Performance Evaluation of Human Brain Structural Segmentation Toolkits

Akanji, W.A., Adelanwa, S.O.A. & Akinrinlola, I.A.
Department of Computer Science
Lagos State Polytechnic
Lagos, Nigeria
E-mails: akanjiwasiu2005@yahoo.com, ayinkaadelanwa@gmail.com, mideanddee@hotmail.com

ABSTRACT

The human brain is a complex nervous system, which receives information from the external natural world via the sensory nerves, integrates it and processed it in order to send the reaction information to the body via the motor nerves. Medical Image Analysis, involving the MRI scan of the human brain, relies heavily on segmentation. Where the technique or algorithm or toolkit used affects the outcome of the segmentation, the analysis is always susceptible to error. Such output cannot be relied upon for medical diagnosis or pre-clinical activities. This research evaluates the capabilities of three different toolkits and their respective algorithms in adequately segmenting human brain tissue from skull for the purpose of structural analysis of the various lobes of the subcortical structures. BrainSuite, Medical Image Processing and Visualization (MIPAV), and FMBRI Software Library (FSL) were evaluated using five evaluation metrics that contrasted our T2-weighted MRI scan segmented on each of the three toolkits with a Ground Truth: Dice Coefficient, Jaccard, Hausdorff distance, Mutual Information Intraclass and correlation coefficient. Our results depict that the Brain Extraction Tool (BET) on MIPAV outperforms FSL’s BET and BrainSuite’s Brain Surface Extraction (BSE).

Keywords: Skull Stripping, Magnetic Resonance Imaging (MRI), Image Segmentation, FSL, MIPAV, BrainSuite, Brain Tissues.

1. INTRODUCTION

The brain is a complex nervous system owned by animals, which steer the movement of animals and reactions in the natural environment. Generally, brain receives information from the external natural world through the sensory nerves that is integrated and processed by the brain, and the reaction information sent to the body through the motor nerves [1]. Human brain is an organ that is made up of a large mass of nerve tissue that is protected within the skull. It plays a role in just about every major body system. The Human brain is a complex system which made up of neurons and glial cells. Due to its complex nature, the structural part of the brain is very complex in nature [2].
It is the central organ of the human nervous system, and with the spinal cord, makes up the central nervous system. It controls most of the activities of the body, processing, integrating, and coordinating the information it receives from the sense organs, and making decisions as to the instructions sent to the rest of the body [3]. The human brain is enclosed within the skull, which provides lateral, frontal and dorsal protection. It also consists of 22 bones, 14 of which form the facial bones and the remaining eight (8) forms the cranial bones [4]. Anatomically, the brain is contained within the cranium and is surrounded by the cerebrospinal fluid (CSF). Cerebrospinal Fluid (CSF) is a fluid that circulates within the skull and spinal cord, filling up hollow spaces on the surface of the brain. The human brain can be divided into hundreds of parts, and in this project, we will introduce three anatomical regions (including their structures, locations and functions) following the standard neuroanatomy hierarchies, including brainstem, cerebellum, diencephalon and cerebrum are included [5].

Medical imaging refers to several different technologies that are used to view the human body in order to diagnose, monitor, or treat medical conditions. Each type of technology gives different information about the area of the body being studied or treated. The information obtained from the images that the technologist produces is used to assess possible disease, injury, or the effectiveness of medical treatment. It is a type of medical imaging technique that utilizes a certain physical mechanism to detect patient internal signals that reflect either anatomical structures or physiological events.

Segmentation is the act of dividing or partitioning into segments, in the analysis of medical image, it is essential that organs be distinguished or segmented from their background and adjacent structures for ease of analysis. Image segmentation technique is the most important task in medical image analysis and is often the first and mostly useful in many clinical applications. It is a method of dividing or partitioning an image into parts, called segments. Segmentation techniques have gained importance in the quantitative analysis of medical images and in image guided interventional procedures over the past decades. Manual, semi-automated and automatic segmentations are the three basic segmentation techniques used in analyzing medical images [6, 7]. The goal of this paper is to separate the skull from the tissue using three automated toolkits: Functional Magnetic Resonance Imaging of the Brain, Medical Image Processing Analysis and Visualization and BrainSuite on T2-weighted MRI scans and analyze the output under five evaluation metrics.

2. RELATED WORK

In the analysis of medical images several authors and scholars have worked on skull stripping; some manual, semi or fully automated, it is often essential that objects/organisms be stripping or segmented from their background. In the study over the past decades now, segmentation techniques have gained importance in the quantitative analysis of medical images and in image guided interventional procedures. Many approaches have now been investigated and some have been implemented in commercial imaging and analysis systems. The literature on segmentation techniques is rich with innovative ideas and algorithms. Kumbhar & Holambe [8] proposed a global threshold applicable when the intensity distribution of objects and background pixels are sufficiently distinct in which a single threshold value is used to represent the whole image. The study showed that the pixel values of the components and that of background are fairly consistent in their respective values over the entire image when the global thresholding is used.

Guo & Pandit [9] proposed a threshold selection histogram based technique for estimating the threshold value that separates the two homogeneous region of the foreground and background of an image. The technique is suitable for large homogeneous images with ability to separate regions where all area of the objects and background are homogenous except the area between the objects and the background.
Wang et al. [10] proposed an iterative heuristic thresholding technique in which the initial threshold value is set to the average brightness with ability to improve the anti-noise capability. The method requires no specific knowledge about the image. Vala & Baxi [11] proposed Otsu method used to overcome the drawback of iterative thresholding by identifying the optimal threshold using histogram of the image in finding the optimal value for the global threshold.

Balla-Arabé & Gao [12] proposed a multi-thresholding technique that determines more than one threshold for the given image and segments the image into certain brightness regions, which correspond to one background and several objects. The method works effectively for objects with colored or complex backgrounds, on which bi-level thresholding generates unsatisfactory results. Fuzzy c-means (FCM) clustering is an unsupervised technique that has been successfully applied to feature analysis, clustering, and classifier designs in fields such as astronomy, geology, medical imaging, target recognition, and image segmentation. Fuzzy C-means its improvement methods algorithm and strategies for remote sensing image segmentation can offer less iterations times to converge to global optimal solution was proposed by Bezdek et al. [13].

Likas et al. [14] had proposed a K-means clustering as an efficient method of threshold selection. Using this algorithm, the image is divided into k segments using (k-1) thresholds and minimizing the total variance within each segment. Senthilkumaran & Vaithiogi [15] introduced a niblack and sauvola thresholding method for images with non-uniform background by calculating several thresholds for every pixel using the mean and standard deviation defined by a window centered on the pixel.

Yanowits and Bruckstein’s method, a threshold surface is constructed by finding the edge points of the smoothed image. The gradient magnitude image is computed and thinned to one pixel-wide line to identify edge points. An iterative interpolation process is employed to get a smooth surface passing through the edge points. The image is thresholded by the constructed surface. In an iterative interpolation process, the interpolated surface is set at image gray scale at the edge points and 0 at the other points proposed by Blayvas et al. [16].

Adaptive thresholding typically takes a grayscale or color image as input and, in the simplest implementation, outputs a binary image representing the segmentation. For each pixel in the image, a threshold has to be calculated. If the pixel value is below the threshold it is set to the background value, otherwise it assumes the foreground value. In adaptive thresholding, different threshold values for different local areas are used as proposed by Roy et al. [17]. The concept of entropy is to compute the threshold value but in different perspective. Two probability distributions were considered instead of one. One probability was for the foreground of the image and another one was for the background. The sum of the individual entropy of the foreground and the background was maximized later. This process created equal-probability gray levels in every region with the desired threshold capable of gray level minimization of relative entropy as proposed by Manic et al. [18].

The Magnetic Resonance Images (MRI) of head scans revealed the details of human brain and its surrounding tissues. Skull stripping is the process of eliminating all non-brain tissues from brain images. It removes extra cerebral tissues such as skull, sclera, fat and skin from the brain images. It is an essential preprocessing step in neuroimage analysis. However, the presence of image artifacts, anatomical variability, varying contrast properties and poor registration make this process a difficult one. The brain portion must be skull stripped before other image processing algorithms such as registration, tissue classification or compression can be applied as proposed by Park & Lee [19].
A number of manual, automated and semi-automated skull stripping algorithms are available in literature. Each of the existing skull stripping technique has its own strengths and weaknesses. Brain Extraction Tool (BET) employs deformable model that evolves to fit the brain’s surface by the application of a set of locally adaptive model forces. The Brain Surface Extraction (BSE) proposed by Fennema [20] used a combination of operations such as anisotropic diffusion filtering, Marr-Hildreth edge detection and morphological operations, to separate brain from non-brain tissues.

The Watershed Algorithm (WAT), intensity based approach that operates under the assumption of white matter connectivity and segment the image into brain and non-brain components. But it often produces over segmentation and is sensitive to noise present in the image. A hybrid method for skull stripping, Hybrid Watershed Algorithm (HWA) proposed by Zhang & Cheng [21] combine the watershed techniques and a deformable surface model. This method first localizes a single white matter voxel in T1-weighted image, and uses it to create a global minimum in the white matter before applying a watershed algorithm.

Zhao & Zhuang [22] developed an automatic skull stripping algorithm called Model-based Level Set (MLS) based on active curve to remove the skull and intracranial tissues surrounding the brain in MR brain images. Skull stripping using Graph Cuts (GCUT) relies on graph theoretic image segmentation techniques to position cuts that serve to isolate and remove dura. A statistical shape model for automatic skull stripping of brain images proposed is based on surface model of the brain boundary and it is hierarchically represented by a set of overlapping surface patches, each of which has elastic property and deformation range that is learned from a training set. Park and Lee developed a skull stripping method for coronal T1-weighted images based on region growing. It aims to automatically detect two seed regions of the brain and non-brain by using a mask produced by morphological operations. Then the seed regions were expanded using 2D region growing algorithm.

Smith [23] developed a fully automatic brain extraction algorithm using diffusion to run length encoding and region labeling generated for skull stripping in T2 weighted axial MR brain images. Several studies have been reported for skull stripping based on anisotropic diffusion filter and morphological processing seed growth and threshold technique sing techniques, fuzzy-ASM based, and deformable surface and tissue classification. Most of these methods are applicable to T1 weighted MR brain images, and does not extract the brain completely in all the slices. Moreover, none of these existing methods give satisfactory performance when evaluated for large scale data set. It is due to the complexity of the human brain, varying image contrast properties, image artifacts such as undersampling, noise factor, variations in the image orientations and types. There appears to be no single method that works for all the three types of T1 and T2 weighted images. In this paper, we propose a fully automatic method for skull stripping in T1 and T2.

3. METHODOLOGY

The major objective is to evaluate the performances of the algorithms on these toolkits’ vis a vis their respective outputs. The research seeks to use automated technique and methods to segment the human’s brain. The toolkits selected for evaluation are FSL, BrianSuite and MIPAV for processing and analysing the human brain on MRI scans.

3.1 The BrianSuite
It has sophisticated visualization tools, such as an MRI providing three orthogonal views (either separately or in 3D) and an overlay surface visualization of cortex, skull, and scalp. It is an integrated image-guided surgical system which provides clear delineation of the major white matter tracts, coloured 3D reconstructions with graphic resolution of relationships between tracts and gray matter, as well as 3D reconstructions of cerebral vasculature
This system provides high-quality images requiring short scan times and makes various sequences available intraoperatively, including MRI angiography and DTI tractography. In brief, this fully integrated architecture facilitates straightforward workflow including patient transport, image scanning, image fusion, surgical planning, registration for navigation, and updating of navigation images. Hereunder are the steps required to perform skull stripping using BrainSuite. BrainSuite contains tools to automatically extract the brain and then generate 3D models of its cortical surface that can be used for further analysis.

BrainSuite steps are

**Step 1:** Open BrainSuite software
**Step 2:** Open Dataset, via File > Open Volume
**Step 3:** Load the Dataset
**Step 4:** Open the CSE dialog via Cortex > Cortical Surface Extraction Sequence
**Step 5:** Click Cortical Surface Extraction Sequence” dialog
**Step 6:** Click on it three more times, until the Skull Stripping Tool dialog says “finished skull stripping”
**Step 7:** Click back three times
**Step 8:** Change the edge constant to 0.5, click “Step >” until the dialog once again says “finished skull stripping”. Again click “Step >” until skull-stripping is completed.
**Step 9:** Close the Skull Stripping dialog box and return to the Cortical Surface Extraction Sequence dialog. The next stage, “Skull and Scalp”, produces surfaces for the skull and scalp, which are useful when you are working with EEGs and MEGs. This scan has been defaced for privacy protection, and so running this stage will not be helpful. Uncheck the box next to it.
**Step 10:** Click “Stage >>” three times to finish skull-stripping and perform bias-field non-uniformity correction and tissue classification.
**Step 11:** Click “Stage” to finish the cerebrum labeling. Notice that the mask now excludes the cerebellum and brainstem, and much of the non-brain tissue.
**Step 12:** “Stage” four times to run Inner Cortical Mask, Scrub Mask, Topology Correction, and Wisp Removal. These tools produce and clean up the white matter mask that will act as a foundation for producing the 3D inner cortical surface that is in turn used as the foundation for generating the pial surface.
**Step 13:** Click “Stage >>” to produce the inner cortical surface, which will appear in the bottom right surface viewer.

The final stage after pial surface generation is splitting and labeling the hemispheres. Click “Stage >>” to complete it. Final output of CSE, with cross-hairs off and volume slices hidden in surface view Observe the results. When the CSE Sequence is completed, the surface viewer will display a 3D model of the pial surface of the skull-stripped scan.

### 3.2. Functional Magnetic Resource Imaging Of Brian [Fmrib] Software Library (FSL)

FMRIB Software Library (FSL) is a comprehensive library of analysis tools for FMRI, MRI and DTI brain imaging data. The most widely used software packages in the neuroimaging community contain automated segmentation routines that utilize iterative approaches to classify human brain images into the three tissue classes. These packages include: Intracranial segmentation commonly referred to as “skull-stripping” removes extra-cerebral tissues such as skull, eyeballs, and skin. Skull-stripping facilitates image processing such as surface rendering, cortical flattening, image registration, and tissue segmentation. Thus, as the first step of the FSL segmentation, the “Brain Extraction Tool version 2.1 (BET)” integrated in FSL software is used to perform skull stripping and remove non-brain parts of the image. FSL version 4.1 is employed for the whole process of registration, skull stripping and brain tissue segmentation. Since the FSL-FAST brain tissue segmentation requires skull stripped version of input MRI data, as the first step of tissue segmentation, skull stripping is performed using the FSL’s own brain extraction tool.
Then, in the second step, FAST tool using probability maps as its default settings is used to segment the brain into three tissue classes of GM, WM and CSF and performing bias correction. FAST (FMRIB's Automated Segmentation Tool) segments a 3D image of the brain into different tissue types (Grey Matter, White Matter, CSF, etc.), whilst also correcting for spatial intensity variations (also known as bias field or RF in homogeneities). The underlying method is based on a hidden Markov random field model and an associated Expectation-Maximization algorithm. The whole process is fully automated and can also produce a bias field-corrected input image and probabilistic and or partial volume tissue segmentation. It is robust and reliable, compared to most finite mixture model-based methods, which are sensitive to noise.

**Step 1:** Install or download the FSL on Linux window.

**Step 2:** Run FSL. From the main FSL menu bar, select 'Plugins-> Install plugin'. The program may ask you to restart the window.

**Step 3:** Starting with the quality T2-weighted MR image volume, the next step in the FSL processing is to load the data set.

**Step 4:** Step one of Select the good quality T2-weighted MR image volume

**Step 5:** Click on OK button to load the T2-weighted MR image volume

The image in Fig. 3.1 depicts the image before it is been segmented.

![Fig 3.1: Loaded Un-segmented Image (Skull and Tissue)](image)

Fig. 3.2 shows the output of the proceed image of Weighted T2

![Fig 3.2: Segmented Image/output](image)
3.3. Medical Image Processing, Analysis And Sualization (Mipav)

MIPAV is a multifaceted, platform independent, quantification, and visualization application for biomedical images. MIPAV supports numerous medical and research file formats including Zeiss, Biorad, TIFF, DICOM3.0, NIFTI, AFNI, MINC, GE, Siemens, Philips, AVI, Interfile, MIPAV XML, and new formats are added as needed. The software tools were developed as Java plug-ins for MIPAV. During initialization, MIPAV automatically loads installed plug-ins and builds the plug-in menu, allowing them to be utilized as if they were native algorithms. The steps for accomplished skull stripping on MIPAV are

Step 1: Install MIPAV: run the MIPAV installer
Step 2: Run MIPAV. From the main MIPAV menu bar, select ‘Plugins-> Install plugin’. The program may ask you to restart MIPAV, in which case you close MIPAV, restart, and repeat. You will get the ‘Install Plugin’ window:
Step 3: select the directory for the CBS Tools, JIST and TOADS-CRUISE with the ‘Browse’ button.
Step 4: Open the MIPAV software from the computer system
Step 5: Starting with the good quality T2-weighted MR image volume to load the data set
Step 6: Select and load the T2 weighted image from the storage
Step 7: Click on toolbar and Algorithm then select BET as shown below
Step 8: Extraction in process

4. SIMULATION RESULTS

Table 4.1 shows the results generated from comparison of Ground Truth T2 weighted image with BrainSuite, MIPAV and FSL using MATLAB software based on some selected evaluation metrics which includes Dice Coefficient, Jaccard, Hausdorff distance, Mutual information and Intraclass correlation coefficient.

<table>
<thead>
<tr>
<th>S/N</th>
<th>Image Evaluation Metric</th>
<th>Ground Truth Vs BrainSuite</th>
<th>Ground Truth Vs MIPAV</th>
<th>Ground Truth Vs FSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sorensen Dice coefficient</td>
<td>0.89607</td>
<td>0.91755</td>
<td>0.8944</td>
</tr>
<tr>
<td>2</td>
<td>Jaccard</td>
<td>0.81171</td>
<td>0.84766</td>
<td>0.8090</td>
</tr>
<tr>
<td>3</td>
<td>Hausdorff distance</td>
<td>4.69040</td>
<td>4.24260</td>
<td>4.8990</td>
</tr>
<tr>
<td>4</td>
<td>Mutual information</td>
<td>0.60560</td>
<td>0.99070</td>
<td>0.6019</td>
</tr>
<tr>
<td>5</td>
<td>Intraclass correlation coefficient</td>
<td>0.26950</td>
<td>0.25340</td>
<td>0.2682</td>
</tr>
</tbody>
</table>

Fig. 4.1 and Fig. 4.2 show the graphical results of BrainSuite, FSL and MIPAV based on the image evaluation metrics.
5. DISCUSSION

**Dice And Jaccard**

DSC is the quotient of similarity and ranges between 0 and 1. It can be viewed as a similarity measure over sets. The degree of similarity increases as the result tends to 1.

**Intraclass Correlation Coefficient**

Values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability. Moreover, the ICC estimate obtained from a reliability study is only an expected value of the true ICC. It is logical to determine the level of reliability (i.e., poor, moderate, good, and excellent) by testing whether the obtained ICC value significantly exceeds the suggested values mentioned above using statistical inference.
Mutual Information
With reference to our results, as shown in Table 4.1, comparing five evaluation metrics for each T2 segmented image and the Ground Truth; we observe that the BrainSuite, MIPAV and FSL do not segment exactly the same, even when BET algorithms are used on MIPAV and FSL. The project employed MATLAB software for our results generation, MATLAB have the capability to compute the value for each metric used in evaluation. For Dice coefficient (DSC), the value is 0.91755 for MIPAV, making it the best amongst the rest that is closest to 1. For Jaccard, the value is 0.84766 for MIPAV, making it the best amongst the rest that is closest to 1. For Hausdorff distance, the value is 4.24260 for MIPAV, being the smallest, makes it the best amongst the rest. For Mutual information, the value is 0.99070 for MIPAV, making it the best amongst the rest. For Intraclass correlation coefficient, the value is 0.25340 for MIPAV, making it the best amongst the rest.
Based on the above analysis, for BrainSuite, MIPAV and FSL, our result shows that the segmentation done with MIPAV, using BET Algorithm, gives higher accuracy and reliability compared than the rest.

6. CONCLUSION
Image segmentation is an important step in many medical applications involving 3D visualization, computer-aided diagnosis, measurements, and registration. This research work has provided a brief introduction to the fundamental concepts of MRI segmentation of the human brain on three toolkits/algorithms that are commonly used. During the past decades, various image segmentation algorithms and toolkits have been proposed with different levels of accuracy, computational time, complexity, and efficiency. The accuracy of brain MRI segmentation should be increased to reduce false diagnosis. Brain MRI segmentation and toolkits are one of the most active research areas in image processing. Large amount of MRI data is generated daily and requires computer aided processing for decision making.

The paper presented the basic algorithms of MRI segmentation methods, such as local thresholding, adaptive thresholding, and global thresholding. In addition, steps to segment brain from skull for MRI have been described using BrainSuite, FSL, and MIPAC. The paper also provides insight into evaluating toolkits or algorithms using five valuation metrics that have been used frequently. Our results depict that MIPAV toolkit using BET algorithm is most reliable and efficient in segmenting human brain. It was also observed that out of the five evaluation metrics used, Sorensen Dice coefficient has universal acceptability in comparing ground Truth with segmented MRI T2 image.

REFERENCES


