResNet model's Training performance curves of Region42

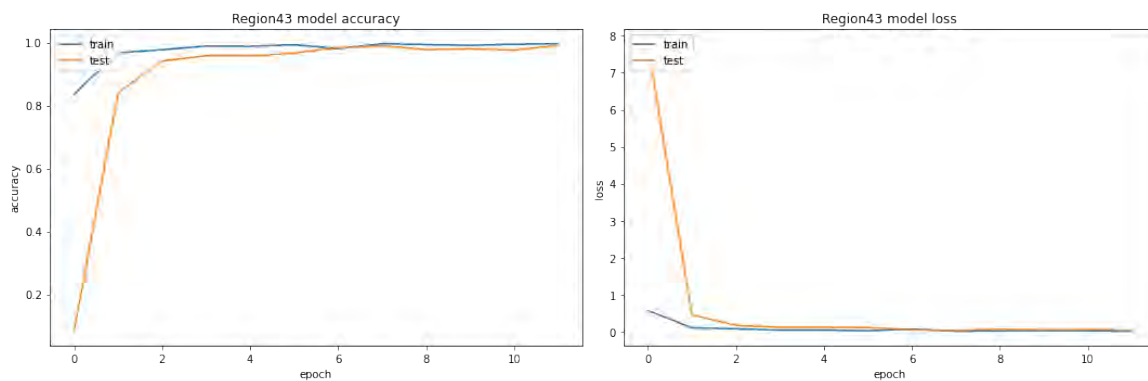


Figure A.43: ResNet model's Training performance curves of Region43

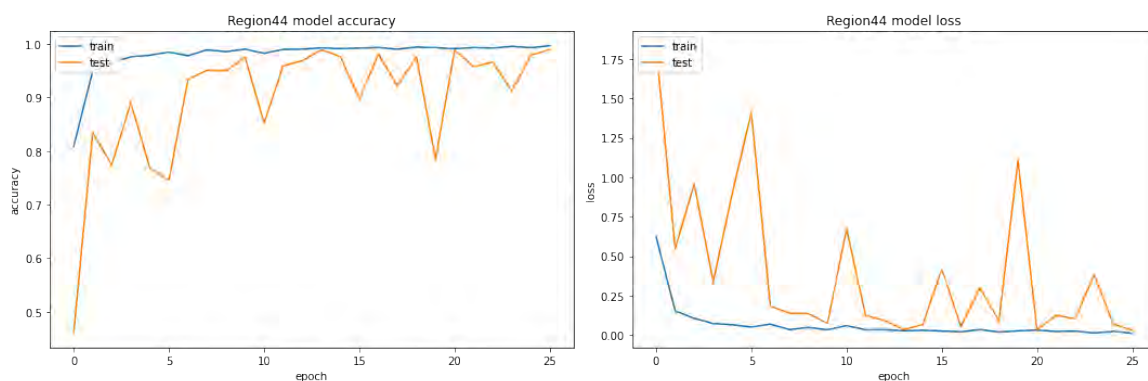


Figure A.44: ResNet model's Training performance curves of Region44

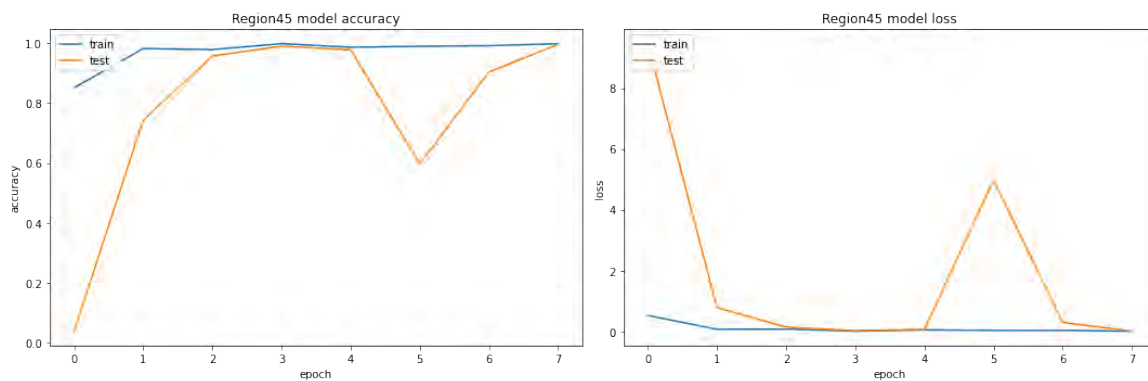


Figure A.45: ResNet model's Training performance curves of Region45

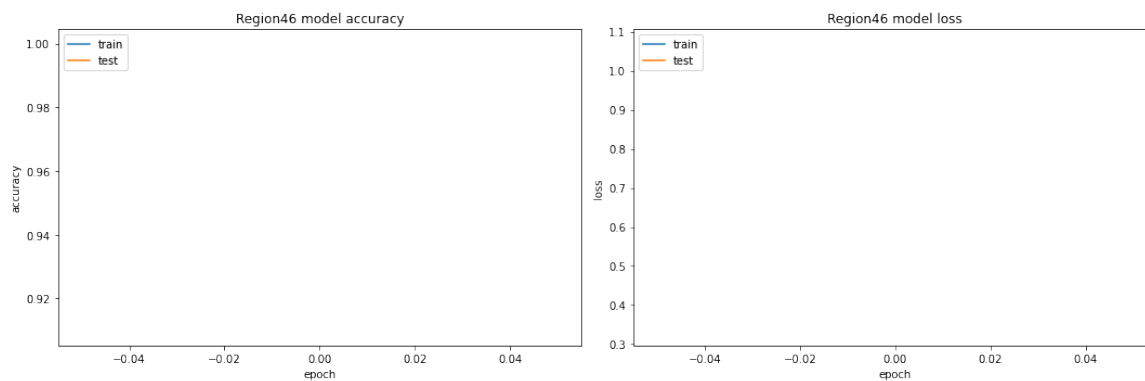


Figure A.46: ResNet model's Training performance curves of Region46

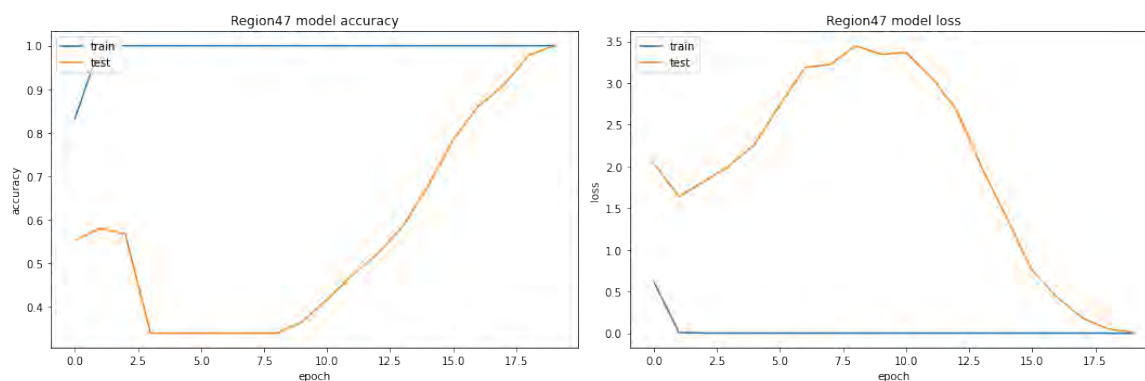


Figure A.47: ResNet model's Training performance curves of Region47

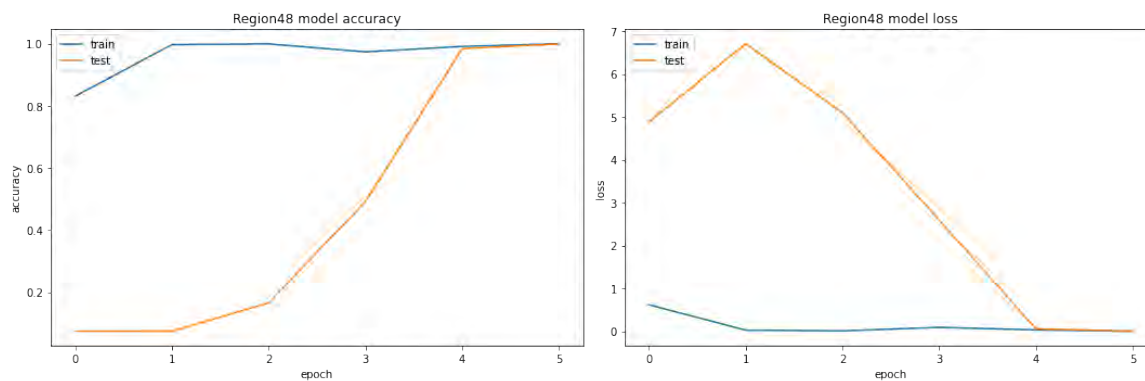


Figure A.48: ResNet model's Training performance curves of Region48

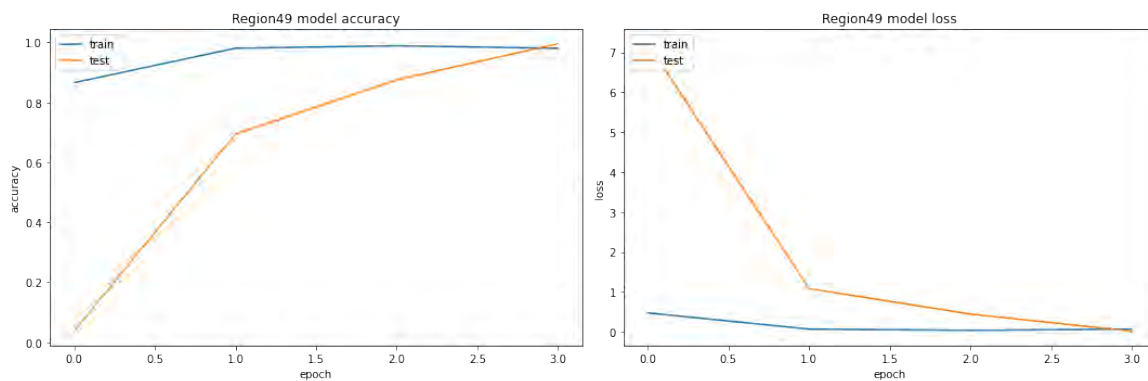


Figure A.49: ResNet model's Training performance curves of Region49

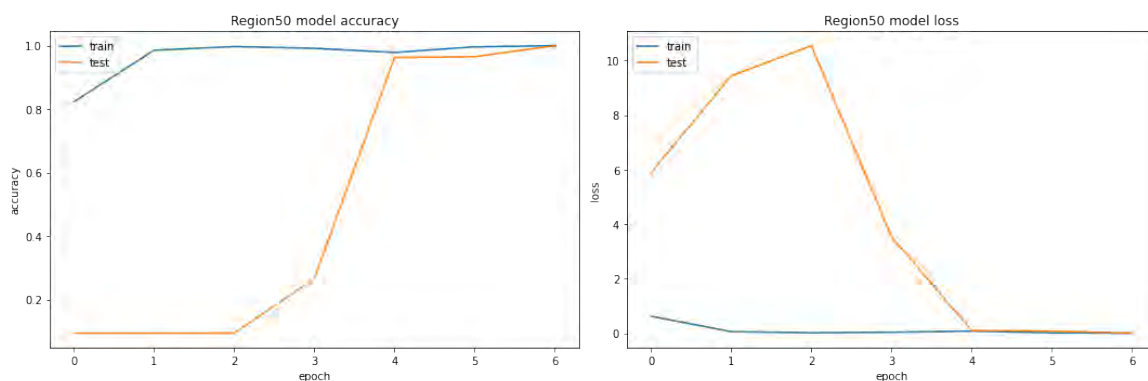


Figure A.50: ResNet model's Training performance curves of Region50

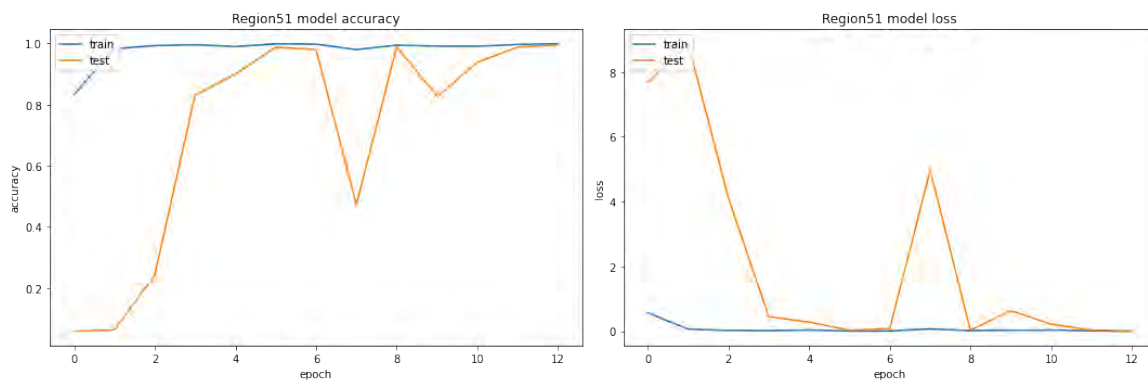


Figure A.51: ResNet model's Training performance curves of Region51

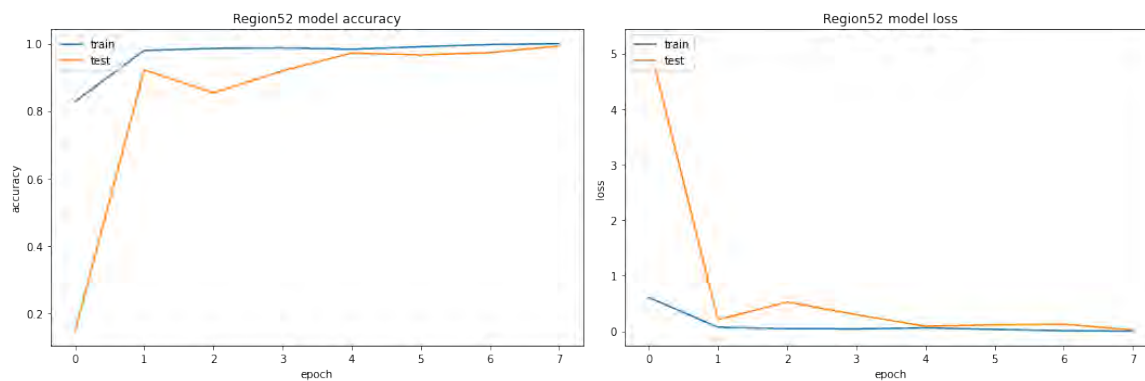


Figure A.52: ResNet model's Training performance curves of Region52

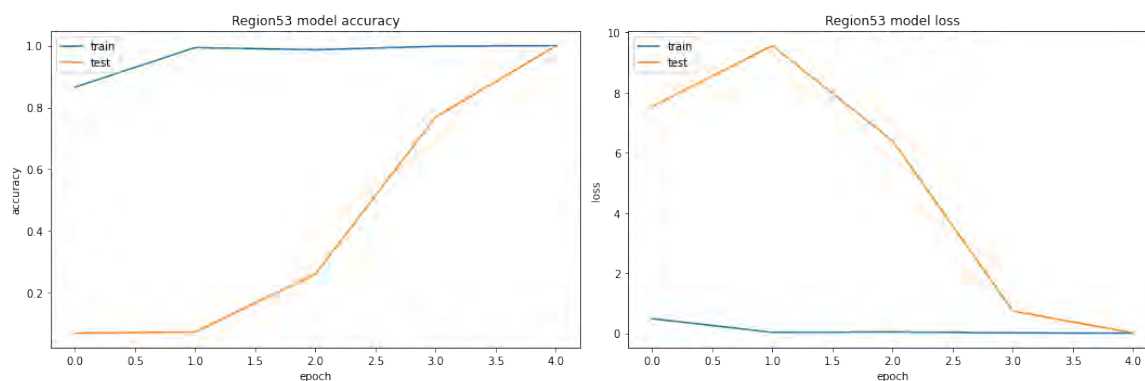


Figure A.53: ResNet model's Training performance curves of Region53

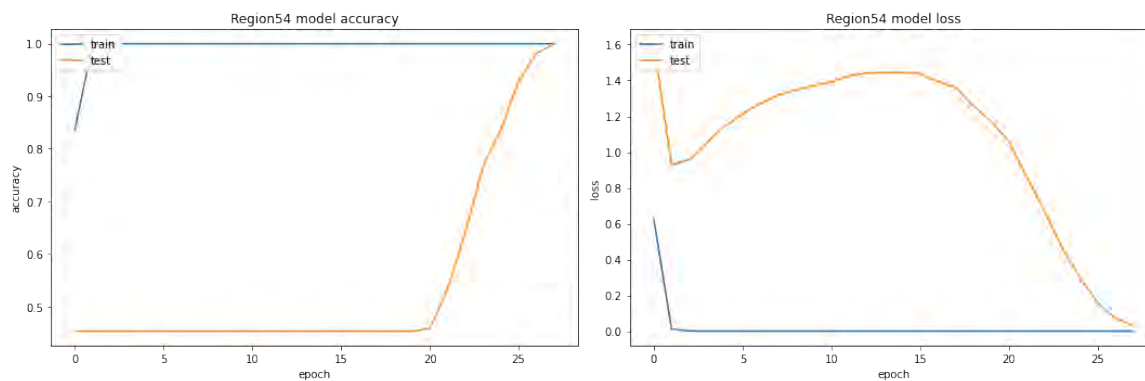


Figure A.54: ResNet model's Training performance curves of Region54

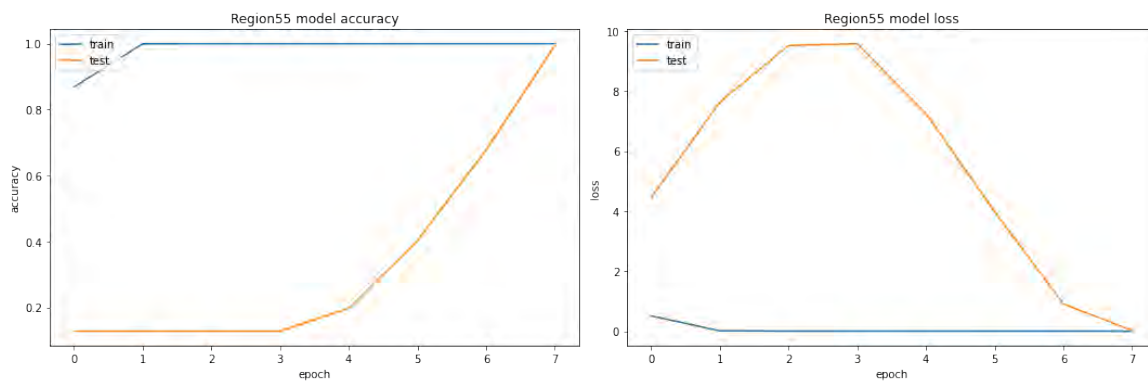


Figure A.55: ResNet model's Training performance curves of Region55

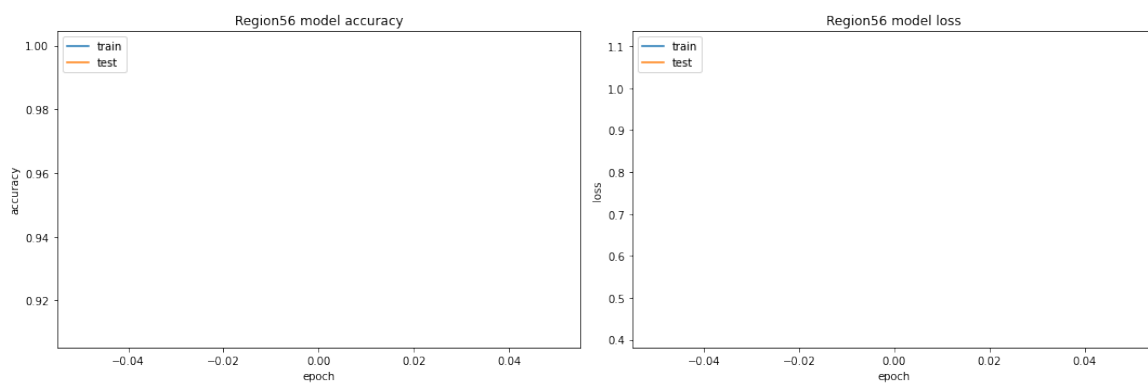


Figure A.56: ResNet model's Training performance curves of Region56

REFERENCES

- [1] M. Greiner. Earthquake: Scientists measure dream content for the first time. [Online]. Available: https://www.mpg.de/4621589/measuring_dream_content
- [2] V. Morin. The devil is in the details: Predicting alzheimer's. [Online]. Available: <https://www.everythingzoomer.com/health/2013/10/30/predicting-alzheimers/>
- [3] F. Liu, S. Wang, J. Rosenberger, J. Su, and H. Liu, "A sparse dictionary learning framework to discover discriminative source activations in eeg brain mapping," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [4] S. I. Serengil. Handwritten digit recognition using cnn with keras. [Online]. Available: <https://sefiks.com/2017/11/05/handwritten-digit-recognition-using-cnn-with-keras/>
- [5] F. G. Awan, O. Saleem, and A. Kiran, "Recent trends and advances in solving the inverse problem for eeg source localization," *Inverse Problems in Science and Engineering*, pp. 1–16, 2018.

- [6] B. He, A. Sohrabpour, E. Brown, and Z. Liu, “Electrophysiological source imaging: A noninvasive window to brain dynamics,” *Annual review of biomedical engineering*, vol. 20, pp. 171–196, 2018.
- [7] A. Sorrentino, L. Parkkonen, A. Pascarella, C. Campi, and M. Piana, “Dynamical meg source modeling with multi-target bayesian filtering,” *Human brain mapping*, vol. 30, no. 6, pp. 1911–1921, 2009.
- [8] X. Chen, S. Särkkä, and S. Godsill, “A bayesian particle filtering method for brain source localisation,” *Digital Signal Processing*, vol. 47, pp. 192–204, 2015.
- [9] M. S. Hämäläinen and R. J. Ilmoniemi, “Interpreting magnetic fields of the brain: minimum norm estimates,” *Medical & biological engineering & computing*, vol. 32, no. 1, pp. 35–42, 1994.
- [10] K. Uutela, M. Hämäläinen, and E. Somersalo, “Visualization of magnetoencephalographic data using minimum current estimates,” *NeuroImage*, vol. 10, no. 2, pp. 173–180, 1999.
- [11] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, “Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain,” *International Journal of psychophysiology*, vol. 18, no. 1, pp. 49–65, 1994.
- [12] S. Iwaki and S. Ueno, “Weighted minimum-norm source estimation of magnetoencephalography utilizing the temporal information of the measured data,” *Journal of Applied physics*, vol. 83, no. 11, pp. 6441–6443, 1998.

- [13] R. D. Pascual-Marqui *et al.*, “Standardized low-resolution brain electromagnetic tomography (sloreta): technical details,” *Methods Find Exp Clin Pharmacol*, vol. 24, no. Suppl D, pp. 5–12, 2002.
- [14] J. Bosch-Bayard, P. Valdes-Sosa, T. Virues-Alba, E. Aubert-Vazquez, E. R. John, T. Harmony, J. Riera-Diaz, and N. Trujillo-Barreto, “3d statistical parametric mapping of eeg source spectra by means of variable resolution electromagnetic tomography (vareta),” *Clinical Electroencephalography*, vol. 32, no. 2, pp. 47–61, 2001.
- [15] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, “Localization of brain electrical activity via linearly constrained minimum variance spatial filtering,” *IEEE Transactions on biomedical engineering*, vol. 44, no. 9, pp. 867–880, 1997.
- [16] J. C. Mosher and R. M. Leahy, “Recursive music: a framework for eeg and meg source localization,” *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 11, pp. 1342–1354, 1998.
- [17] C. Phillips, J. Mattout, M. D. Rugg, P. Maquet, and K. J. Friston, “An empirical bayesian solution to the source reconstruction problem in eeg,” *NeuroImage*, vol. 24, no. 4, pp. 997–1011, 2005.
- [18] J. Mattout, C. Phillips, W. D. Penny, M. D. Rugg, and K. J. Friston, “Meg source localization under multiple constraints: an extended bayesian framework,” *NeuroImage*, vol. 30, no. 3, pp. 753–767, 2006.

- [19] Y. Petrov, “Harmony: Eeg/meg linear inverse source reconstruction in the anatomical basis of spherical harmonics,” *PLoS one*, vol. 7, no. 10, p. e44439, 2012.
- [20] I. F. Gorodnitsky, J. S. George, and B. D. Rao, “Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm,” *Electroencephalography and clinical Neurophysiology*, vol. 95, no. 4, pp. 231–251, 1995.
- [21] M.-a. Sato, T. Yoshioka, S. Kajihara, K. Toyama, N. Goda, K. Doya, and M. Kawato, “Hierarchical bayesian estimation for meg inverse problem,” *NeuroImage*, vol. 23, no. 3, pp. 806–826, 2004.
- [22] T. Auranen, A. Nummenmaa, M. S. Hämäläinen, I. P. Jääskeläinen, J. Lampinen, A. Vehtari, and M. Sams, “Bayesian analysis of the neuromagnetic inverse problem with p-norm priors,” *NeuroImage*, vol. 26, no. 3, pp. 870–884, 2005.
- [23] J. M. Zumer, H. T. Attias, K. Sekihara, and S. S. Nagarajan, “A probabilistic algorithm integrating source localization and noise suppression of meg and eeg data,” in *Advances in Neural Information Processing Systems*, 2007, pp. 1625–1632.
- [24] C. Stahlhut, M. Mørup, O. Winther, and L. K. Hansen, “Simultaneous eeg source and forward model reconstruction (sofomore) using a hierarchical bayesian approach,” *Journal of Signal Processing Systems*, vol. 65, no. 3, pp. 431–444, 2011.

- [25] S. Baillet and L. Garnero, “A bayesian approach to introducing anatomofunctional priors in the eeg/meg inverse problem,” *IEEE transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 374–385, 1997.
- [26] T. Limpiti, B. D. Van Veen, H. T. Attias, and S. S. Nagarajan, “A spatiotemporal framework for estimating trial-to-trial amplitude variation in event-related meg/eeg,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 633–645, 2009.
- [27] M. Dannhauer, E. Lämmel, C. H. Wolters, and T. R. Knösche, “Spatio-temporal regularization in linear distributed source reconstruction from eeg/meg: a critical evaluation,” *Brain topography*, vol. 26, no. 2, pp. 229–246, 2013.
- [28] J.-M. Lina, R. Chowdhury, E. Lemay, E. Kobayashi, and C. Grova, “Wavelet-based localization of oscillatory sources from magnetoencephalography data,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 8, pp. 2350–2364, 2014.
- [29] S. Castaño-Candamil, J. Höhne, J.-D. Martínez-Vargas, X.-W. An, G. Castellanos-Domínguez, and S. Haufe, “Solving the eeg inverse problem based on space–time–frequency structured sparsity constraints,” *NeuroImage*, vol. 118, pp. 598–612, 2015.
- [30] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*. siam, 2005, vol. 89.

- [31] A. Solin, P. Jylänki, J. Kauramäki, T. Heskes, M. A. van Gerven, and S. Särkkä, “Regularizing solutions to the meg inverse problem using space-time separable covariance functions,” *arXiv preprint arXiv:1604.04931*, 2016.
- [32] D. Kuzin, O. Isupova, and L. Mihaylova, “Spatio-temporal structured sparse regression with hierarchical gaussian process priors,” *IEEE Transactions on Signal Processing*, vol. 66, no. 17, pp. 4598–4611, 2018.
- [33] E. Pirondini, B. Babadi, G. Obregon-Henao, C. Lamus, W. Q. Malik, M. S. Hämäläinen, and P. L. Purdon, “Computationally efficient algorithms for sparse, dynamic solutions to the eeg source localization problem,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 6, pp. 1359–1372, 2018.
- [34] A. Al Hilli, L. Najafizadeh, and A. P. Petropulu, “A weighted approach for sparse signal support estimation with application to eeg source localization,” *IEEE Transactions on Signal Processing*, vol. 65, no. 24, pp. 6551–6565, 2017.
- [35] F. Costa, H. Batatia, L. Chaari, and J.-Y. Tournier, “Sparse eeg source localization using bernoulli laplacian priors,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 12, pp. 2888–2898, 2015.
- [36] T. Mannepalli and A. Routray, “Certainty-based reduced sparse solution for dense array eeg source localization,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 2, pp. 172–178, 2019.

- [37] A. Gramfort, M. Kowalski, and M. Hämäläinen, “Mixed-norm estimates for the m/eeg inverse problem using accelerated gradient methods,” *Physics in Medicine & Biology*, vol. 57, no. 7, p. 1937, 2012.
- [38] I. F. Gorodnitsky and B. D. Rao, “Sparse signal reconstruction from limited data using focuss: A re-weighted minimum norm algorithm,” *IEEE Transactions on signal processing*, vol. 45, no. 3, pp. 600–616, 1997.
- [39] U. R. Abeyratne, Y. Kinouchi, H. Oki, J. Okada, F. Shichijo, and K. Matsumoto, “Artificial neural networks for source localization in the human brain,” *Brain Topography*, vol. 4, no. 1, pp. 3–21, 1991.
- [40] U. Abeyratne, Y. Kinouchi, H. Oki, J. Okada, F. Shichijo, and K. Matsumoto, “A new dipole localization method using a backpropagation neural network,” *Brain Topography*, vol. 3, no. 1, p. 221, 1990.
- [41] Y. Kinouchi, G. Ohara, H. Nagashino, T. Soga, F. Shichijo, and K. Matsumoto, “Dipole source localization of meg by bp neural networks,” *Brain Topography*, vol. 8, no. 3, pp. 317–321, 1996.
- [42] G. Van Hoey, J. De Clercq, B. Vanrumste, R. Van de Walle, I. Lemahieu, M. D’Havé, and P. Boon, “Eeg dipole source localization using artificial neural networks,” *Physics in Medicine & Biology*, vol. 45, no. 4, p. 997, 2000.
- [43] S. C. Jun, B. A. Pearlmutter, and G. Nolte, “Fast accurate meg source localization using a multilayer perceptron trained with real brain noise,” *Physics in Medicine & Biology*, vol. 47, no. 14, p. 2547, 2002.

- [44] S. C. Jun and B. A. Pearlmutter, “Fast robust subject-independent magnetoencephalographic source localization using an artificial neural network,” *Human brain mapping*, vol. 24, no. 1, pp. 21–34, 2005.
- [45] J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, “One network to solve them all—solving linear inverse problems using deep projection models,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5888–5897.
- [46] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.
- [47] J. Adler and O. Öktem, “Solving ill-posed inverse problems using iterative deep neural networks,” *Inverse Problems*, vol. 33, no. 12, p. 124007, 2017.
- [48] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, “Using deep neural networks for inverse problems in imaging: beyond analytical methods,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, 2018.
- [49] F. Wang, A. Eljarrat, J. Müller, T. R. Henninen, R. Erni, and C. T. Koch, “Multi-resolution convolutional neural networks for inverse problems,” *Scientific reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [50] Z. Wei and X. Chen, “Deep-learning schemes for full-wave nonlinear inverse scattering problems,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 1849–1860, 2018.

- [51] H. R. Tamaddon-Jahromi, N. K. Chakshu, I. Sazonov, L. M. Evans, H. Thomas, and P. Nithiarasu, “Data-driven inverse modelling through neural network (deep learning) and computational heat transfer,” *Computer Methods in Applied Mechanics and Engineering*, vol. 369, p. 113217, 2020.
- [52] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, “A deep cascade of convolutional neural networks for dynamic mr image reconstruction,” *IEEE transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, 2017.
- [53] D. Pantazis and A. Adler, “Meg source localization via deep learning,” *arXiv preprint arXiv:2012.00588*, 2020.
- [54] L. Hecker, R. Rupprecht, L. T. van Elst, and J. Kornmeier, “Convdip: A convolutional neural network for better m/eeg source imaging,” *bioRxiv*, 2020.
- [55] A. Razorenova, N. Yavich, M. Malovichko, M. Fedorov, N. Koshev, and D. V. Dylov, “Deep learning for non-invasive cortical potential imaging,” in *Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-oncology*. Springer, 2020, pp. 45–55.
- [56] R. Sun, A. Sohrabpour, S. Ye, and B. He, “Sifnet: Electromagnetic source imaging framework using deep neural networks,” *bioRxiv*, 2020.
- [57] C. Dinh, J. G. Samuelsson, A. Hunold, M. S. Hämäläinen, and S. Khan, “Contextual minimum-norm estimates (cmne): a deep learning method for source estimation in neuronal networks,” *arXiv preprint arXiv:1909.02636*, 2019.

- [58] S. Cui, L. Duan, B. Gong, Y. Qiao, F. Xu, J. Chen, and C. Wang, “Eeg source localization using spatio-temporal neural network,” *China Communications*, vol. 16, no. 7, pp. 131–143, 2019.
- [59] E. Giraldo-Suarez, J. D. Martínez-Vargas, and G. Castellanos-Dominguez, “Reconstruction of neural activity from eeg data using dynamic spatiotemporal constraints,” *International Journal of Neural Systems*, vol. 26, no. 07, p. 1650026, 2016.
- [60] P. Broca, “Remarques sur le siège de la faculté du langage articulé, suivies d’une observation d’aphémie (perte de la parole),” *Bulletin et Memoires de la Societe anatomique de Paris*, vol. 6, pp. 330–357, 1861.
- [61] N. F. Dronkers, O. Plaisant, M. T. Iba-Zizen, and E. A. Cabanis, “Paul broca’s historic cases: high resolution mr imaging of the brains of leborgne and lelong,” *Brain*, vol. 130, no. 5, pp. 1432–1441, 2007.
- [62] [Online]. Available: <https://hiddentalents.org/brain/113-maps.html>
- [63] M. A. Jatoi and N. Kamel, *Brain Source Localization Using EEG Signal Analysis*. CRC Press, 2017.
- [64] R. Oostenveld. Eeg signal background and real-time processing. [Online]. Available: <https://pt.slideshare.net/RobertOostenveld/eeg-signal-background-and-realtime-processing>
- [65] R. G. Baraniuk, “Compressive sensing,” *IEEE signal processing magazine*, vol. 24, no. 4, 2007.

- [66] S. Li, L. Da Xu, and X. Wang, “Compressed sensing signal and data acquisition in wireless sensor networks and internet of things,” *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2177–2186, 2013.
- [67] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [68] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [69] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [70] D. Zachariah, S. Chatterjee, and M. Jansson, “Dynamic subspace pursuit,” in *2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2012, pp. 3605–3608.
- [71] D. Needell and R. Vershynin, “Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit,” *IEEE Journal of selected topics in signal processing*, vol. 4, no. 2, pp. 310–316, 2010.
- [72] D. Needell and J. A. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and computational harmonic analysis*, vol. 26, no. 3, pp. 301–321, 2009.

- [73] J. Haupt and R. Nowak, “Signal reconstruction from noisy random projections,” *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.
- [74] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, “Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit,” *IEEE transactions on Information Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.
- [75] P. Schniter, L. C. Potter, and J. Ziniel, “Fast bayesian matching pursuit,” in *2008 Information Theory and Applications Workshop*. IEEE, 2008, pp. 326–333.
- [76] C. M. Bishop and M. Tipping, “Variational relevance vector machines,” *arXiv preprint arXiv:1301.3838*, 2013.
- [77] M. E. Tipping, “Sparse bayesian learning and the relevance vector machine,” *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244, 2001.
- [78] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” *IEEE Journal of selected topics in signal processing*, vol. 1, no. 4, pp. 586–597, 2007.
- [79] S. Ji and L. Carin, “Bayesian compressive sensing and projection optimization,” in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 377–384.
- [80] E. G. Larsson and Y. Selén, “Linear regression with a sparse parameter vector,” *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 451–460, 2007.

- [81] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Transactions on information theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [82] M. Yuan and Y. Lin, “Model selection and estimation in regression with grouped variables,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.
- [83] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, “Block-sparse signals: Uncertainty relations and efficient recovery,” *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.
- [84] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in *Proceedings of 27th Asilomar conference on signals, systems and computers*. IEEE, 1993, pp. 40–44.
- [85] Y. C. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.
- [86] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and computational harmonic analysis*, vol. 27, no. 3, pp. 265–274, 2009.

- [87] V. Cevher, C. Hegde, M. F. Duarte, and R. G. Baraniuk, “Sparse signal recovery using markov random fields,” RICE UNIV HOUSTON TX DEPT OF ELECTRICAL AND COMPUTER ENGINEERING, Tech. Rep., 2009.
- [88] A. A. Quadeer and T. Y. Al-Naffouri, “Structure-based bayesian sparse reconstruction,” *IEEE transactions on signal processing*, vol. 60, no. 12, pp. 6354–6367, 2012.
- [89] Z. Zhang and B. D. Rao, “Recovery of block sparse signals using the framework of block sparse bayesian learning,” in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 3345–3348.
- [90] M. Masood and T. Y. Al-Naffouri, “Support agnostic bayesian matching pursuit for block sparse signals,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 4643–4647.
- [91] —, “Support agnostic bayesian recovery of jointly sparse signals,” in *2014 22nd European Signal Processing Conference (EUSIPCO)*. IEEE, 2014, pp. 1741–1745.
- [92] D. Wipf and S. Nagarajan, “A unified bayesian framework for meg/eeg source imaging,” *NeuroImage*, vol. 44, no. 3, pp. 947–966, 2009.
- [93] J. W. Phillips, R. M. Leahy, and J. C. Mosher, “Meg-based imaging of focal neuronal current sources,” *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 338–348, 1997.

- [94] G. Obozinski, M. J. Wainwright, and M. I. Jordan, “Support union recovery in high-dimensional multivariate regression,” *The Annals of Statistics*, vol. 39, no. 1, pp. 1–47, 2011.
- [95] A. S. Cichocki, “Blind source separation: New tools for extraction of source signals and denoising,” in *Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks III*, vol. 5818. International Society for Optics and Photonics, 2005, pp. 11–25.
- [96] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.
- [97] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, “Simultaneous sparse approximation via greedy pursuit,” in *Proceedings.(ICASSP’05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 5. IEEE, 2005, pp. v–721.
- [98] D. P. Wipf and B. D. Rao, “An empirical bayesian strategy for solving the simultaneous sparse approximation problem,” *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.
- [99] M. Mishali and Y. C. Eldar, “Reduce and boost: Recovering arbitrary sets of jointly sparse vectors,” *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4692–4702, 2008.

- [100] Z. Zhang and B. D. Rao, “Sparse signal recovery in the presence of correlated multiple measurement vectors,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2010, pp. 3986–3989.
- [101] Y. Jin and B. D. Rao, “Support recovery of sparse signals in the presence of multiple measurement vectors,” *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 3139–3157, 2013.

VITAE

- Name: Elmubarak M. Imam
- Nationality: Sudanese
- Date of Birth: October 19, 1992
- Email: *elmubarakimam@gmail.com*
- Permanent Address: Almanshia B/36, Khartoum, B.O.: 11111, Sudan
- Publications:
 - Elhassan, Bashar S., et al. "Fault diagnosis using cross-wavelet transform for electricity pre-payment meter." 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE). IEEE, 2017.
 - Imam, Elmubarak, Mohamed Elhafiz Mohamed Ahmed, and Ghassan Abdalla. "Design and implementation of discrete cosine transform algorithm on fpga device." 2016 Conference of Basic Sciences and Engineering Studies (SGCAC). IEEE, 2016.