BABEŞ BOLYAI UNIVERSITY, CLUJ NAPOCA, ROMÂNIA
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Advancements in Skull Stripping and Multiple Sclerosis Lesion Segmentation for Neuroimaging Applications

PhD Thesis Summary

PhD student: Maria Popa
Scientific supervisor: Prof. dr. Anca Andreica

2025

Acknowledgments

I would like to begin this thesis by expressing my sincere gratitude to several people whose guidance, support, and encouragement have made this work possible.

I am deeply thankful to my advisor, Prof. Dr. Anca Andreica, for her invaluable guidance, encouragement, and continuous support throughout this journey.

My sincere appreciation also goes to the faculty members, professors, and colleagues of the Department of Mathematics & Computer Science at Babeș-Bolyai University, whose advice, collaboration, and support have been extremely helpful.

I am especially grateful to Dr. Ing. Liviu Badea, Senior Researcher and Head of the Artificial Intelligence and Bioinformatics Laboratory at ICI Bucharest, for his generous support, collaboration, and the inspiring work we accomplished together.

I would also like to extend my sincere thanks to Dr. Gabriela Vișa and Dr. Ciprian Șofariu for their advice, support, and fruitful collaboration throughout this journey.

Finally, I am deeply grateful to my friends and family for their patience, love, and encouragement throughout this endeavor.

Abstract

The challenges of analyzing large volumes of MRI data for early diagnisus and monitoring of neurological diseases, such as Multiple Sclerosis (MS), Alzheimer's disease, and Glioblastoma, highlight the need for accurate and automated systems to assist medical professionals. Since MRI is the most widely used method, a preliminary step is skull stripping, a process that removes the skull to better view the brain tissue.

This thesis aims to contribute to the development of methods for these problems. First, an unsupervised graph-based skull stripping method is presented, which is improved and evaluated on several datasets, showing performance comparable to state-of-the-art methods. Then, the research continues with Multiple Sclerosis lesion segmentation, inspecting the generalizability of the models.

Since datasets are a major limitation to these models' performance, we address the critical lack of pediatric data by introducing two new pediatric datasets, PediMS and PediDemi. PediMS contains patients diagnosed with MS, while PediDemi contains pediatric patients with demyelinating lesions from pathologies other than MS. Lastly, to address the critical issue of misidentifying small lesions, which are often missed by standard metrics like the Dice score, we proposed a novel size-aware metric that groups lesions based on their size and computes the Dice score for each size category.

Ultimately, the thesis presents a set of methods and publicly available datasets that could accelerate the development of more accurate and generalizable deep learning models for diagnosing and monitoring neurological diseases in both adults and children.

Contents

1 Introduction **4**
- 1.1 Motivation 4
- 1.2 Objectives 5
- 1.3 Original contributions 5

2 Background **10**

3 Novel graph-based skull stripping methods **12**
- 3.1 New unsupervised graph-based skull stripping methods 12
- 3.2 Exploratory Study of Graph convolutional neural networks for skull stripping 13

4 Demyelinating lesion segmentation **14**

5 Lesion-wise metric **15**

Conclusions and future work **16**

Bibliography **19**

List of Publications

Publications were ranked based on CNATDCU standards for doctoral students enrolled after October 1, 2018. The rankings follow the classifications for Computer Sciences journals and conferences, referencing UEFISCDI for journals[1] and CORE for conferences[2].

Publications in Web of Science

[45] Maria Popa, Gabriela Adriana Vișa, and Ciprian Radu Șofariu. Pedims: A pediatric multiple sclerosis lesion segmentation dataset. *Scientific Data*, 12(1):1184, Jul 2025 (**AIS Quartile Q1, WoS category: Multidisciplinary sciences**)
Rank A, 8 points

[6] Liviu Badea and Maria Popa. Toward generalizable multiple sclerosis lesion segmentation models. *IEEE Access*, 13:97859–97869, 2025 (**AIS Quartile Q2, WoS category: Mathematical & Computational Biology - SCIE**)
Rank B, 4 points

Publications in Web of Science, Conference Proceedings Citation Index

[39] Maria Popa. An 3d mri unsupervised graph-based skull stripping algorithm. *Procedia Computer Science*, 225:1682–1690, 2023. 27th International Conference on Knowledge Based and Intelligent Information and Engineering Sytems (KES 2023)
Rank B, 4 points

---

[1] https://uefiscdi.ro/premierea-rezultatelor-cercetarii-articole

[2] http://portal.core.edu.au/conf-ranks/

[43] Maria Popa and Liviu Badea. Measuring spatial specificity of multiple sclerosis lesion segmentation using dice spectra. *Procedia Computer Science*, 246:3467–3474, 2024. 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024)
Rank B, 4 points

[42] Maria Popa and Anca Andreica. Towards an improved unsupervised graph-based mri brain segmentation method. In Mohamed Sellami, Maria-Esther Vidal, Boudewijn van Dongen, Walid Gaaloul, and Hervé Panetto, editors, *Cooperative Information Systems*, pages 480–487, Cham, 2024. Springer Nature Switzerland
Rank B, 2.66 points

[40] Maria Popa. Graph convolutional neural networks for skull stripping. In Wolfgang Osten, editor, *Seventeenth International Conference on Machine Vision (ICMV 2024)*, volume 13517, page 135170T. International Society for Optics and Photonics, SPIE, 2025
Rank C, 2 points

[41] Maria Popa and Anca Andreica. Improving unsupervised graph-based skull stripping: Enhancements and comparative analysis with state-of-the-art methods. In Natalia Shakhovska, Jianbo Jiao, Ivan Izonin, and Stéphane Chrétien, editors, *Proceedings of the 7th International Conference on Informatics & Data-Driven Medicine, Birmingham, United Kingdom, November 14-16, 2024*, volume 3892 of *CEUR Workshop Proceedings*, pages 20–31. CEUR-WS.org, 2024
Rank C, 2 points

[44] Maria Popa and Gabriela Adriana Vișa. Pedidemi - a pediatric demyelinating lesion segmentation dataset. *MELBA Journal*, Open-Data 2025 - MICCAI 2025, 2025. **Accepted**
Rank D, 1 point

Publications score: 27.66 points.

Chapter 1

Introduction

The aim of this thesis is to address key challenges in brain-related medical imaging through the application of both unsupervised and supervised techniques.

With the rise of neurological disorders, there is a stronger demand for cutting-edge algorithms that can be integrated into computer-aided tools. Enhanced by recent technological advancements, these monitoring systems are essential for disease identification and mortality reduction [48]. They are particularly vital in the context of neurological conditions, where early and accurate diagnosis can significantly improve patient outcomes.

1.1 Motivation

In the last years, there has been a noticeable increase in the prevalence of different brain-related issues, necessitating regular screenings and substantial efforts from medical personnel. The statistics underscore the demand for precise computer-assisted systems that could streamline medical tasks. According to [4, 11], each year, about 5.26 out of every 100,000 people are diagnosed with glioblastoma, with clinical trial data showing a median survival of 12 to 14 months. The survival rate of those diagnosed patients is fewer than 10% in the second year [5]. From 2014 to 2018, an alarming 87000 individuals developed glioblastoma, a highly aggressive brain tumor, as cited in [30]. According to [35], between 2015 and 2020 Alzheimer disease, the most prevalent form of dementia, was estimated at 760.5 per 100,000 inhabitants. Additionally, more than 2 million individuals are estimated to suffer from Multiple Sclerosis (MS), an autoimmune disorder impacting various cognitive, emotional, motor, sensory, or visual functions [24, 18]. Epilepsy, another chronic noncommunicable brain disorder, affects individuals of all ages, with an estimated

50 million people worldwide experiencing this condition [61]. These represent just a few of the challenges confronting the population that necessitate comprehensive brain screenings.

Given the increasing prevalence of brain pathologies, there is a rising demand for computer-aided tools that can benefit both patients and healthcare professionals. These systems have the potential to assist medical staff in prioritizing patient care, monitoring disease progression, and detecting various lesions more effectively [20].

1.2 Objectives

The focus of this thesis is on applying both unsupervised and supervised methods to address challenges related to brain imaging. Titled "Advancements in Skull Stripping and Multiple Sclerosis Lesion Segmentation for Neuroimaging Applications", the PhD thesis aims to develop and analyze advanced brain segmentation models, addressing key challenges in this research domain. The main research objectives are as follows:

RO1 *Skull stripping* Develop skull stripping methods and assess their performance across various datasets.

RO2 *Generalizability of multiple sclerosis lesion segmentation models* Investigate the generalizability of multiple sclerosis lesion segmentation models across various datasets to enhance their utility in real clinical scenarios.

RO3 *Development of a new evaluation metric for lesion segmentation* Create a novel evaluation metric for lesion segmentation, which evaluates the model's performance across different lesion size categories.

RO4 *Enhancement of open MS datasets* Enhance the availability and quality of publicly accessible datasets for multiple sclerosis to facilitate comparative studies and model development.

1.3 Original contributions

Our original contributions are discussed in Chapters 3, 4, and 5, and are summarized as follows:

1. **Skull Stripping Methods**

   Unsupervised techniques are widely used in the medical domain, particularly for segmentation tasks. Among these, graph-based methods are a key focus due to their ability to efficiently represent image structure.

A critical preliminary step in analyzing brain images is skull stripping, which involves separating the brain from the surrounding skull [36]. We propose the following original approaches for skull stripping:

(a) We developed an unsupervised graph-based segmentation method called Sphere-GUBS, which was presented in our original paper [39]. This method transforms a user-dependent and dataset-dependent approach into a user- and dataset-independent one. This approach based on the Graph-Based Unsupervised Brain Segmentation (GUBS) method presented in [31]. GUBS methods require user intervation, because specific skull-limit and brain-limit thresholds must be defined by the user for each dataset. Our proposed approach addresses these challenges by using a sphere to select the nodes inside the brain, thereby eliminating the need for user intervention and dataset-dependent parameters. The method was tested on the Neurofeedback Skull-stripped (NFBS) [46] repository.

(b) Although experiments demonstrated that this method outperforms the original GUBS method, we developed an improved version, Ellipsoid-GUBS, which is detailed in [42]. This version utilizes an ellipsoid instead of a sphere for the node selection process. The idea for using an ellipsoid was inspired by the brain extraction technique BET* [59], as an ellipsoid provides a closer approximation to the brain's shape compared to a sphere. Ellipsoid-GUBS addresses specific limitations of Sphere-GUBS, particularly in accurately modeling brain structures. The method was assessed using the same dataset, and experiments showed further improvements in performance with this enhanced approach.

(c) Given that the center of mass of the image is not always aligned with the center of the image, we developed an enhanced version of Ellipsoid-GUBS, positioning the center of the ellipsoid at the image center. Initially, we tested this new approach on the NFBS dataset, as in previous experiments, and observed slight improvements in performance. Considering these enhancements, we deemed this version optimized and subsequently tested it on four additional datasets containing diverse patient data. A comparison was also performed with two state-of-the-art methods, with the new approach showing slightly better performance. These findings are detailed in [41].

(d) Given our focus on skull stripping and graph-based methods, we explored the application of Graph Convolution Neural Networks (GCNs) [66] as a pioneering test for this task. While GCNs are typically employed for more complex neurological problems, our interest in their potential led us to apply this approach to the skull stripping task. However, as this method is relatively novel and not yet widely used for this specific purpose, it does not currently produce similar results to state-of-the-art methods. Despite this, we plan to explore GCNs further for more complex neurological tasks in the future. Our initial findings were presented in [40].

2. **Multiple Sclerosis Lesion Segmentation**

   Identifying Multiple Sclerosis (MS) lesions is difficult even for medical professionals, because of their diverse shapes and sizes [14]. Deep learning-based segmentation models excel in numerous domains; however, the state-of-the-art for MS lesion segmentation remains suboptimal [6]. Competitions like ISBI2015 [8] and MICCAI2016 [10] have focused on developing segmentation algorithms for individual datasets. In our paper [6], jointly with Dr. Liviu Badea (ICI Bucuresti), we explored the generalizability of these models and conducted a comprehensive analysis of segmentation methods systematically trained and tested on multiple publicly available datasets. The key contributions are the following:

   (a) One of the limitations of existing literature for which we propose an original and effective solution is the diversity of dataset distributions and image variation caused by differences in scanners and imaging protocols, which we mitigate by using quantile normalization [2, 7]. To support our findings, we performed an ablation study.

   (b) To demonstrate generalization, we collected the largest set of publicly available MS lesion segmentation datasets and performed all possible tests on independent datasets (i.e. datasets not seen during training). More precisely, we always tested on datasets not seen during training and we trained on all other datasets that we had at our disposal. Therefore, we think this is the most extended study of generalizability.

   (c) Using the state-of-the-art UNet++ [67] architecture, we trained models on individual datasets and tested them on other datasets. We chose the UNet++ architecture because of the slightly improved results compared to the UNet architecture [51]. We performed an ablation study comparing the two, UNet++ showing slightly improved results (but still statistically significant by a Wilcoxon test). Additionally, we trained models on combinations of datasets and tested them on the remaining ones. Without specific dataset optimization, the models trained on a combination of three datasets even slightly outperformed the winner of the MICCAI2016 competition.

(d) In collaboration with the Clinical Pediatric Hospital Sibiu, we developed and publicly released two datasets containing pediatric patients (PediMS [45] and PediDemi [44]). PediMS [45] is a longitudinal dataset for cerebral lesion segmentation in multiple sclerosis (MS), comprising 9 patients, each with between one and six time points, totaling 28 MRI scans. The second dataset, named PediDemi [44], focuses on demyelinating lesions in patients who exhibit cerebral demyelination but have not yet been diagnosed with MS. This dataset includes several cases of acute disseminated encephalomyelitis (ADEM). Both datasets were annotated by medical professionals.

3. **Lesion-wise metric**

Metrics are essential for assessing the performance of a method [55]. For evaluating model performance in medical image segmentation, Dice score [12] is the most widely used metric [9, 33]. In our paper [43], jointly with Dr. Liviu Badea (ICI Bucuresti), we introduced a new metric, the Dice Spectrum, which, for the first time, presents an extensive evaluation of the modelâ€™s performance across various lesion size categories. The major contributions are outlined below:

(a) Although a single score is produced by the standard Dice metric, it does not capture detailed performance across lesion size variations, most notably in detecting small lesions important for early disease identification. We addressed this by proposing the Dice Spectrum, a novel metric that is especially relevant for evaluating models' ability to detect small lesions. Even when a model achieves high overall Dice scores, it may still underperform on smaller lesions. The Dice Spectrum analyzes the multiple sclerosis lesions in a dataset and categorizes them by size, ensuring that each category contains an equal number of lesions. For each image slice, the lesions are grouped into their respective size categories, and a Dice score is computed for each group.

(b) This method was evaluated across multiple datasets and models, demonstrating that the models generally perform poorly on small lesions. Moreover, the Dice Spectrum is not limited to multiple sclerosis lesion segmentation but can be applied in other medical imaging contexts as well.

Chapter 2

Background

Image segmentation is a subfield of computer vision that partitions an image into regions, either based on pixel similarity or according to whether the regions represent areas of interest [37, 19, 26, 53, 65]. Image segmentation has become increasingly popular across various industries, from automotive [21, 52] to medical fields [34, 64, 60, 25].

Methods for segmentation are typically grouped into supervised and unsupervised. Unsupervised methods operate on unlabelled datasets, partitioning images based on inherent patterns such as pixel similarity or connectivity. Supervised methods, by contrast, utilize labelled datasets where examples of input data are paired with corresponding outcomes to guide the model in identifying meaningful patterns and relationships [63]. Machine Learning (ML) plays a central role in both of these approaches by enabling the extraction of patterns, structures, and insights from complex data. Unsupervised learning, a sub-field of ML, seeks to uncover hidden structures within unlabelled data, operating without direct feedback or predefined labels [15]. Supervised learning, conversely, relies on labeled data to train models for accurate pattern recognition [13, 47].

Traditional image segmentation techniques fall under the unsupervised category, relying on pixel-level characteristics for segmentation without the need for training. These approaches are generally classified into thresholding, clustering, edge-based, region-based, and graph theory-based methods [22, 29].

The rapid advancements in deep learning, have demonstrated impressive performance in image segmentation [17, 49, 16].

The state-of-the-art in supervised medical image segmentation is *UNet* [51]. The UNet architecture is designed to contract an image, analyze its features, and then reconstruct the image [49, 40]. It employs an encoder-decoder mechanism, with the encoder compressing the input data into a latent-space representation, and the decoder transforms this representation back into the desired output [17]. The UNet architecture uses a two-part design: a contracting path to extract image features and an expansive path to pinpoint their exact locations [6, 51].

A powerful version of UNet is *UNet++* [67], which was inspired by DesNet [23, 49]. It redesigns the standard UNet by replacing its skip connections with nested, dense skip pathways [56]. These pathways use dense convolution blocks to more effectively connect feature maps from the encoder to the decoder, reducing the semantic gap before their combination [6].

Graph Neural Networks (GNNs) are a type of network designed to process irregular data structures and aggregate pairwise relationships between entities [66, 40].

Graph Convolutional Neural Networks (GCNs), also known as Convolutional Graph Neural Networks (ConvGNNs) [62], are a category of GNN that generalizes the convolution operation from grids (e.g., images) to graphs [54]. Similar to unsupervised graph-based approaches, GCNs encode information within nodes [66] but additionally enable nodes to learn from their neighbors through graph-specific convolutional operations [28, 40].

In the medical field, segmentation is crucial for interpreting magnetic resonance (MR) or computed tomography (CT) images, which can be used to delineate organs [32, 50], bones [32, 27], lesions, and abnormalities [32, 3]. Neuroimaging, a subfield of medical imaging, focuses specifically on brain structure and function [58, 1]. MRI images are acquired using scanners with different parameter settings, resulting in various image types known as sequences or modalities. Each modality highlights different tissue properties, aiding in comprehensive brain analysis[57]. Common MRI modalities include: T1-weighted (T1w), T2-weighted (T2w), FLAIR, and PD-weighted.

In this work, we focus on skull stripping and multiple sclerosis (MS) lesion segmentation. Skull stripping is a preliminary technique used to eliminate the skull and other non-brain tissues, leaving only the brain on MRI scans [38]. The brain's structure is typically well-defined with clear anatomical boundaries, and due to few labelled datasets, we focus on traditional methods in skull stripping. Conversely, multiple sclerosis lesions vary in shape, size, and location, presenting a far more complex segmentation challenge. We find supervised deep learning approaches more suitable for MS lesion segmentation, as they can learn to detect these lesions from labeled data.

Chapter 3

Novel graph-based skull stripping methods

This chapter outlines our contributions to developing skull stripping methods with the focus on graph-based methods. The results presented in this chapter are detailed in the following papers [39, 42, 41, 40].

3.1 New unsupervised graph-based skull stripping methods

This text outlines the development of a brain segmentation method that is an evolution of the GUBS approach. The key goal was to create a more robust and automated process by eliminating the need for human intervention, which was a significant limitation of the original GUBS method. The core of all these methods is the use of a graph representation of an MRI image, followed by the application of a minimal spanning tree to segment the brain.

The first major improvement was Sphere-GUBS. This version removed the need for manual parameter settings and reduced computation time. It defined a spherical region to automatically select nodes for the graph, unlike GUBS which required manual "seed" points. When tested on the NFBS dataset, Sphere-GUBS outperformed the original GUBS, achieving a notable 25% increase in precision and a 20% increase in the Dice coefficient when using original image sizes.

A further enhancement led to Ellipsoid-GUBS. This method improved upon Sphere-GUBS by using an ellipsoid for node selection, which proved to be more effective. When tested on the NFBS dataset, Ellipsoid-GUBS improved upon its predecessor, showing a 3% increase in precision over Sphere-GUBS and a 28% increase over the original GUBS.

The final proposed method addressed the limitations of Ellipsoid-GUBS by centering the sampling ellipsoid at the image's geometric center instead of the center of mass. This change aimed to improve robustness across datasets with different head positions. This final version was tested on four datasets (NFBS, FMS, IXI, and QIN) and compared against two widely-used state-of-the-art tools, BET2 and BSE. The results were promising: on the QIN (infant) dataset, the method showed significant improvements over both BET2 and BSE, including a 14% increase in precision over BET2 and a 6% increase in the Dice coefficient over BSE. It also performed better than BET2 and BSE on the FMS dataset. While BSE performed best on the NFBS dataset and BET2 was slightly better on the IXI dataset, the final proposed method generally showed competitive or superior performance, especially for specific datasets.

3.2 Exploratory Study of Graph convolutional neural networks for skull stripping

This research proposes a graph neural network (GNN) for skull stripping. The method converts MRI slices into graphs, where each pixel is a node, to perform brain segmentation using a binary graph convolutional neural network (GCN).

The approach was evaluated on two datasets: NFBS and Infant T1. To manage data size and complexity, only 5 slices (slices 50-54) from each of the 125 NFBS subjects were used, totaling 625 2D images. The data was split for training (60%), testing (20%), and validation (20%), and the model was trained for 100 epochs using the Adam optimizer.

On the NFBS dataset, the GCN model achieved a sensitivity of 0.56%. When tested on 16 scans from the Infant T1 dataset (using slices 100-105), the model's performance metrics were a precision of 0.62, a Dice coefficient of 0.48, and an accuracy of 0.45.

The results for both datasets are promising but are based on a limited number of axial slices rather than the full MRI volumes, which is a key limitation of the evaluation.

Chapter 4

Demyelinating lesion segmentation

This chapter presents a comprehensive analysis of MS lesion segmentation, with a focus on building generalizable models. The study's main contributions include a systematic, cross-dataset evaluation of model performance and the public release of two new, rare pediatric MS datasets [6].

The researchers used the UNet++ architecture with a resnet18 encoder for its superior performance. They trained models on individual datasets and then tested them on others to assess their generalizability. The results showed consistent performance across different datasets, marking this as the most extensive cross-dataset analysis in the field to date. The best performance, a Dice score of 0.698, was achieved by a model trained and tested on specific subsets of the MSSEG2016 dataset, which outperformed even top state-of-the-art models [6].

A significant contribution is the creation of two public datasets from a pediatric hospital in Romania: PediMS (pediatric patients with MS) and PediDemi (pediatric patients with other demyelinating conditions). These datasets address a critical gap in publicly available data, as they represent a rare, underrepresented patient population. The PediDemi dataset, in particular, is unique because it includes non-MS demyelinating cases, providing a valuable resource for developing models that can differentiate between various pathologies [6].

The study also highlights two major challenges in achieving higher performance: insufficient dataset size and low-quality expert annotations. The authors note that the limited size of current MS datasets is a significant bottleneck. Furthermore, they found that even expert annotations have low inter-rater agreement (around 60-70%), suggesting that improving annotation quality is essential for developing more accurate models. The authors conclude that these limitations, rather than model architecture, are the main barriers to achieving higher segmentation scores [6].

Chapter 5

Lesion-wise metric

This chapter introduces the Dice spectrum, a new metric for evaluating lesion segmentation models. Unlike traditional metrics that provide a single, overall performance score, the Dice spectrum offers a more detailed view of a model's behavior by analyzing its performance on lesions of different sizes. This is crucial because lesions, particularly in conditions like Multiple Sclerosis, vary significantly in size [43].

The Dice spectrum works by grouping lesions into different size categories and calculating a separate Dice score for each group. This approach allows for a more precise understanding of a model's strengths and weaknesses, helping to determine its suitability for specific clinical applications like early lesion detection or monitoring larger lesions [43].

Evaluation of several models using the Dice spectrum on three public datasets (MSSEG-2016, ISBI-2015, and 3D-MR-MS) revealed a consistent trend: all models performed significantly better on larger lesions and very poorly on smaller ones. The average Dice scores for the smallest lesions ranged from 0.02 to 0.13, while intermediate lesions scored an average of 0.35. This finding highlights a major limitation of current segmentation models—they are ill-suited for detecting small lesions—and underscores the need for more sophisticated training methods. The Dice spectrum can serve as a valuable tool for evaluating these future models [43].

Conclusions and future work

This thesis contributes to advancing the field of neuroimaging, a critical tool for diagnosing and monitoring neurological disorders, by addressing two fundamental challenges: skull stripping and multiple sclerosis (MS) lesion segmentation. The primary goal was to enhance segmentation algorithms for brain MRI, improving their accuracy and generalizability for clinical applications.

Chapter 3 presented the proposed methods for skull stripping. One of the key contribution is the Sphere-GUBS, a novel skull stripping method based on spherical approximations. This approach provided a foundational framework for brain extraction, improving segmentation accuracy without the need for user intervention or preset parameters. Building on this, we developed the Ellipsoid-GUBS method, which further refined segmentation precision by using ellipsoidal models. By varying the center of the ellipsoid and testing it on multiple datasets, Ellipsoid-GUBS demonstrated competitive performance when compared to state-of-the-art methods such as BET2 and BSE, marking a considerable advancement over earlier spherical models. Furthermore, we pioneered the use of Graph Convolutional Networks (GCNs) for skull stripping, which, while showing potential, faced challenges due to the time-consuming nature of graph construction. The presented results were published in the following papers: [39, 42, 41, 40]. The work on skull stripping methods, particularly the improvements in segmentation precision through the use of ellipsoidal models, lays a strong foundation for future research. These approaches can be further optimized and integrated into clinical imaging workflows to improve neuroimaging outcomes. The promise of GCNs remains an exciting avenue for exploration; however, future work should focus on exploring the GCNs on more complex neurological problems, such as brain connectivity and EEG emotion recognition. Collaborating with clinical researchers will also be essential to validate the real-world applicability of these methods and to address any practical challenges encountered in clinical settings.

Chapter 4 focused on the MS lesion segmentation, namely addressing the challenge of generalizability across diverse datasets. By training a state-of-the-art UNet++ architecture on a wide range of datasets, we demonstrated that model performance improves with larger and more heterogeneous data. Additionally, we employed quantile normalization to mitigate scanner-related distribution shifts, which contributed to improved performance across datasets. This work emphasized the importance of dataset diversity and normalization techniques in developing segmentation models that can generalize to real-world clinical data. Standardizing axial slices further enhanced out-of-distribution performance, bringing these models closer to practical application. The results are presented in paper [6]. We also introduced two pediatric datasets [45, 44], one containing MS patients and one containing demyelinating lesions, for patients that have not been diagnosed with MS yet. By this we wanted to contribute to further developement of MS methods.

While progress has been made in improving generalizability, challenges remain, particularly in achieving the level of performance necessary for clinical applications. Current models tend to favor larger lesions, limiting their effectiveness in early detection of small lesions, which are crucial for early-stage MS diagnosis. Future work should prioritize collecting additional data from a broader range of scanners and patient cohorts, as well as addressing the variability in expert annotations to improve the quality and consistency of training datasets. Furthermore, exploring novel loss functions that explicitly target smaller lesions during training could significantly enhance the detection of early-stage MS lesions, making these models more suitable for clinical screening applications.

Chapter 5 introduced the Dice spectrum [43], a novel metric for assessing the performance of MS lesion segmentation models across different lesion sizes. The Dice spectrum offers a more detailed evaluation, particularly valuable in clinical scenarios where detecting small lesions is critical, such as in early-stage MS. By computing separate Dice scores for different lesion size categories, the Dice spectrum provides insights into model performance that traditional global metrics may overlook, particularly in detecting smaller, early-stage lesions. While the Dice spectrum represents a valuable advancement in evaluating lesion segmentation models, its application is currently limited to the evaluation stage. Future research should aim to integrate size-specific performance considerations into the training process itself. By incorporating metrics like the Dice spectrum into the model training pipeline, segmentation models can be better tailored to address the challenges posed by varying lesion sizes. Additionally, expanding the Dice spectrum to other performance metrics, such as Sensitivity and Specificity, could further enhance the evaluation and selection of models for specific clinical applications. Investigating the influence of different training methods and loss functions on size-specific performance will enable the development of more targeted and clinically effective segmentation models.

In conclusion, this thesis has introduced new methods and evaluation metrics that advance the state of neuroimaging, particularly in skull stripping and MS lesion segmentation. By addressing these critical challenges and proposing future directions for research, this work paves the way for more accurate and clinically applicable tools that could significantly enhance the diagnosis and treatment of neurological disorders, ultimately improving patient outcomes.

In addition to the future work directions outlined at the end of each chapter, several other research avenues are being considered. One direction involves optimizing the proposed unsupervised methods for skull stripping and integrating them into comprehensive neuroimaging preprocessing pipelines suitable for clinical use. Evaluating these methods on more diverse datasets-including pathological cases and non-adult populations-will be essential to assess their broader applicability.

Another important direction focuses on the use of Graph Convolutional Networks (GCNs), particularly in domains where connectivity is a key factor, such as functional MRI, diffusion tractography, or EEG-based emotion recognition. Future work will explore how GCNs can be embedded into tools intended for clinical applications.

In the domain of multiple sclerosis lesion segmentation, this thesis has emphasized the need for models that generalize across institutions, scanners, and patient cohorts. Future work should aim to build larger and more diverse annotated datasets. Additionally, segmentation research should expand beyond MS to include other demyelinating conditions such as ADEM. Segmenting spinal cord lesions is another critical direction, as patients diagnosed with MS who present brain lesions frequently also exhibit spinal cord involvement.

A further promising direction lies in the development of models tailored to detect small lesions, which are crucial for early-stage MS diagnosis. Future research could focus on custom loss functions that give greater weight to small lesions, and on progressive training strategies that guide the model from easier to more complex detection tasks. Incorporating temporal information about lesion evolution may also help improve sensitivity in early diagnosis.

Bibliography

[1] A decade of infant neuroimaging research: What have we learned and where are we going? *Infant Behavior and Development*, 58:101389, 2020.

[2] Dhammika Amaratunga and Javier Cabrera. Analysis of data from viral dna microchips. *Journal of the American Statistical Association*, 96:1161–1170, 2001.

[3] Syed Muhammad Anwar, Muhammad Majid, Adnan Qayyum, Muhammad Awais, Majdi Alnowami, and Muhammad Khurram Khan. Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, 42:1–13, 2018.

[4] Harshit Arora, Marco Mammi, Naisargi Manishkumar Patel, Dea Zyfi, Hema Reddy Dasari, Ismael Yunusa, Thomas Simjian, Timothy R Smith, and Rania A Mekary. Dexamethasone and overall survival and progression free survival in patients with newly diagnosed glioblastoma: a meta-analysis. *Journal of Neuro-Oncology*, 166(1):17–26, 2024.

[5] Harshit Arora, Marco Mammi, Naisargi Manishkumar Patel, Dea Zyfi, Hema Reddy Dasari, Ismael Yunusa, Thomas Simjian, Timothy R Smith, and Rania A Mekary. Dexamethasone and overall survival and progression free survival in patients with newly diagnosed glioblastoma: a meta-analysis. *Journal of Neuro-Oncology*, 166(1):17–26, 2024.

[6] Liviu Badea and Maria Popa. Toward generalizable multiple sclerosis lesion segmentation models. *IEEE Access*, 13:97859–97869, 2025.

[7] B.M. Bolstad, R.A Irizarry, M. Astrand, and T.P. Speed. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–193, 01 2003.

[8] Aaron Carass, Snehashis Roy, Amod Jog, Jennifer L. Cuzzocreo, Elizabeth Magrath, Adrian Gherman, Julia Button, James Nguyen, Pierre-Louis Bazin,

Peter A. Calabresi, Ciprian M. Crainiceanu, Lotta M. Ellingsen, Daniel S. Reich, Jerry L. Prince, and Dzung L. Pham. Longitudinal multiple sclerosis lesion segmentation data resource. *Data in Brief*, 12:346–350, 2017.

[9] Yuan Chang and Zheng Li. Attention-enhanced dual-stream registration network via mixed attention transformer and gated adaptive fusion. *Medical Image Analysis*, 105:103713, 2025.

[10] Olivier Commowick, Audrey Istace, Michaël Kain, Baptiste Laurent, Florent Leray, Mathieu Simon, Sorina Pop, Pascal Girard, Roxana Ameli, Jean-Christophe Ferré, Anne Kerbrat, Thomas Tourdias, Frederic Cervenansky, Tristan Glatard, Jeremy Beaumont, Senan Doyle, Florence Forbes, Jesse Knight, April Khademi, and Christian Barillot. Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Scientific Reports*, 8:13650–13666, 09 2018.

[11] PD Delgado-López and EM Corrales-García. Survival in glioblastoma: a review on the impact of treatment modalities. *Clinical and Translational Oncology*, 18(11):1062–1071, 2016.

[12] Lee Raymond Dice. Measures of the amount of ecologic association between species. *Ecology*, 26:297–302, 1945.

[13] Christina M Eckhardt, Sophia J Madjarova, Riley J Williams, Mattheu Ollivier, Jón Karlsson, Ayoosh Pareek, and Benedict U Nwachukwu. Unsupervised machine learning methods and emerging applications in healthcare. *Knee Surgery, Sports Traumatology, Arthroscopy*, 31(2):376–381, 2023.

[14] Daniel García-Lorenzo, Simon Francis, Sridar Narayanan, Douglas L. Arnold, and D. Louis Collins. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical Image Analysis*, 17(1):1–18, 2013.

[15] Zoubin Ghahramani. *Unsupervised Learning*, pages 72–112. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[16] Ian Goodfellow. Deep learning, 2016.

[17] Mayuri Gupta and Ashish Mishra. A systematic review of deep learning based image segmentation to detect polyp. *Artificial Intelligence Review*, 57(1):7, 2024.

[18] Maha Haki, Haeder A Al-Biati, Zahraa Salam Al-Tameemi, Inas Sami Ali, and Hany A Al-Hussaniy. Review of multiple sclerosis: Epidemiology, etiology, pathophysiology, and treatment. *Medicine*, 103(8):e37297, 2024.

[19] Hamirul'Aini Hambali, Nursuriati Jamil, Sharifah Lailee Syed Abdullah, and Hazaruddin Harun. A rule-based segmentation method for fruit images under natural illumination. In *2014 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*, pages 13–18, 2014.

[20] Jianxing He, Sally L Baxter, Jie Xu, Jiming Xu, Xingtao Zhou, and Kang Zhang. The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, 25(1):30–36, 2019.

[21] Markus Hofmarcher, Thomas Unterthiner, José Arjona-Medina, Günter Klambauer, Sepp Hochreiter, and Bernhard Nessler. Visual scene understanding for autonomous driving using semantic segmentation. *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, pages 285–296, 2019.

[22] Lei Hua, Yi Gu, Xiaoqing Gu, Jing Xue, and Tongguang Ni. A novel brain mri image segmentation method using an improved multi-view fuzzy c-means clustering algorithm. *Frontiers in Neuroscience*, 15:662674, 2021.

[23] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[24] Dejan Jakimovski, Stefan Bittner, Robert Zivadinov, Sarah A Morrow, Ralph HB Benedict, Frauke Zipp, and Bianca Weinstock-Guttman. Multiple sclerosis. *The Lancet*, 403(10422):183–202, 2024.

[25] Sandra Jardim, João António, and Carlos Mora. Image thresholding approaches for medical image segmentation-short literature review. *Procedia Computer Science*, 219:1485–1492, 2023.

[26] Xu Ji, João F. Henriques, and Andrea Vedaldi. Invariant information clustering for unsupervised image classification and segmentation, 2019.

[27] Baris Kayalibay, Grady Jensen, and Patrick van der Smagt. Cnn-based segmentation of medical imaging data. *arXiv preprint arXiv:1701.03056*, 2017.

[28] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017.

[29] Subbiahpillai Neelakantapillai Kumar, Alfred Lenin Fred, and Paul Sebastin Varghese. An overview of segmentation algorithms for the analysis of anomalies on medical images. *Journal of Intelligent Systems*, 29(1):612–625, 2019.

[30] Justin T Low, Quinn T Ostrom, Gino Cioffi, Corey Neff, Kristin A Waite, Carol Kruchko, and Jill S Barnholtz-Sloan. Primary brain and other central nervous system tumors in the United States (2014-2018): A summary of the CBTRUS statistical report for clinicians. *Neuro-Oncology Practice*, 9(3):165–182, 02 2022.

[31] Simeon Mayala, Ida Herdlevær, Jonas Bull Haugsøen, Shamundeeswari Anandan, Nello Blaser, Sonia Gavasso, and Morten Brun. Gubs: Graph-based unsupervised brain segmentation in mri images. *Journal of Imaging*, 8(10), 2022.

[32] Maciej A. Mazurowski, Haoyu Dong, Hanxue Gu, Jichen Yang, Nicholas Konz, and Yixin Zhang. Segment anything model for medical image analysis: An experimental study. *Medical Image Analysis*, 89:102918, 2023.

[33] Ibrahim Khalil Muhammad, Tehsin Samabia, Humayun Mamoona, N.Z Jhanjhi, and A. AlZain Mohammed. Multi-scale network for thoracic organs segmentation. *Computers, Materials & Continua*, 70(2):3251–3265, 2022.

[34] Dominik Müller and Frank Kramer. Miscnn: a framework for medical image segmentation with convolutional neural networks and deep learning. *BMC medical imaging*, 21:1–11, 2021.

[35] Javier Olazaran, Cristobal Carnero-Pardo, Juan Fortea, Pascual Sanchez-Juan, Guillermo Garcia-Ribas, Felix Vinuela, Pablo Martinez-Lage, and Merce Boada. Prevalence of treated patients with alzheimer,s disease: current trends and covid-19 impact. *Alzheimer,s Research & Therapy*, 15, 08 2023.

[36] Abbas Omidi, Aida Mohammadshahi, Neha Gianchandani, Regan King, Lara Leijser, and Roberto Souza. Unsupervised domain adaptation of mri skull-stripping trained on adult data to newborns. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 7718–7727, January 2024.

[37] Nikhil R Pal and Sankar K Pal. A review on image segmentation techniques. *Pattern Recognition*, 26(9):1277–1294, 1993.

[38] Linmin Pei, Murat Ak, Nourel Hoda M Tahon, Serafettin Zenkin, Safa Alkarawi, Abdallah Kamal, Mahir Yilmaz, Lingling Chen, Mehmet Er, Nursima

Ak, et al. A general skull stripping of multiparametric brain mris using 3d convolutional neural network. *Scientific Reports*, 12(1):10826, 2022.

[39] Maria Popa. An 3d mri unsupervised graph-based skull stripping algorithm. *Procedia Computer Science*, 225:1682–1690, 2023. 27th International Conference on Knowledge Based and Intelligent Information and Engineering Sytems (KES 2023).

[40] Maria Popa. Graph convolutional neural networks for skull stripping. In Wolfgang Osten, editor, *Seventeenth International Conference on Machine Vision (ICMV 2024)*, volume 13517, page 135170T. International Society for Optics and Photonics, SPIE, 2025.

[41] Maria Popa and Anca Andreica. Improving unsupervised graph-based skull stripping: Enhancements and comparative analysis with state-of-the-art methods. In Natalia Shakhovska, Jianbo Jiao, Ivan Izonin, and Stéphane Chrétien, editors, *Proceedings of the 7th International Conference on Informatics & Data-Driven Medicine, Birmingham, United Kingdom, November 14-16, 2024*, volume 3892 of *CEUR Workshop Proceedings*, pages 20–31. CEUR-WS.org, 2024.

[42] Maria Popa and Anca Andreica. Towards an improved unsupervised graph-based mri brain segmentation method. In Mohamed Sellami, Maria-Esther Vidal, Boudewijn van Dongen, Walid Gaaloul, and Hervé Panetto, editors, *Cooperative Information Systems*, pages 480–487, Cham, 2024. Springer Nature Switzerland.

[43] Maria Popa and Liviu Badea. Measuring spatial specificity of multiple sclerosis lesion segmentation using dice spectra. *Procedia Computer Science*, 246:3467–3474, 2024. 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024).

[44] Maria Popa and Gabriela Adriana Vișa. Pedidemi - a pediatric demyelinating lesion segmentation dataset. *MELBA Journal*, OpenData 2025 - MICCAI 2025, 2025. **Accepted**.

[45] Maria Popa, Gabriela Adriana Vișa, and Ciprian Radu Șofariu. Pedims: A pediatric multiple sclerosis lesion segmentation dataset. *Scientific Data*, 12(1):1184, Jul 2025.

[46] Benjamin Puccio, James P Pooley, John S Pellman, Elise C Taverna, and R Cameron Craddock. The preprocessed connectomes project repository of

manually corrected skull-stripped T1-weighted anatomical MRI data. *Giga-Science*, 5(1), 10 2016. s13742-016-0150-5.

[47] Bashar Rajoub. Chapter 3 - supervised and unsupervised learning. In Walid Zgallai, editor, *Biomedical Signal Processing and Artificial Intelligence in Healthcare*, Developments in Biomedical Engineering and Bioelectronics, pages 51–89. Academic Press, 2020.

[48] Meghavi Rana and Megha Bhushan. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia Tools and Applications*, 82(17):26731–26769, 2023.

[49] Md. Eshmam Rayed, S.M. Sajibul Islam, Sadia Islam Niha, Jamin Rahman Jim, Md Mohsin Kabir, and M.F. Mridha. Deep learning for medical image segmentation: State-of-the-art advancements and challenges. *Informatics in Medicine Unlocked*, 47:101504, 2024.

[50] Blaine Rister, Kaushik Shivakumar, Tomomi Nobashi, and Daniel L Rubin. Ct-org: Ct volumes with multiple organ segmentations [dataset]. *The Cancer Imaging Archive*, 21, 2019.

[51] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.

[52] Abhinav Sagar and RajKumar Soundrapandiyan. Semantic segmentation with multi scale spatial attention for self driving cars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2650–2656, 2021.

[53] Lars Schmarje, Monty Santarossa, Simon-Martin Schroder, and Reinhard Koch. A survey on semi-, self- and unsupervised learning for image classification. *IEEE Access*, 9:82146–82168, 2021.

[54] Martin Simonovsky and Nikos Komodakis. Dynamic edge-conditioned filters in convolutional neural networks on graphs. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3693–3702, 2017.

[55] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427–437, 2009.

[56] Danail Stoyanov, Zeike Taylor, Gustavo Carneiro, Tanveer Syeda-Mahmood, Anne Martel, Lena Maier-Hein, João Manuel RS Tavares, Andrew Bradley, João Paulo Papa, Vasileios Belagiannis, et al. *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, dlmia 2018, and 8th international workshop, ml-cds 2018, held in conjunction with miccai 2018, granada, spain, september 20, 2018, proceedings*, volume 11045. Springer, 2018.

[57] Rezq Muhammed Thabet, Dina Khattab, Howida A. Shedeed, and Maryam Al-Berry. The effects of diverse brain mri modalities on multiple sclerosis lesions segmentation: Comprehensive analysis. In *2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, pages 571–576, 2024.

[58] Robert T. Thibault. Conclusion. In Amir Raz and Robert T. Thibault, editors, *Casting Light on the Dark Side of Brain Imaging*, pages 159–160. Academic Press, 2019.

[59] Liping Wang, Ziming Zeng, and Reyer Zwiggelaar. An improved bet method for brain segmentation. In *2014 22nd International Conference on Pattern Recognition*, pages 3221–3226, 2014.

[60] Ruxin Wang, Shuyuan Chen, Chaojie Ji, Jianping Fan, and Ye Li. Boundary-aware context neural network for medical image segmentation. *Medical image analysis*, 78:102395, 2022.

[61] World Health Organization. Epilepsy: A public health imperative. `https://www.who.int/news-room/fact-sheets/detail/epilepsy`, 2019. Accessed: 2024-08-30.

[62] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, 2021.

[63] Kalidas Yeturu. Chapter 3 - machine learning algorithms, applications, and practices in data science. In Arni S.R. Srinivasa Rao and C.R. Rao, editors, *Principles and Methods for Data Science*, volume 43 of *Handbook of Statistics*, pages 81–206. Elsevier, 2020.

[64] Hongfeng You, Long Yu, Shengwei Tian, and Weiwei Cai. Dr-net: dual-rotation network with feature map enhancement for medical image segmentation. *Complex & Intelligent Systems*, pages 1–13, 2021.

[65] Nida M. Zaitoun and Musbah J. Aqel. Survey on image segmentation techniques. *Procedia Computer Science*, 65:797–806, 2015. International Conference on Communications, management, and Information technology (ICCMIT'2015).

[66] Lin Zhang, Yan Zhao, Tongtong Che, Shuyu Li, and Xiuying Wang. Graph neural networks for image-guided disease diagnosis: A review. *iRADIOLOGY*, 1(2):151–166, 2023.

[67] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. *CoRR*, abs/1807.10165, 2018.