

Article

GUBS: Graph-Based Unsupervised Brain Segmentation in MRI Images

Simeon Mayala ^{1,*}, Ida Herdlevær ^{2,3}, Jonas Bull Haugsøen ^{2,3}, Shamundeewari Anandan ^{2,3},
Nello Blaser ⁴, Sonia Gavasso ^{2,3} and Morten Brun ¹

¹ Department of Mathematics, University of Bergen, 5020 Bergen, Norway

² Department of Clinical Medicine, University of Bergen, 5021 Bergen, Norway

³ Neuro-SysMed, Department of Neurology, Haukeland University Hospital, 5053 Bergen, Norway

⁴ Department of Informatics, University of Bergen, 5020 Bergen, Norway

* Correspondence: simeon.mayala@uib.no

Abstract: Brain segmentation in magnetic resonance imaging (MRI) images is the process of isolating the brain from non-brain tissues to simplify the further analysis, such as detecting pathology or calculating volumes. This paper proposes a Graph-based Unsupervised Brain Segmentation (GUBS) that processes 3D MRI images and segments them into brain, non-brain tissues, and backgrounds. GUBS first constructs an adjacency graph from a preprocessed MRI image, weights it by the difference between voxel intensities, and computes its minimum spanning tree (MST). It then uses domain knowledge about the different regions of MRIs to sample representative points from the brain, non-brain, and background regions of the MRI image. The adjacency graph nodes corresponding to sampled points in each region are identified and used as the terminal nodes for paths connecting the regions in the MST. GUBS then computes a subgraph of the MST by first removing the longest edge of the path connecting the terminal nodes in the brain and other regions, followed by removing the longest edge of the path connecting non-brain and background regions. This process results in three labeled, connected components, whose labels are used to segment the brain, non-brain tissues, and the background. GUBS was tested by segmenting 3D T1 weighted MRI images from three publicly available data sets. GUBS shows comparable results to the state-of-the-art methods in terms of performance. However, many competing methods rely on having labeled data available for training. Labeling is a time-intensive and costly process, and a big advantage of GUBS is that it does not require labels.

Keywords: brain tissues; non-brain tissues; segmentation; minimum spanning tree



Citation: Mayala, S.; Herdlevær, I.; Haugsøen, J.B.; Anandan, S.; Blaser, N.; Gavasso, S.; Brun, M. GUBS: Graph-Based Unsupervised Brain Segmentation in MRI Images. *J. Imaging* **2022**, *8*, 262. <https://doi.org/10.3390/jimaging8100262>

Academic Editor: Caroline Petitjean

Received: 17 August 2022

Accepted: 22 September 2022

Published: 27 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The brain is a complex organ that makes the central nervous system together with the spinal cord. It is divided into forebrain (sensory processing, higher reasoning), midbrain (motor movement, audio/visual processing), and hindbrain (autonomic functions such as sleep and respiratory rhythms). Over the past years, non-invasive imaging techniques have gained momentum, in assessing brain injury and studying brain anatomy. In particular, magnetic resonance technology is widely used in the diagnosis of brain diseases such as brain tumors, multiple sclerosis, hematomas and to find the cause of conditions such as epilepsy and headaches [1]. The technology produces magnetic resonance imaging (MRI) data, which can be processed to produce 3D volumetric data with the intensity of voxels varying according to the properties of different tissues. MRI images are most commonly presented as a stack of two-dimensional slices. Analysis of such high quality complex MRI data is a tedious process. Recent advances with computer aided-tools have overcome the major pitfalls of manual MRI data analysis. Brain MRI segmentation is an important step in image processing as it highly influences the outcome of the entire analysis, which is

crucial in the case of surgical planning, delineating lesions and image-guided interventions. To segment any target structure in the brain, it is common to perform a preprocessing step to isolate the brain from non-brain tissues such as the skull, dura and scalp [2,3].

Methods for segmenting the brain from non-brain tissues can be classified as manual, semi-automated, and fully automated [4,5]. Manual brain extraction gives high precision but is labor-intensive [6]. Semi-automated methods involve a certain degree of user intervention, whereas automated methods do not depend on any human interaction. Most automated methods for brain extraction can be classified in categories, such as mathematical morphology-based, intensity-based, deformable surface-based, atlas-based, and hybrid methods [2,7,8]. Machine learning techniques, including neural networks are also widely used for skull stripping [9,10]. We review different skull stripping techniques before describing GUBS.

Louis Lemieux et al. [11] proposed a fully automated method for segmenting the whole brain in T1-weighted volume. It is fast and based on foreground thresholding and morphological operations. It performs 3D connected component analysis at each level. Furthermore, Brain Extraction Tool (BET) is an automated method for segmenting Magnetic Resonance head images to separate brain and non-brain tissues. It uses a deformable model that evolves to fit the brain's surface by applying a set of locally adaptive model forces. The method is robust, fast and it does not require any pre-registration or preprocessing [12].

FreeSurfer is an open-source software that implements various image processing tools for both structural and functional MRI data sets. Skull stripping is one of the tasks that FreeSurfer provides for processing MRI images. It performs automatic skull stripping from intensity normalized images through a deformation of a tessellated ellipsoidal template into the shape of the inner surface of the skull [13]. David W. Shattuck and Richard M. Leahy [14] present an MRI analysis tool that produces cortical surface representations with spherical topology from MR images of the human brain. One of the tools that include skull stripping is a brain surface extractor (BSE). It breaks connections between the brain and non-brain tissues by using a morphological erosion operation, and then it identifies the brain using a connected component operation followed by a dilation operation to undo the effects of the erosion [2]. The final step is closing small holes that may occur in the brain surface.

The work of [15] presents a skull-stripping algorithm based on a hybrid approach (HWA). It combines watershed algorithms and deformable surface models. The method uses a localized voxel in T1-weighted white matter to create a global minimum in the white matter. The global minimum is created before applying the watershed algorithm with a pre-flooding height. The Robust Brain Extraction (ROBEX) is another method for performing skull stripping. It is a learning-based brain extraction system that combines discriminative and generative models. It is trained to detect brain boundaries using the random forest approach and ensures that the result is plausible. To obtain the final segmentation, the imperfect shape presented by the model is refined by using a graph cut [16]. The work presented in [17] proposed an automatic skull stripping method based on deformable models and histogram analysis named Simplex Mesh and Histogram Analysis Skull Stripping (SMHASS) method. It defines the starting point for deformation using a rough segmentation based on thresholds and morphological operators. The deformable model is based on a simplex mesh, whereas its deformation is controlled by local gray levels of the image and the information obtained on the gray level modeling of the rough-segmentation.

A simple skull stripping algorithm (S3) is proposed in [18]. The method is based on brain anatomy and image intensity characteristics. It is a knowledge-based algorithm and works by using adaptive intensity thresholding and then morphological operations. Oeslle Lucena et al. [9] proposed a Convolutional Neuron Network (CNN) for brain extraction in MRI trained with "silver standard" masks. The method generates silver standard masks which are used as inputs by using the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm and then implementing a tri-planar method using parallel 2D U-Net-based CNNs, named as CONSNet. Anam Fatima et al. [19] proposed a skull

stripping method and evaluated it on a 2D slice-based multi-view U-Net (MVU-Net) architecture. It performs as well as a 3D model while using fewer computational resources. In [20], a Single-Input Multi-Output U-Net (SIMOU-Net) was developed for segmenting foetal brain. Different from the original U-Net, the SIMOU-Net has a deeper architecture and takes account of the features extracted from each side output. Furthermore, Ref. [6] proposed a 3D-UNet for skull stripping to address the entire brain extraction problem satisfactory for diverse datasets. The work in [21] proposed a graph-based method for skull stripping, which uses intensity thresholding on the input image to obtain a preliminary mask. Finally, the method removes narrow connections using graph cuts followed by post-processing.

In this paper, we propose a graph-based approach that represents an MRI image using a graph, before segmenting the brain. The method uses the minimum spanning tree (MST) of the constructed adjacency graph to separate the brain, non-brain tissues, and the background. The approach involves the following steps; preprocessing of the MRI image, sampling points within and outside the brain, MST construction from the graph, isolating brain, non-brain tissues, and the background, followed by post-processing. The general idea of using a graph-based method for skull stripping relates to the ideas of [21]. We define an image into adjacency graph in the same way, but our approach differs significantly regarding the segmentation criteria. We perform skull stripping using the MST of the graph, in which each vertex has a minimal number of connections instead of using the whole graph. Different from our approach, the method presented by [21] depends on an estimate of the region bounded by white matter obtained by region growth. We define edge weights differently, and our approach does not depend on region growth. The method presented by [21] uses an initial mask obtained by thresholding and uses a graph cut to disconnect the narrow connections. GUBS depends on the MST constructed from the adjacency graph and separates the brain, non-brain tissues, and the background by disconnecting paths connecting these regions.

2. Materials and Methods

2.1. Materials

In this paper, we analyzed three different data sets. The first data set is from the Open Access Series of Imaging Studies (OASIS). It consist of T1W images from 77 subjects with an isotropic voxel resolution of 1.0 mm and a shape of $176 \times 208 \times 176$ [22]. The ground truth segmentation provided with this data set was created using a custom method based on registration to an atlas and then revised by human experts. Twenty of the subjects in this dataset suffered from Alzheimer's disease, a degenerative disease characterized by the loss of brain tissue [9,16].

The second data set consists of 20 simulated T1W MRI images mimicking healthy brains collected from the BrainWeb (BW) database [23,24]. They are provided as *.MINC* format anatomical models consisting of a set of 3D tissue membership volumes, one for each tissue class. Each label at a voxel in the anatomical model represents the tissue that contributes the most to that voxel. They have dimensions of $362 \times 434 \times 362$ and have a 0.5 mm isotropic voxel size.

The third data set consists of T1W MRI images from 18 subjects provided to the Internet Brain Segmentation Repository (IBSR) by the Center for Morphometric Analysis at Massachusetts General Hospital [25]. The images are stored in the NIfTI format. The shape of the images is $256 \times 256 \times 128$ and the voxel resolution is $0.94 \text{ mm} \times 0.94 \text{ mm} \times 1.5 \text{ mm}$ [26]. The database contains a manual segmentation of the gray matter (GM), white matter (WM) and cerebral spinal fluid (CSF), as well as skull stripped images.

Figure 1 shows coronal sections of brains sampled from a 3D MRI of a representative subject from each of the three data sets.

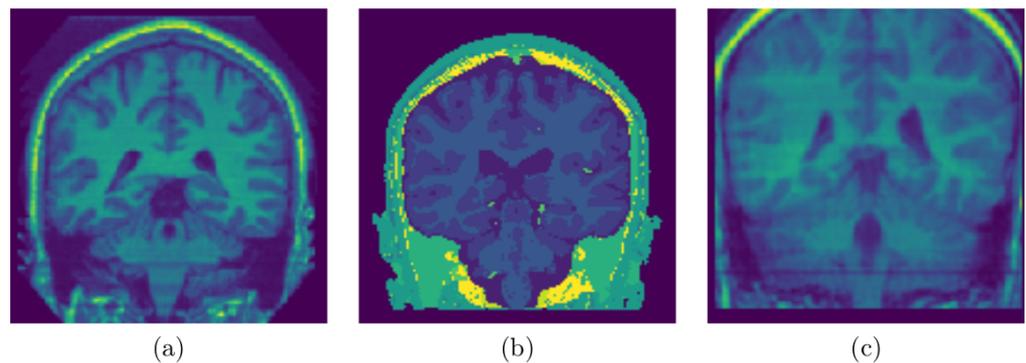


Figure 1. Sampled Slices: (a) Sampled slice coronal section from a 3D MRI volume OASIS data set, (b) Sampled slice coronal section from a 3D MRI volume BW data set, (c) Sampled slice coronal section from a 3D MRI volume IBSR data set. Notice the differences and the quality of data sets.

2.2. Methods

In this section, a graph-based method for brain segmentation using MST is presented. An adjacency graph is constructed from a preprocessed 3D MRI and then its MST is applied for separating the brain, non-brain tissues, and the background. The summarized steps for brain segmentation are presented in Figure 2.

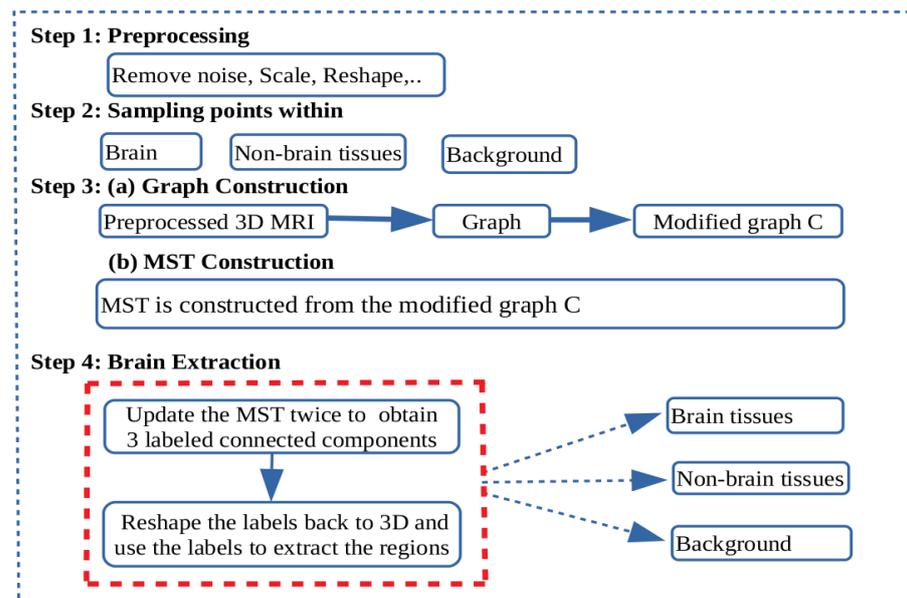


Figure 2. Schematic diagram: Flow diagram showing the steps for brain extraction. Step 1: Pre-processing to remove noise, scale values in the range of 0 and 1 and reshape the 3D magnetic resonance imaging (MRI) volume. Step 2: Sampling points within the brain, non-brain tissues, and the background. Step 3: An adjacency graph weighted by absolute intensity differences is constructed from the preprocessed 3D MRI volume. Then, nodes in the adjacency graph corresponding to the sampled points in step 2 are collapsed in their respective regions to form a graph C. From the modified graph C, a minimum spanning tree is constructed. Step 4: Brain segmentation. Nodes in C representing each of the regions of interest are the terminal nodes for the paths to be disconnected to separate the regions. First, the minimum spanning tree (MST) is modified by removing the edge with highest edge weight in the path connecting the representative nodes to separate the brain and non-brain subtrees. Again the MST is modified by removing the edge with highest edge weight in the path connecting the representative nodes to separate the non-brain and background subtrees. New labels are assigned and reshaped back to the shape of the 3D MRI.

Let A be an MRI image with dimension $M \times N \times L$ such that $A_{(i,j,k)}$ gives the intensity value at position (i, j, k) for $i = 0, 1, \dots, M - 1; j = 0, 1, \dots, N - 1$ and $k = 0, 1, \dots, L - 1$. Voxels at position (i, j, k) and (i', j', k') are called adjacent if $(i - i')^2 + (j - j')^2 + (k - k')^2 = 1$. We can mathematically define

$$A = [A_{(i,j,k)}] \in \mathbb{R}^{M \times N \times L} \tag{1}$$

and each point changes along the coordinate axes with respect to the position of the adjacent voxels.

2.2.1. Preprocessing

We read all the image data sets using a *NiBabel* package that can read different common medical images file formats [27] and then retrieve a 3D image using a *nilearn* package [28]. Finally, every MRI image is convert into *.npy* using a *numpy* package [29]. Each image was resized to the size of $128 \times 128 \times 128$. The intensities were scaled in the range of 0 to 1 using

$$A_{(i,j,k)}^{scaled} = \frac{A_{(i,j,k)} - A_{(j,k)}^{min}}{A_{(j,k)}^{max} - A_{(j,k)}^{min}}, \tag{2}$$

where $A_{(j,k)}^{min} = \min_{i \in \{0, \dots, M-1\}} \{A_{(i,j,k)}\}$ and $A_{(j,k)}^{max} = \max_{i \in \{0, \dots, M-1\}} \{A_{(i,j,k)}\}$. We adapt the formula and its implementation from *scikit-learn* [28]. Since the implementation requires only number of samples and number of features, the 3D MRI are reshaped into 2D before scaling and then reshape back to 3D after scaling. Whenever $A_{(j,k)}^{max}$ is approximately or equal to $A_{(j,k)}^{min}$ the situation is taken care by setting scales of near constant features to 1 to avoid division by very small number or zero values.

We remove noise and make sure that there is no spot in the background (outside the head) by removing small objects. Some data sets may need additional preprocessing steps to accelerate the separation of the regions. Since the brain is connected to the rest of the head by dark, thin segments [11], a small threshold may be applied to disconnect more the brain from the non-brain tissues. This is done by setting all values less than the threshold to 0. For IBSR and OASIS data sets, the values 0.25 and 0.32 were used in the preprocessing step, respectively. For data sets with good contrast between the brain and non-brain tissues such as BW, thresholding is unnecessary.

2.2.2. Edge Surface Detection

The edge surfaces in 3D images are defined as the structural boundaries of objects in the image. However, the true edge-surfaces of 3D images are continuous surfaces rather than discrete 3D edges like points which are detected by edge detectors [30]. These are points sampled at the surface that form the boundary between the object and the background. In this context the surface of the object is represented by intensity changes in the data volume [31]. The changes will be detected between edges.

The edges are detected by highlighting the local variation between the adjacent voxels. Example of the two adjacent voxels (i, j, k) and (i', j', k') located on the opposite side of an edge. We use the finite difference method (particularly forward difference) to compute the change between two adjacent voxels from the binary image. We define B to be the binary image of A . The change is computed by

$$\varepsilon = B_{(i',j',k')} - B_{(i,j,k)}. \tag{3}$$

For a 3D binary image the edge is detected when the changes between two neighboring voxels are

$$\varepsilon = \begin{cases} +1 & \text{if } B_{(i',j',k')} = 1 \text{ and } B_{(i,j,k)} = 0 \\ -1 & \text{if } B_{(i',j',k')} = 0 \text{ and } B_{(i,j,k)} = 1 \end{cases} \tag{4}$$

So, when ε is +1 or -1 the forward difference method asserts the presence of an edge. The edge surfaces of the 3D image consist of all points satisfying the changes in Equation (4) along a given axis. Then, the discrete 3D edge-like points represents the set of the edge points detected from 3D binary image B by satisfying the condition ε for the neighboring voxels.

In the implementation we perform two steps to obtain the discrete 3D edge-like points. Image A is binarized to obtain image B by either setting voxels to 1 if the gray value is greater than zero or setting to 0 elsewhere. We fill all possible holes in the binary image B by using a function for binary closing from Scipy package [32]. Then, we use a *diff* function from numpy package [29] to compute the discrete difference of the binary image by using the forward difference formula. Then, the points satisfying condition ε are discrete 3D edge-like points distributed at the surface of the head. These points are very useful in the next subsection when sampling points within the brain and non-brain tissues.

2.2.3. Sampling Points within the Brain, Non-Brain Tissues and Background

Let T be the set of points in the MRI image A , i.e.,

$$T = \{(i, j, k) : 0 \leq i \leq M - 1, 0 \leq j \leq N - 1, 0 \leq k \leq L - 1\}. \tag{5}$$

Next, we will sample points T_B , T_{NB} and T_{BG} in the brain, non-brain tissues and background, respectively. We define, $T_B \subset T$, $T_{NB} \subset T$ and $T_{BG} \subset T$ such that T_B , T_{NB} and T_{BG} are mutually exclusive but not exhaustive. That is,

$$T_B \cap T_{NB} = T_{NB} \cap T_{BG} = T_B \cap T_{BG} = \emptyset \tag{6}$$

and

$$T_B \cup T_{NB} \cup T_{BG} \subsetneq T. \tag{7}$$

We do the following things to obtain points in T_B : From the discrete 3D edge-like points distributed at the surface of the head we take points above the selection line which is obtained by visual inspection (see Figure 3). Then, pull these points from the surface of the head to the brain by a certain distance $\delta_1 \in \mathbb{N}$. Then, compute the mean point of the pulled points. For each pulled point a distance is computed from the mean point and get the maximum distance. Then, the maximum value times 0.75 is the a distance threshold value for removing the points with the distance above the threshold value. This last step is performed to restrict the pulled points to lie within the brain and remove points close to the skull.

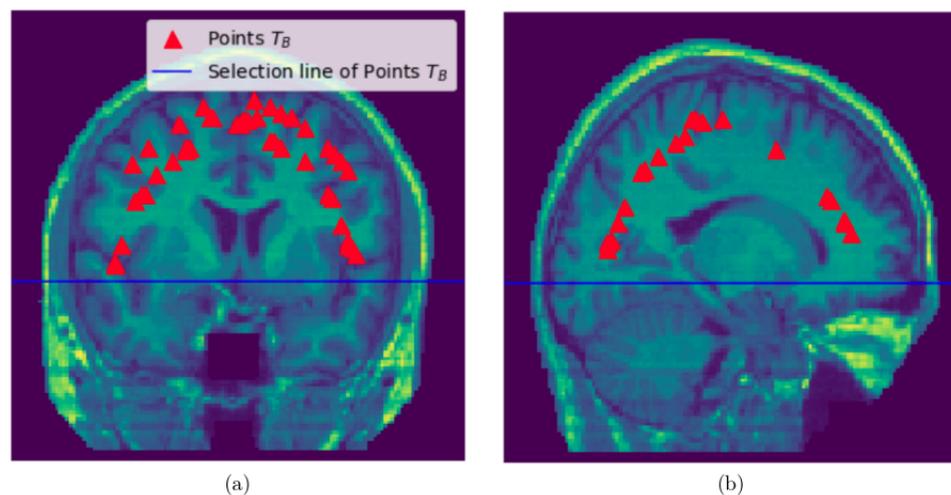


Figure 3. Sampling points T_B within the brain: (a) Sampled slice coronal section showing the selection line and points sampled within the brain, (b) Sampled slice sagittal section section showing the selection line and points sampled within the brain.

To obtain points T_{NB} in the non-brain tissues we use all the discrete 3D edge-like points distributed at the surface of the head and pull them inside towards the non-brain tissue by $\delta_2 \in \mathbb{N}$. The pulled points are rechecked if there is any point with the intensity value of zero it is removed. See Figures 4 and 5.

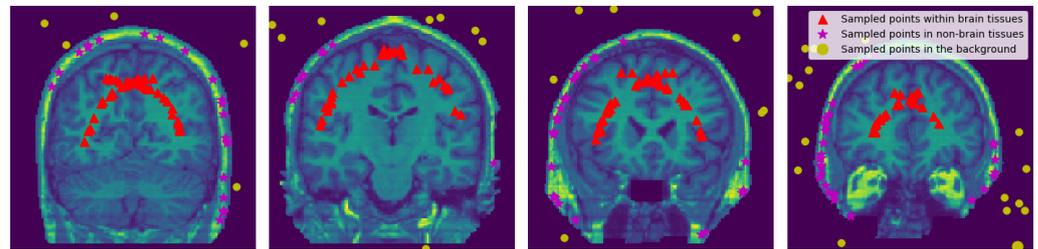


Figure 4. Sampled Points: 2D visualization of the representative coronal section from MRI image of a single subject (OASIS data) showing the sampled points within the brain, within the non-brain, and in the background region. δ_1 and δ_2 values were set to 15 and 3, respectively.

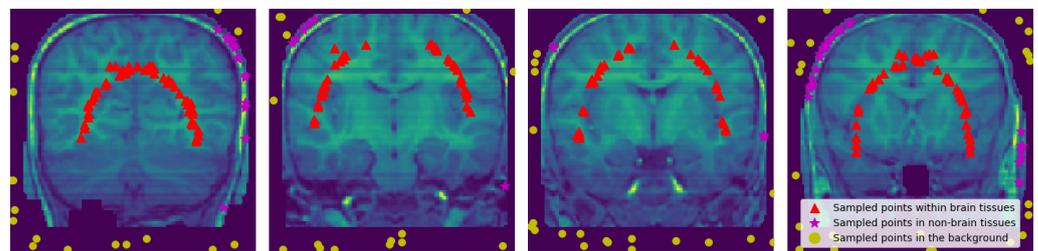


Figure 5. Sampled Points: 2D visualization of the representative coronal section from MRI image of a single subject (IBSR data) showing the sampled points within the brain region, within non-brain tissues, and in the background. δ_1 and δ_2 values were set to 15 and 3, respectively. Notice the removed part of the skull and the brain from some slices.

To obtain points T_{BG} sampled in the background we use the binary image after closing all the possible holes (this is the same binary image described in the last paragraph of Section 2.2.2).

First, we compute all points in the binary image whose points have an intensity values of 0. We also find points at the six surfaces of the 3D cube of the binary image whose points have an intensity values of 0. We combine these points to increase the chance of getting representative points from all the sides. We then sample 20,000 of these points uniformly at random to get the T_{BG} points.

2.2.4. Graph

Let (V, E) be the voxel adjacency graph constructed from A such that V is the set of nodes and E represents a set of weighted edges. Each node represents a voxel at location (i, j, k) in A , and each edge connects two adjacent voxels. The edge's weight is the absolute value of the difference between the intensity values of the voxels it connects.

2.2.5. Collapsing Nodes

Given a graph G and a subset H of the nodes in G we construct a graph where all the nodes in H have been collapsed to a single node. In order to ease the construction of this graph we represent the single collapsed node by a node $h \in H$.

The graph G_H^h where all nodes in H have been collapsed to a single node h in H is constructed as follows:

1. For every edge e in G , if both end nodes of e appear in H the edge e is discarded.
2. For every remaining edge e , if a node v in H appears in e , the edge e is modified by replacing v by h .
3. Remove all the nodes in H except the node h from the modified graph to obtain the graph G_H^h .

2.2.6. Segmentation Criteria

The central problem addressed by GUBS is to segment an image using samples of nodes from the regions of interest. From the voxels adjacency graph, we construct a new graph by successively collapsing the nodes sampled from each of the regions. The main idea of GUBS is to use the minimum spanning tree of this new graph instead of the minimum spanning tree of the voxels adjacency graph.

Given an MRI image of the brain, we want to separate the voxel adjacency graph into three regions representing the brain, non-brain tissues, and background.

Let $V_B \subset V$, $V_{NB} \subset V$ and $V_{BG} \subset V$ be sets of nodes in a graph G corresponding to sets of coordinate points T_B , T_{NB} and T_{BG} sampled from brain, non-brain tissues, and background, respectively. We construct the graphs G_B , G_{NB} and G_{BG} so that the nodes V_B , V_{NB} and V_{BG} are in G_B , G_{NB} and G_{BG} , respectively.

Let $v_B \in V_B$, $v_{NB} \in V_{NB}$ and $v_{BG} \in V_{BG}$ be randomly sampled nodes. Let $A = G_{V_B}^{v_B}$, $B = A_{V_{NB}}^{v_{NB}}$, $C = B_{V_{BG}}^{v_{BG}}$. In the modified graph C each of the sets V_B , V_{NB} and V_{BG} have been collapsed to a single node.

The minimum spanning tree (MST) of the graph C is the spanning tree with the least total edge weight among all possible spanning trees of the adjacency graph [33].

A path P in MST is a sequence of nodes in which each pair of consecutive nodes are connected by an edge. Note that since MST is a tree every pair of nodes is connected by a unique path.

The MST is constructed from the modified graph C and the components representing brain, non-brain tissues and the background are extracted as follows: First, we modify the MST by removing the edge with highest edge weight in the path connecting the nodes v_B and v_{NB} to separate the brain (G_B) and non-brain (G_{NB}) subtrees. Next, we separate the non-brain region and the background by removing the edge with highest edge weight in the path connecting the nodes v_{NB} and v_{BG} . This gives us the connected components of the tree with three different labels. Finally, the labels are reshaped back to the shape of the input MRI image.

3. Results

For visualization purposes, we perform a padding of background slices on the MRI image. The visualization of the results in this paper are performed using Matplotlib [34] and Seaborn [35]. The segmented 3D brain from OASIS, BW and IBSR data sets are visualized in Figures 6–8, respectively.

Figures 6–8 represent visual comparison of the extracted 3D brain using the GUBS approach to the gold standard segmented brain. The general quality of the predicted brain is sufficient, except for small variation in finer details. The small variation can be attributed to different reasons such as the quality of the data (see Figure 1), differences when preparing the gold standard segmentation, the patients conditions, such as disease or aging, and possibly parameter tuning for certain data.

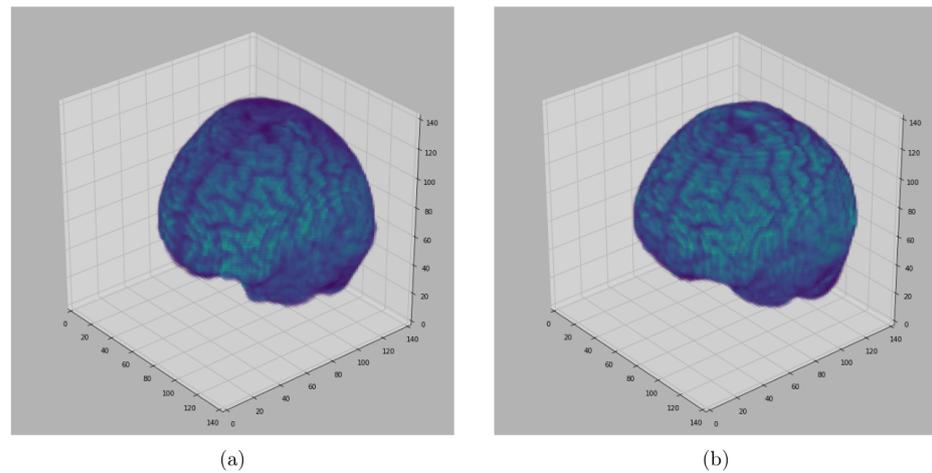


Figure 6. Segmented Brain (OASIS data): One representative subject representing (a) 3D brain segmented using GUBS approach (predicted), (b) 3D brain (ground truth). The masks were segmented using a custom method based on registration to an atlas, and then revised by human experts.

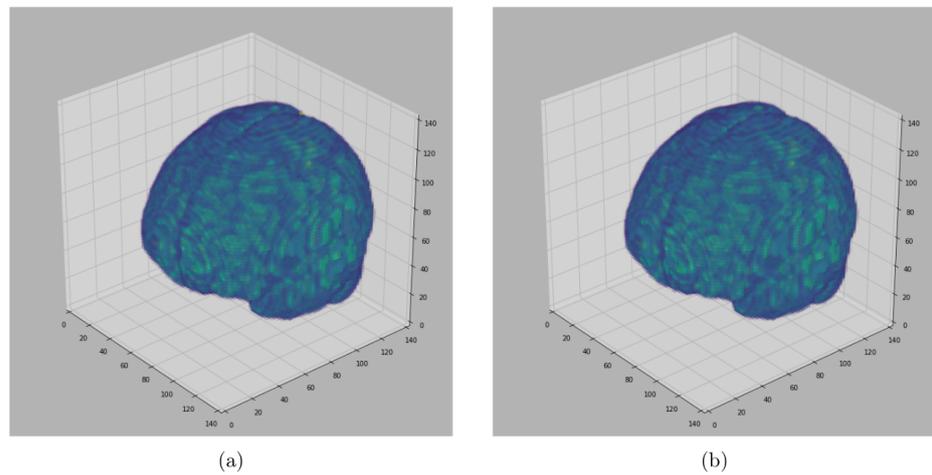


Figure 7. Segmented Brain (BW data): One representative subject representing (a) 3D brain segmented using GUBS approach (predicted), (b) 3D brain (ground truth). The ground truth was obtained from the labels representing cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM).

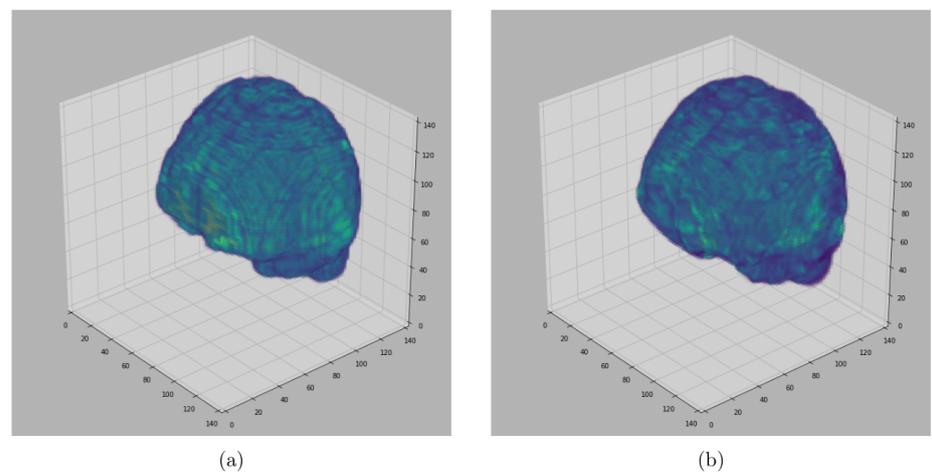


Figure 8. Segmented Brain (IBSR data): One representative subject representing (a) 3D brain segmented using GUBS approach (predicted), (b) 3D brain (ground truth). The ground truth was obtained by manual-guided expert segmentation.

Figures 9 and 10 present examples of a visual comparison of 2D slices segmented from two different subjects from IBSR and OASIS data sets, respectively. Notice that for Figure 9, the original sagittal sections show that some parts of the 3D head were cut. Figure 10 also present representative sagittal sections from a single subject. Since the proposed approach depends on sampling points within different regions, tuning the parameters for extracting the brain from the two compared MRI volumes (Figures 9 and 10) can differ significantly.

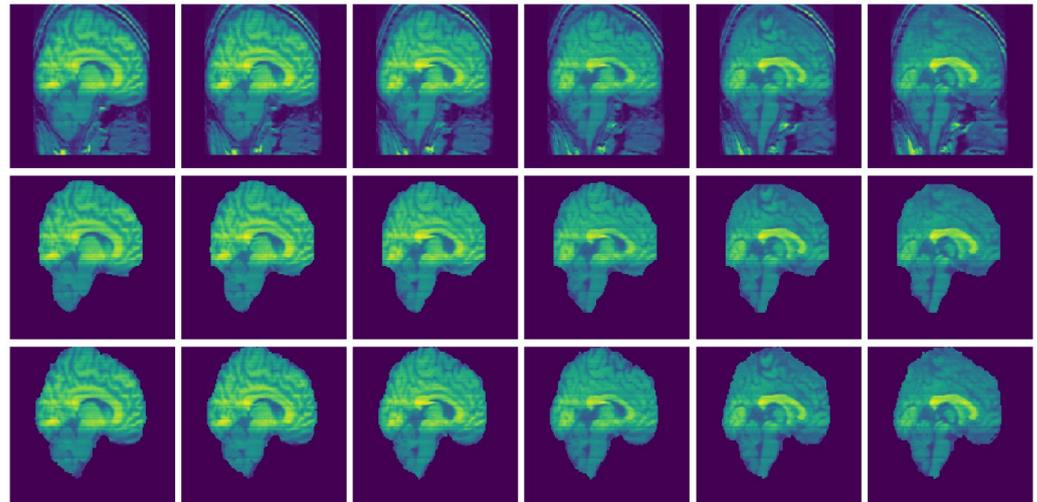


Figure 9. Selected MRI slices (IBSR data set): Sagittal MRI plane segmented brain. Row one: Input images, Row two: Predicted brain and Row three: Ground truth brain.

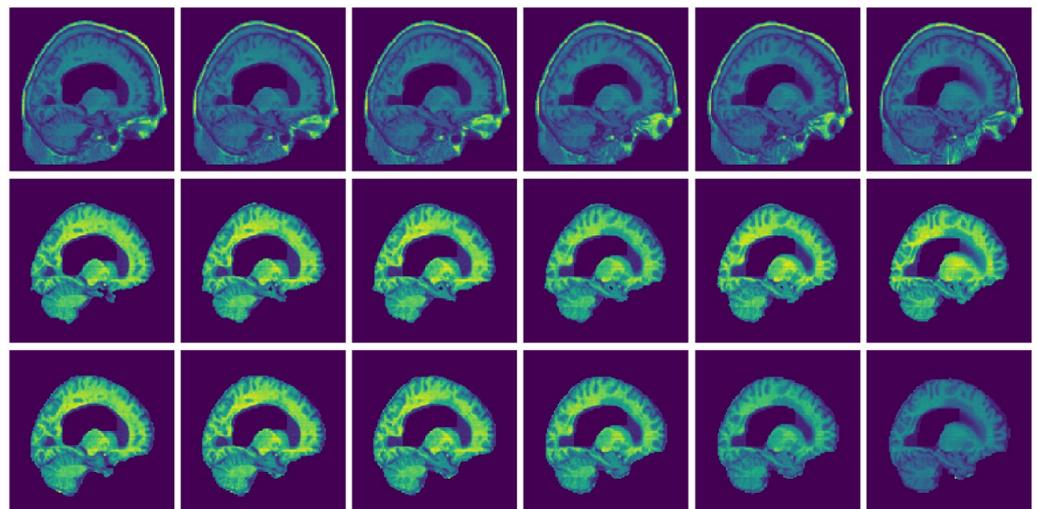


Figure 10. Selected MRI slices (OASIS data set): Sagittal MRI plane segmented brain. Row one: Input images, Row two: Predicted brain and Row three: Ground truth brain.

3.1. Performance Analysis

In this subsection, we present a quantitative analysis using Jaccard Index (JI), Dice Similarity Coefficient (DSC), sensitivity, specificity, accuracy and precision to assess the performance of GUBS. We evaluate the performance by computing the listed measures of similarity between the predicted and the ground truth of the MRI image of the binary labels.

The voxels in the binary labels that are correctly classified as brain are represented as true positive (TP). The voxels that are incorrectly classified as brain are represented as false positive (FP). The voxels that are correctly classified as non-brain tissues are represented by true negative (TN). The voxels that are in the brain region but incorrectly classified

as non-brain tissues are represented by false negative (FN). The Jaccard Index and Dice Similarity Coefficient are computed by

$$JI = \frac{TP}{TP + FP + FN}, DSC = \frac{2TP}{2TP + FP + FN}. \tag{8}$$

The sensitivity and specificity which show the percentage of brain and non-brain voxels are computed by

$$Sensitivity = \frac{TP}{TP + FN}, Specificity = \frac{TN}{TN + FP}. \tag{9}$$

The precision gives the ratio of the correctly positive identified labels against all the labels, whereas accuracy gives the ratio of the correctly identified.

$$Precision = \frac{TP}{TP + FP}, Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{10}$$

The values for the measures of similarity are presented in the range of 0 to 1. A measure of similarity with 0 value shows that there is no overlap between the predicted brain and the ground truth. The measure of similarity with 1 indicates that there is a perfect overlap between ground truth and the predicted brain segmentation.

The quantitative evaluation of GUBS compared to the state of the art method is presented in Tables 1–3 for OASIS, BW and IBSR data sets, respectively. The general performance of GUBS is satisfactory. For some data sets GUBS does not outperform other state of the art methods, but the results obtained by GUBS is consistent with previous results. Looking at the performance from the individual data sets GUBS does not perform well on sensitivity for OASIS and IBSR, but it performs well on specificity from all the three data sets.

Table 1. Performance analysis of GUBS compared to STAPLE and CONSNets [9] methods based on the OASIS data set by presenting the average in each measure of similarity.

	JI (mean ± sd)	DSC (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
STAPLE	-	0.960960 ± 0.0070	0.989830 ± 0.0060	0.951880 ± 0.0200
CONSNets	-	0.955480 ± 0.0100	0.990550 ± 0.0060	0.939800 ± 0.0280
GUBS	0.872633 ± 0.0148	0.931918 ± 0.0084	0.937179 ± 0.0256	0.974223 ± 0.0101

Table 2. Performance analysis of GUBS compared to BSE, HWA and SMHASS [17] methods based on the BW data by presenting the average in each measure of similarity.

	JI (mean ± sd)	DSC (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
BSE	0.875000 ± 0.0490	0.932000 ± 0.0310	0.991000 ± 0.0040	0.979000 ± 0.0120
HWA	0.685000 ± 0.0170	0.813000 ± 0.0120	1.000000 ± 0.0010	0.928000 ± 0.0050
SMHASS	0.904000 ± 0.0110	0.950000 ± 0.0060	0.990000 ± 0.0030	0.985000 ± 0.0020
GUBS	0.982396 ± 0.0271	0.990927 ± 0.0141	0.984012 ± 0.0268	0.999356 ± 0.0005

Table 3. Performance analysis of GUBS compared to HWA, SMHASS [17] and multi-view U-Net (MVU-Net) [19] methods based on the IBSR data set by presenting the average in each measure of similarity.

	JI (mean ± sd)	DSC (mean ± sd)	Sensitivity (mean ± sd)	Specificity (mean ± sd)
HWA	0.814000 ± 0.0360	0.897000 ± 0.0220	1.000000 ± 0.0000	0.966000 ± 0.0120
SMHASS	0.905000 ± 0.0300	0.950000 ± 0.0170	0.992000 ± 0.0100	0.985000 ± 0.0090
MVU-Net	-	0.908100	0.941400	0.989400
GUBS	0.859300 ± 0.0176	0.924229 ± 0.0102	0.918936 ± 0.0334	0.980869 ± 0.0104

Consistency Analysis of the GUBS across Different Data Sets

We have segmented three different data sets from OASIS, BW, and IBSR. They are significantly different in terms of quality and quantity. The OASIS data sets were collected from participants of different age groups, healthy and unhealthy subjects suffering from dementia. The data sets from BW are normal simulated brains whereas the IBSR data sets were collected from healthy subjects. To better understand the performance of the GUBS for differences in brain shapes across ages, possibly diseased and non-diseased brains, we combined all the results to explore the relationship between different measures of similarity. Figure 11 presents plots in matrix format showing the relationship between different measures of similarity on the three data sets. The diagonal subplots show the distributions for a specific measure of similarity in the individual data sets.

From Figure 11, the distribution in the diagonal show that the accuracy, DSC, and sensitivity from the OASIS data set have a lower deviation from the mean value compared to BW and IBSR data sets. It also indicates that specificity and precision from BW data have a lower deviation from the mean compared to the specificity and precision from OASIS and IBSR. Furthermore, Figure 11 shows that BW has the highest mean values for all measures of similarity compared to OASIS and IBSR data sets. Measures of similarity from IBSR show a higher deviation from the mean values compared to the measures from other data sets.

The pair plots show that the DSC are positively correlated with the accuracies obtained from all the data sets. The sensitivity obtained from BW and IBSR is positively correlated with accuracy and DSC. Specificity and precision obtained from BW show that they are not correlated with other measures of similarity whereas specificity and precision obtained from IBSR and OASIS are negatively correlated with sensitivity obtained from these data sets. For all data sets the obtained precision is positively correlated with the specificity.

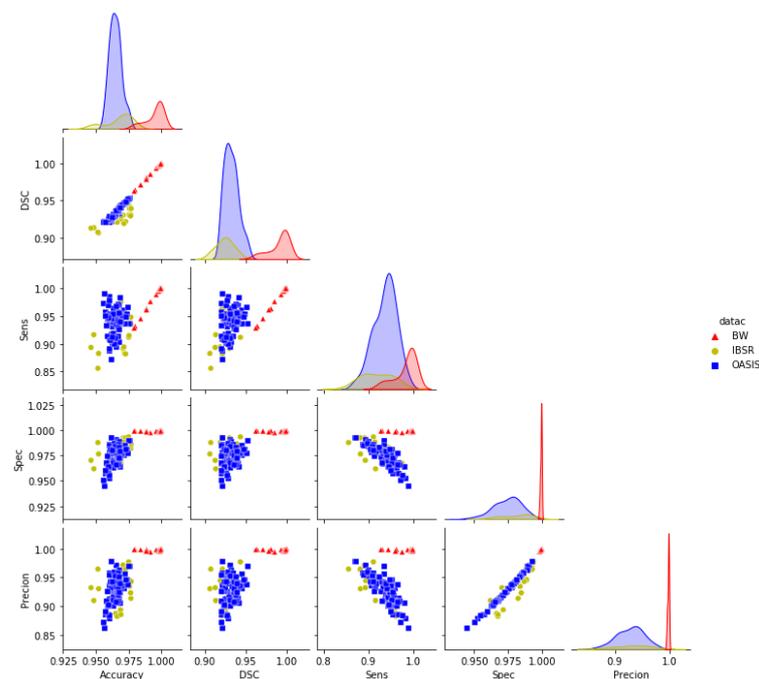


Figure 11. Pair plots for the measures of similarity (Combined data sets): Pair plots showing the pairwise relationship between different measures of similarity for the results obtained using GUBS across the combined data sets. DSC = Dice Similarity coefficients, Sens = Sensitivity, Spec = Specificity.

Figure 12d shows a separation between the measures with and without outliers after combining the results from the three data sets. It shows many outliers compared to accuracy and DSC. Based on the results presented in Tables 1–3 it is possible that the noted outliers come from BW results. Furthermore, we note that sensitivity, specificity, and precision

do not have outliers even though specificity shows high performance for all the data sets compared to sensitivity and precision.

Looking at the size of the box plots we note that the measures of accuracy, DSCs, and specificity in their distribution show that the lower and upper quartile are close to each other compared to other measures.

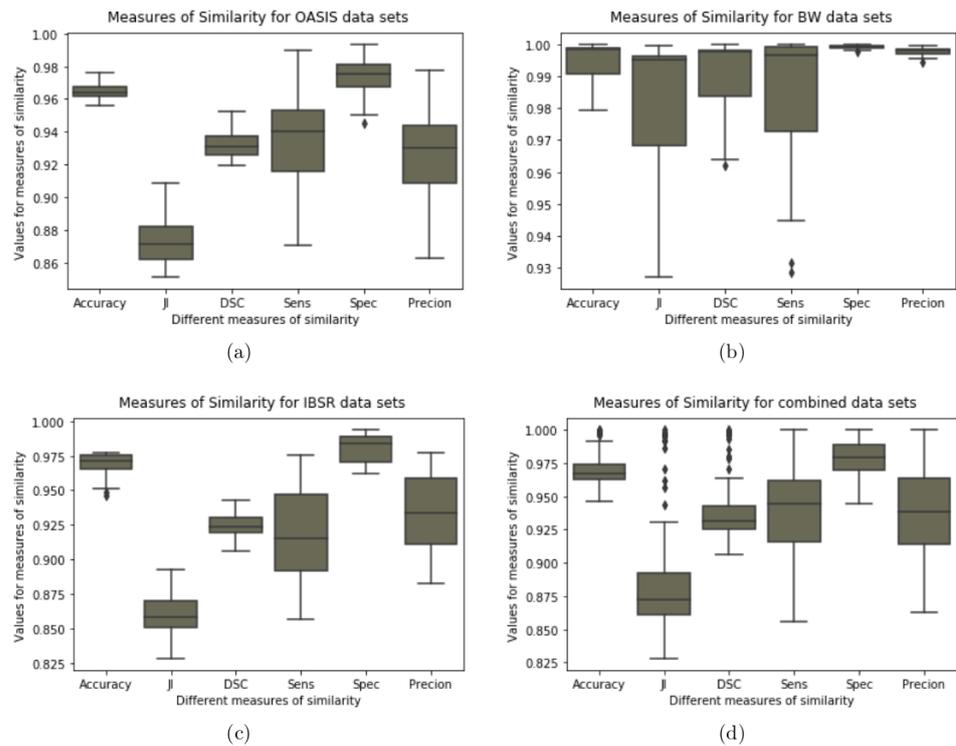


Figure 12. Boxplot for the measures of similarity (Combined Data sets): Boxplot showing variability for the measures of similarity for the results obtained using GUBS method across the combined data sets. JI = Jaccard Indices, DSC = Dice Similarity coefficients, Sens = Sensitivity, Spec = Specificity.

3.2. Parameter Selection

Due to high variation within and across image data sets, it is challenging to get a single set of parameters that works to produce the best possible segmentation results for different image data sets. In this work, parameter tuning was initially done on trial and error bases to obtain a value that could achieve good results. For each data set, the GUBS is run by testing different parameter values and choosing the parameters which lead to the separation of the tree into the separated components. Figures 13 and 14 show experimentation for different threshold parameters with same nodes sample size of 20,000, where Figures 15 and 16 show experimentation for different nodes sample sizes and same threshold. Note that BW data (see Figure 16) does not need threshold.

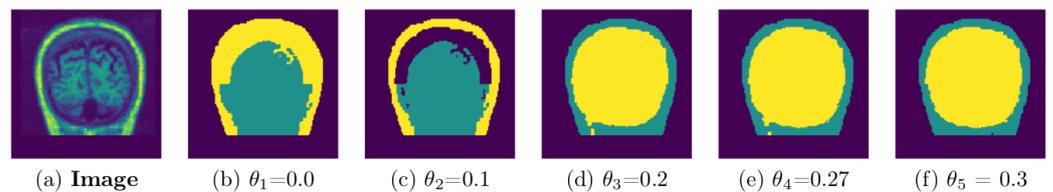


Figure 13. A representative coronal section from 3D MRI showing the separation of components for different thresholds: MRI from a single subject (IBSR data set) experimented using different thresholds and nodes sample size is 20,000. GUBS is run on 3D and for visualization we take 2D at the same location for all experiments.

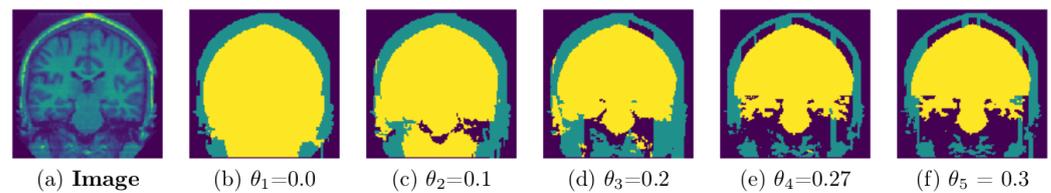


Figure 14. A representative coronal section from 3D MRI showing the separation of components for different thresholds: MRI from a single subject (OASIS data set) experimented using different thresholds and the nodes sample size is 20,000. GUBS is run on 3D and for visualization we take 2D at the same location for all experiments.

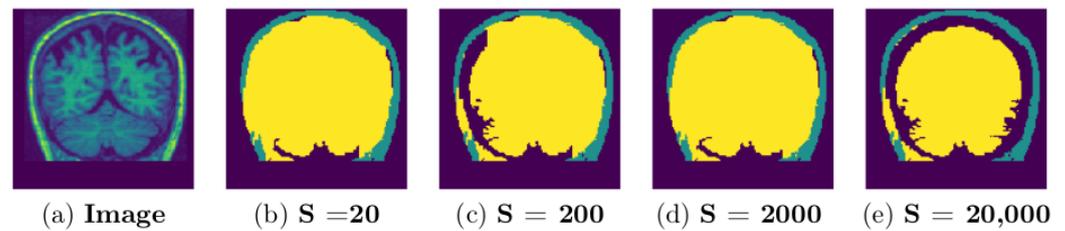


Figure 15. A representative coronal section from 3D MRI showing the separation of components for different sample size: MRI from a single subject (IBSR data set) experimented using threshold $T = 0.27$ and different size of the sampled nodes. GUBS is run on 3D and for visualization we take 2D at the same location for all experiments.

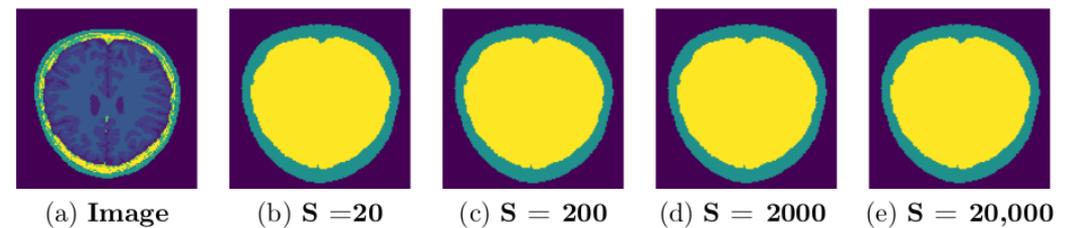


Figure 16. A representative axial section from 3D MRI showing the separation of components for different sample size: MRI from a single subject (BW data set) experimented without threshold, different size of the sampled nodes. GUBS is run on 3D and for visualization we take 2D at the same location for all experiments.

The selection line, δ_1 and δ_2 can highly vary based on the size of the images. The selection of parameters for the selection line depends on the visual inspection because the 3D MRI undergo different preprocessing steps after acquisition of the signals. For OASIS, IBSR and BW data sets the selection line values used were 60, 65 and 35, respectively. For all the three data sets $\delta_1 = 15$. For IBSR and OASIS $\delta_2 = 1$, for BW $\delta_2 = 3$.

Figures 9 and 10 show experimental result of one subject from IBSR and OASIS, respectively. The challenging part is to sample within the non-brain tissues because some parts of the brain and non-brain tissues have been removed (see row one in Figure 9). Some points sampled from the non-brain tissues are likely to be taken from the brain when δ_2 is applied. For Figures 9 and 10, $\delta_1 = 15$ and $\delta_2 = 1$.

3.3. Experimental Timing

The time complexity analysis for implementing GUBS is divided into three parts. These include adjacency graph construction, MST construction, and the time is taken for disconnecting the MST into different connected components. The adjacency graph size is defined by the number of voxels in the MRI image, corresponding to vertices in the adjacency graph. The experimental timing for the adjacency graph and MST construction is efficient and presented in [36]. For the adjacency graph constructed from MRI image $134 \times 134 \times 134$ (with 2,406,104 nodes in the adjacency graph), we update the MST twice to disconnect a path connecting the brain and the non-brain tissues as well as disconnect

the path connecting non-brain tissues and the background. The experimental time ranges between 12 and 25 s for disconnecting a path. The time spent on separating the tree into different connected components depends on the length of the path between the terminal nodes. The implementation was done by writing scripts in the python programming language, and it was run on a PC processor (Core i7-8650UCPU @ 1.90GHz×8).

4. Discussion

We extended the segmentation criteria used in the paper [36] by collapsing subgraphs of the voxels adjacency graph before constructing the MST and presented the GUBS method for segmenting the brain from MRI images. GUBS works by representing MRI volume into an adjacency graph, followed by collapsing representative nodes sampled from the brain, no-brain, and background regions, respectively. Then, the MST is constructed from the modified graph. The collapsed nodes are used as terminal nodes for disconnecting the paths in the MST to separate the brain, non-brain tissues, and background.

The approach was tested by segmenting the brain from three different MRI data sets. The results are compared to the ground truth to assess the performance of GUBS. The experimental results show that GUBS successfully segments brain with high performance from different data sets regardless of the differences of these data sets. Moreover, the results obtained using GUBS are compared to the results obtained using different state of the art methods. GUBS gives competitive results in terms of performance. However, unlike different state of the art methods which require labeled images for training, GUBS does not require labels. In most cases the labeled images are obtained manually by highly qualified individuals. This task is labour intensive and time consuming. That is one of strength of GUBS.

The evaluation of the results obtained using the GUBS approach from the three data sets indicates that the quality of the data highly influences the results. The results obtained from the BrainWeb data set indicate that GUBS outperforms different state of the art methods, whereas results obtained from the OASIS data set provide competitive results compared to the ones obtained previously using neural network approaches. In some cases, the specificity obtained by the GUBS approach outperforms other methods. Similar to other methods, GUBS get good results from IBSR data set. Since the approach depends on sampling points within different regions, the quality of the data might have influenced the IBSR results because some parts of the MRI images were removed. It is likely that nodes sampled in the wrong region will be collapsed together with nodes in a wrong region. Thus, the GUBS approach will be limited in the setting of incomplete 3D brain MRIs. An extension of GUBS to a semi-automated method, in which experts can sample points with high control before running the approach would alleviate this problem.

5. Conclusions

We developed a graph-based method that uses a minimum spanning tree (MST) to segment 3D brain from MRI images. The proposed method was tested by segmenting three different data sets. It is efficient and competes the state-of-the-art methods in terms of performance. The proposed method is simple to adapt and apply on different data sets.

Author Contributions: Conceptualization, S.M. and M.B.; Formal analysis, S.M.; Methodology, S.M. and M.B.; Software, S.M. and M.B.; Supervision, N.B., S.G. and M.B.; Visualization, S.M.; Writing—original draft, S.M.; Writing—review and editing, I.H., J.B.H., S.A., N.B., S.G. and M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research work is funded by the University of Bergen, Norway.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The first data set is from the Open Access Series of Imaging Studies (OASIS) [22], www.oasis-brains.org (accessed on 15 January 2022). The second data set is from

the BrainWeb (BW) database [23,24], <https://brainweb.bic.mni.mcgill.ca/> (accessed on 15 January 2022). The third data set is from the Internet Brain Segmentation Repository (IBSR) [25], <https://www.nitrc.org/projects/ibsr> (accessed on 15 January 2022). Codes for producing results are available here <https://github.com/simeonmayala/GUBS-Brain-segmentation>.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Despotović, I.; Goossens, B.; Philips, W. MRI segmentation of the human brain: Challenges, methods, and applications. *Comput. Math. Methods Med.* **2015**, *2015*, 450341. [[CrossRef](#)] [[PubMed](#)]
- Kalavathi, P.; Prasath, V. Methods on skull stripping of MRI head scan images—A review. *J. Digit. Imaging* **2016**, *29*, 365–379. [[CrossRef](#)] [[PubMed](#)]
- Ramon, C.; Garguilo, P.; Fridgerisson, E.A.; Hauelsen, J. Changes in scalp potentials and spatial smoothing effects of inclusion of dura layer in human head models for EEG simulations. *Front. Neuroeng.* **2014**, *7*, 32. [[CrossRef](#)] [[PubMed](#)]
- Fatima, A.; Shahid, A.R.; Raza, B.; Madni, T.M.; Janjua, U.I. State-of-the-art traditional to the machine-and deep-learning-based skull stripping techniques, models, and algorithms. *J. Digit. Imaging* **2020**, *33*, 1443–1464. [[CrossRef](#)]
- Li, J.; Erdt, M.; Janoos, F.; Chang, T.C.; Egger, J. Medical image segmentation in oral-maxillofacial surgery. In *Computer-Aided Oral and Maxillofacial Surgery*; Academic Press: Cambridge, MA, USA, 2021; pp. 1–27.
- Hwang, H.; Rehman, H.Z.U.; Lee, S. 3D U-Net for skull stripping in brain MRI. *Appl. Sci.* **2019**, *9*, 569. [[CrossRef](#)]
- Kapur, T.; Grimson, W.E.L.; Wells, W.M., III; Kikinis, R. Segmentation of brain tissue from magnetic resonance images. *Med. Image Anal.* **1996**, *1*, 109–127. [[CrossRef](#)]
- Gui, L.; Lisowski, R.; Faundez, T.; Hüppi, P.S.; Lazeyras, F.; Kocher, M. Morphology-driven automatic segmentation of MR images of the neonatal brain. *Med. Image Anal.* **2012**, *16*, 1565–1579. [[CrossRef](#)]
- Lucena, O.; Souza, R.; Rittner, L.; Frayne, R.; Lotufo, R. Convolutional neural networks for skull-stripping in brain MR imaging using silver standard masks. *Artif. Intell. Med.* **2019**, *98*, 48–58. [[CrossRef](#)]
- Rehman, H.Z.U.; Hwang, H.; Lee, S. Conventional and deep learning methods for skull stripping in brain MRI. *Appl. Sci.* **2020**, *10*, 1773. [[CrossRef](#)]
- Lemieux, L.; Hagemann, G.; Krakow, K.; Woermann, F.G. Fast, accurate, and reproducible automatic segmentation of the brain in T1-weighted volume MRI data. *Magn. Reson. Med.* **1999**, *42*, 127–135. [[CrossRef](#)]
- Smith, S.M. BET: Brain Extraction Tool. In *FMRIB TR00SMS2b (Oxford Centre for Functional Magnetic Resonance Imaging of the Brain)*; Department of Clinical Neurology, Oxford University, John Radcliffe Hospital: Headington, UK, 2000; pp. 1–25
- Dale, A.M.; Fischl, B.; Sereno, M.I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **1999**, *9*, 179–194. [[CrossRef](#)] [[PubMed](#)]
- Shattuck, D.W.; Leahy, R.M. BrainSuite: An automated cortical surface identification tool. *Med. Image Anal.* **2002**, *6*, 129–142. [[CrossRef](#)]
- Ségonne, F.; Dale, A.M.; Busa, E.; Glessner, M.; Salat, D.; Hahn, H.K.; Fischl, B. A hybrid approach to the skull stripping problem in MRI. *Neuroimage* **2004**, *22*, 1060–1075. [[CrossRef](#)] [[PubMed](#)]
- Iglesias, J.E.; Liu, C.Y.; Thompson, P.M.; Tu, Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Trans. Med. Imaging* **2011**, *30*, 1617–1634. [[CrossRef](#)]
- Galdames, F.J.; Jaillet, F.; Perez, C.A. An accurate skull stripping method based on simplex meshes and histogram analysis for magnetic resonance images. *J. Neurosci. Methods* **2012**, *206*, 103–119. [[CrossRef](#)]
- Roy, S.; Maji, P. A simple skull stripping algorithm for brain MRI. In Proceedings of the 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), Kolkata, India, 4–7 January 2015; pp. 1–6.
- Fatima, A.; Madni, T.M.; Anwar, F.; Janjua, U.I.; Sultana, N. Automated 2D Slice-Based Skull Stripping Multi-View Ensemble Model on NFBS and IBSR Datasets. *J. Digit. Imaging* **2022**, *35*, 374–384. [[CrossRef](#)]
- Rampun, A.; Jarvis, D.; Griffiths, P.D.; Zwiggelhaar, R.; Scotney, B.W.; Armitage, P.A. Single-Input Multi-Output U-Net for Automated 2D Foetal Brain Segmentation of MR Images. *J. Imaging* **2021**, *7*, 200. [[CrossRef](#)]
- Sadanathan, S.A.; Zheng, W.; Chee, M.W.; Zagorodnov, V. Skull stripping using graph cuts. *NeuroImage* **2010**, *49*, 225–239. [[CrossRef](#)]
- Marcus, D.S.; Wang, T.H.; Parker, J.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.* **2007**, *19*, 1498–1507. [[CrossRef](#)]
- Aubert-Broche, B.; Evans, A.C.; Collins, L. A new improved version of the realistic digital brain phantom. *NeuroImage* **2006**, *32*, 138–145. [[CrossRef](#)]
- Cocosco, C.A.; Kollokian, V.; Kwan, R.K.S.; Pike, G.B.; Evans, A.C. Brainweb: Online interface to a 3D MRI simulated brain database. *NeuroImage* **1997**, *5*, 425.
- Rohlfing, T. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE Trans. Med. Imaging* **2011**, *31*, 153–163. [[CrossRef](#)] [[PubMed](#)]

26. Kleesiek, J.; Urban, G.; Hubert, A.; Schwarz, D.; Maier-Hein, K.; Bendszus, M.; Biller, A. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. *NeuroImage* **2016**, *129*, 460–469. [[CrossRef](#)] [[PubMed](#)]
27. Brett, M.; Markiewicz, C.J.; Hanke, M.; Côté, M.A.; Cipollini, B.; McCarthy, P.; Jarecka, D.; Cheng, C.P.; Halchenko, Y.O.; Cottaar, M.; et al. nipy/nibabel: 3.2.1. 2020. Available online: <https://doi.org/10.5281/zenodo.4295521> (accessed on 1 February 2022). [[CrossRef](#)]
28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
29. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. doi: 10.1038/s41586-020-2649-2. [[CrossRef](#)]
30. Heng, P.; Wang, L.; Wong, T.; Leung, K.; Cheng, J.C.Y. Edge surface extraction from 3D images. In Proceedings of the Medical Imaging 2001: Image Processing, San Diego, CA, USA, 17–22 February 2001; Volume 4322, pp. 407–416.
31. Bomans, M.; Hohne, K.H.; Tiede, U.; Riemer, M. 3-D segmentation of MR images of the head for 3-D display. *IEEE Trans. Med. Imaging* **1990**, *9*, 177–183. [[CrossRef](#)]
32. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]
33. Morris, O.; Lee, M.d.J.; Constantinides, A. Graph theory for image analysis: An approach based on the shortest spanning tree. *IEE Proc. (Commun. Radar Signal Process.)* **1986**, *133*, 146–152. [[CrossRef](#)]
34. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. doi: 10.1109/MCSE.2007.55. [[CrossRef](#)]
35. Waskom, M.; Botvinnik, O.; O’Kane, D.; Hobson, P.; Lukauskas, S.; Gempertine, D.C.; Augspurger, T.; Halchenko, Y.; Cole, J.B.; Warmenhoven, J.; et al. mwaskom/seaborn: V0.8.1 (September 2017). Available online: <https://doi.org/10.5281/zenodo.883859> (accessed on 1 February 2022). [[CrossRef](#)]
36. Mayala, S.; Herdlevær, I.V.; Bull Haugsøen, J.; Anandan, S.; Gavasso, S.; Brun, M. Brain Tumor Segmentation Based on Minimum Spanning Tree. *Front. Signal Process.* **2022**, *2*, 816186. [[CrossRef](#)]