

# Skull Stripping on Multimodal Brain MRI Scans using Thresholding and Morphology

Sajid Y. Bhat, Afnan Naqshbandi and Muhammad Abulaish, *SMIEEE*

**Abstract**—This paper introduces a novel thresholding and morphology-based skull stripping method for different MRI modalities. The proposed method is designed in a way which is easy to use and generates satisfactory results with minimal parameter adjustments. The method is evaluated on three different benchmark datasets and compared with nine state-of-the-art skull stripping methods. The experimental results suggest that the proposed method generates comparable results to some of the best methods in literature. However, unlike many other methods, it works well on different types of MRI scans. Moreover, this method generates the skull mask along with the brain mask that can be used to study various skull pathologies.

**Index Terms**—Medical Imaging, Image Processing, Skull Stripping, Thresholding, Morphology, MRI, Skull, Brain.

## I. INTRODUCTION

SKULL stripping is the process of separation of brain and non-brain areas. It is a widely used technique in brain image analysis. Here, the region of interest is a skull separated brain. This method has proven helpful in the field of medicine and hence gained interest in the domain. It may help in early diagnosis of various neural disorders, which would otherwise be fatal, if not diagnosed early. Skull stripping may be a requirement when only parts of brain like cerebrospinal fluid, grey matter and white matter are the area of interest.

Skull stripping in itself is a challenging task; the hindrance being the intensity similarities of brain and non-brain tissues. Hence, achieving skull stripping may involve a lot of morphological and thresholding processing. Selecting an optimal threshold that separates foreground and background becomes difficult in many cases. Even if machine learning based skull stripping approaches seem to be promising, new types of MRI sequences evolve with time, on which the existing methods may not work without proper re-training of the classification models. The traditional intensity-based skull stripping methods define brain areas based on intensities. However, as seen through the experimental results in reported in [1], it is clear that different MRI scanners produce different fMRI results, and hence we may obtain different output for same input using different scanners. A large number of proposed methods are based on the largest connected component assumption [2-5], wherein the brain is assumed to be the largest connected component after binarization.

However, identifying an optimal threshold to separate the brain and skull using binarization is a challenging task.

This paper presents a novel thresholding and morphology-based skull stripping method. The proposed method differs from traditional skull stripping methods mainly in terms of the approach used for skull stripping. Instead of focusing on brain extraction directly, the proposed method focuses on the detection and extraction of the skull in the first stage. The identified skull region is then subtracted from the original image, which results in the raw-brain area. The raw-brain is then processed using morphological operations to remove various protrusions and cerebrospinal fluid. The subtraction of the skull area from a scan can be considered as a preprocessing step which ensures that the final skull stripped result excludes any possible portions from the skull. This is unlike existing thresholding and morphology based methods that perform thresholding and largest connected component detection (for brain extraction) on the whole scan, often resulting in less accurate results. Moreover, the proposed approach easily generalizes on multimodal scans having varying intensities for the skull and brain regions respectively. It also helps to address the challenge of overfitting as faced by machine learning (ML) and deep learning (DL) approaches which often fail to generate acceptable results on the modalities that they are not trained on. Moreover, unlike the ML and DL methods the requirement of huge pre-labeled ground truth is not faced by the proposed method.

The rest of the paper is organized as follows. Section II presents the related work and section III presents the proposed work. The method is evaluated in section IV wherein it is compared with other state-of-the-art methods on some benchmark datasets. Finally, section V concludes the paper and presents future directions of research.

## II. RELATED WORK

Skull stripping methods can be broadly divided into four main categories [6]: (i) conventional skull stripping approaches, (ii) machine learning-based approaches, (iii) deep learning-based approaches, and (iv) hybrid skull stripping approaches.

The conventional skull stripping approaches may be sub-categorized into morphology-based methods [4-11], intensity-based methods [8, 32, 33], deformable surface-based methods [12-15], and template-based methods [16, 25-28]. Conventional

S. Y. Bhat, is with Department of Computer Sciences, University of Kashmir, Hazratbal, Srinagar, J&K, India (e-mail: [bhatsajid@uok.edu.in](mailto:bhatsajid@uok.edu.in)).

A. Naqshbandi, is an M.Tech. CS student at Department of Computer Sciences, University of Kashmir, Hazratbal, Srinagar, J&K, India (e-mail: [afnaan.naqshbandi@gmail.com](mailto:afnaan.naqshbandi@gmail.com)).

M. Abulaish, is with Department of Computer Science, South Asian University, New Delhi, India (e-mail: [abulaish@ieee.org](mailto:abulaish@ieee.org)).

skull stripping methods are the most frequently used traditional methods. Morphology-based methods are the ones that perform skull extraction by making use of various morphological operations. These methods can be mathematical morphology-based methods [32], binarization methods [9], fuzzy morphological operations-based methods [10], and edge detection-based methods [4].

Roslan et al. [32] proposed a mathematical morphology-based method which makes use of double and Otsu's thresholding to perform binarization. The method assigns a particular set of intensities to cerebral and non-cerebral tissues and defines the binary image accordingly. This convention may fail in some cases, as seen through experimental results in [1], where different MRI scanners produce different outputs on same input. Üsame et al. [9] proposed a simple binarization skull stripping method that segments the MR stack into grey matter, white matter and cerebrospinal fluid. The pixels belonging to these segments are then summed up to create the skull stripping mask. It uses these regions as priori knowledge and converts the MR images into binary format. Even if this method has proven to be robust, it is very rare and difficult to separate brain and non-brain regions using simple binarization because there exist strong connectivity among them. Also, in most of the cases, using different thresholds on same images results in different output. Kader et al. [10] proposed a skull stripping algorithm based on fuzzy morphological operations. It performs binarization using Otsu's thresholding and is based on largest connected component selection assumption, which is a pitfall because one cannot always claim the brain being a largest connected component. At times, finding the largest connected component also becomes difficult due to the strong connectivity between the skull and brain, especially in presence of anomalies like tumors. Swiebocka-Wiek [4] proposed an edge detection-based morphological skull stripping method, which involves image thresholding and brain mask generation using morphological operations. Finally, it is followed by edge detection and selection of the largest connected component as the brain. Experimental results of this method have shown that Cerebrospinal fluid was partially preserved, while it should have been removed.

Intensity-based skull stripping methods separate the brain and non-brain parts based on MRI pixel intensities. Somasudaram and Kalaiselvi [8] proposed a morphology and intensity thresholding-based method, wherein intensity thresholding is used to obtain a binary image and morphological operations are performed to refine the brain boundaries. It uses feature extraction to detect the brain followed by morphological operations to obtain segmentation. The drawback of the intensity-based methods is that they are highly sensitive to intensity variations in brain MRI. To solve this challenging issue of optimal intensity threshold estimation, Walczak et al. [41] use genetic algorithms to progressively tune the parameters for cortical surface extraction from magnetic resonance scans.

Deformable surface-based skull stripping methods are the ones where a model is deformed iteratively based on the image intensity to overlap with the input image brain area. The number of iterations depends on the accuracy of the required solution. Tao and Chang [12] proposed a deformable surface-based skull stripping algorithm, which uses a combination of deformable surface models and fuzzy tissue classification. The method starts with finding the centroid of the brain by analyzing the

input image intensities and initializing an ellipsoidal surface around the identified centroid. It is then followed by tissue classification and bias field estimation. Cox [15] introduced a deformable surface-based skull stripping model, AFNI, which has the ability to view each cardinal plane simultaneously and allows transformation of functional and anatomical scans into stereotaxic coordinates by identifying some anatomical landmarks. However, the main drawback of the deformable methods is their dependence on the surface models used, which affects their working accuracy.

Template-based skull stripping models are the ones wherein the skull stripping is performed based on template or atlas of brain MRI. In these methods, the brain and non-brain separation varies depending on the type and number of templates used. Eskildsen et al. [26] proposed a non-local patch-based segmentation method, BEaST, which makes use of squared differences for estimating the patch distances. Smith et al. [27] proposed a template-based skull stripping method, FSL-BET, to remove the non-brain parts from a head MRI scan. Bauer et al. [28] proposed a template-based skull stripping model, where affine transformation models are used to register the atlas image to patient image. However, main disadvantage of the template-based methods is their dependence on templates used and their registration.

Machine Learning (ML) based skull stripping models [17, 18] use supervised learning to perform skull stripping. Here, a model is trained on different sample datasets. One of the benefits of ML-based methods is the requirement of less human intervention for parameter tuning after the model has been trained. Iglesias et al. [17] proposed a ML-based brain separation method, ROBEX, which uses a combination of discriminative and generative models to perform skull stripping. The former is used to predict the brain boundary and the latter is used to smoothen the results. Even if this method has generated robust results on some datasets, the disadvantage of ML-based algorithms is that they usually work efficiently on the type of scans that were used in model training.

Deep learning skull stripping methods [2, 19] involve the usage of convolutional neural networks. Here the network consists of various convolution layers through which each input passes. The brain and non-brain separation usually involves either of the two approaches:

1. Voxel wise networks.
2. Fully convolutional networks.

Isensee et al. [2] proposed HD-BET, an artificial neural network-based brain extraction method wherein the performance is evaluated by comparing the segmentation results of different brain extraction methods to their referential brain masks, similarity between segmentation masks are measured using Dice similarity coefficient and Hausdorff distance. Even if this method has shown good results in case of pathologically altered brains, the pitfall of deep learning methods is that it required large amount of labelled data for training. To have these methods work on different MRI sequence types, these methods need training on all types of data and involves a lot of manual data labelling. More recent methods for skull stripping involving deep learning based architectures incorporate 2D and 3D variations [34-37] of the

TABLE I  
SUMMARY OF RELATED WORK

Approach	Methods	Pros	Cons
Morphology Based	Roslan et al. [32] Üsame et al. [9] Kader et al. [10] Swiebocka-Wiek [4] Abd El-Kader et al. [7] Somasudaram and Kalaiselvi [8] Ezhilarasan et al. [5] Tchoketch et al. [11]	<ul style="list-style-type: none"> <li>• Usually, fast.</li> <li>• Less parameter tuning.</li> <li>• No groundtruth required.</li> </ul>	<ul style="list-style-type: none"> <li>• Identifying optimal thresholds are often challenging.</li> </ul>
Intensity Based	Somasudaram and Kalaiselvi [8] Roslan et al. [32] Somasundaram et al. [33]	<ul style="list-style-type: none"> <li>• Usually, fast.</li> <li>• Less parameter tuning.</li> <li>• No groundtruth required.</li> </ul>	<ul style="list-style-type: none"> <li>• Intensity characterization is often difficult.</li> </ul>
Traditional Machine Learning Based	Iglesias et al. [17] Kobashi et al. [18]	<ul style="list-style-type: none"> <li>• Less human intervention for parameter tuning for the trained model</li> </ul>	<ul style="list-style-type: none"> <li>• Often over fit on training data.</li> <li>• Do not generalize on multiple modalities.</li> </ul>
Deformable Surface-Based	Tao and Chang [12] Ségonne et al. [13] Smith [14] Cox [15]	<ul style="list-style-type: none"> <li>• Less parameter tuning.</li> </ul>	<ul style="list-style-type: none"> <li>• Dependence on the surface models used affects accuracy.</li> </ul>
Template-Based	Dale et al. [25] Leung et al. [16] Bauer et al. [28] Eskildsen et al. [26] Smith et al. [27]	<ul style="list-style-type: none"> <li>• Fairly good results on T1 modalities.</li> </ul>	<ul style="list-style-type: none"> <li>• Dependence on templates and their registration.</li> <li>• Do not generalize on multiple modalities.</li> </ul>
Hybrid	Se'gonne et al. [13] Shattuck et al. [3]	<ul style="list-style-type: none"> <li>• Aim to achieve high accuracy.</li> </ul>	<ul style="list-style-type: none"> <li>• Base methods incorporating varying parameters.</li> </ul>
Deep Learning Based	Isensee et al. [2] Long et al. [19] Hoopes et al. [34] Lucena et al. [35] Thakur et al. [36] Hwang et al. [37]	<ul style="list-style-type: none"> <li>• High Accuracy on known modalities.</li> <li>• Less human intervention for parameter tuning for the trained model</li> </ul>	<ul style="list-style-type: none"> <li>• Large Amount of labelled training data required.</li> <li>• Possibility of overfitting.</li> </ul>

U-Net-based convolutional neural networks [10]. Hoopes et al. [34] aim to generalize the approach to multiple modalities by generating a synthetic training dataset with varying and more generalized anatomies, intensity distributions, and artifacts based on anatomical segmentations.

Hybrid methods [3, 13, 17] of skull stripping involve a combination of two or more existing skull stripping methods. Se'gonne et al. [13], proposed a hybrid skull stripping approach, which uses a combination of Watershed algorithm and Deformable Surface models. The algorithm starts with localizing the white matter voxel in the MR input T1-w image to build an approximation of brain volume using the Watershed algorithm. The surface deformation model is then used to refine and smoothen the brain mask. Shattuck et al. [3] proposed a hybrid skull stripping method, BrainSuite, which uses a combination of edge detection and morphological processing techniques.

The summary of the reviewed skull stripping methods is provided in table I. For a more comprehensive review of the other state-of-the-art skull stripping methods, the reader can refer to [5, 6, 31, 38].

### III. PROPOSED METHOD

Unlike existing morphology and thresholding-based skull stripping methods, the proposed method aims to model the skull stripping problem as a skull detection and subtraction problem, instead of a direct brain detection problem. The input to the method can be a single grayscale axial MRI scan or a full volume. In the latter case, the method processes each axial scan separately as a single image. The input scan is processed through various stages described in the following sections.

#### A. Blob Generation and Outline Detection

The method starts with identifying the whole skull blob which aims to separate the foreground (both hard and soft tissue) and the background in the input axial image  $I_o$ . This is achieved by thresholding a log-transformed and contrast stretched version of the input image referred as  $I_p$  at a value  $T_b$ . The value of threshold  $T_b$  is determined by finding the local minimum in the grayscale range [130 to 170] of the histogram of  $I_p$ . The result of this binarization represents the raw blob image  $I_b$ .



Often, MRI scans contain annotation labels and artifacts generated by over-exposures. To deal with this, the raw blob image  $I_b$  is processed by performing morphological spur reduction to remove small protrusions, cleaning to remove isolated pixels and closing to fill small gaps and breaks in the skull blob that may appear due to intensity variations. This processed blob image is then hole-filled and the largest connected component is extracted that represents the foreground mask or the whole skull blob ( $Blob_{mask}$ ).

To be able to identify the skull bone from  $I_o$ , some cue regarding the location of the same is required. To get a rough mask indicating the skull bone, an outline-mask ( $Outline_{mask}$ ) is extracted from the  $Blob_{mask}$  using the Bradley's adaptive thresholding [20] with the sensitivity threshold value set to 0.2. This process is illustrated in figure 1 wherein figure 1a is the input image, figure 1b is the raw blob, figure 1c is the processed blob mask, and figure 1d is the extracted outline-mask.

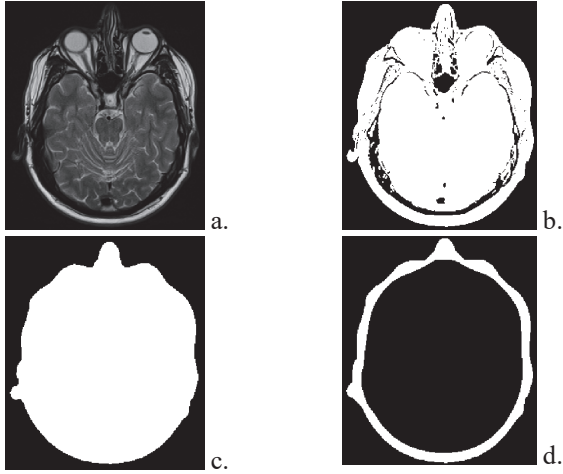


Fig. 1. Illustration of Blob generation and outline detection. a) Original image. b) Raw blob image  $I_b$ . c) Whole skull blob  $Blob_{mask}$  d) Outline mask image  $Outline_{mask}$ .

## B. Skull Extraction

The  $Blob_{mask}$  extracted in the previous section is used to extract the region of interest (i.e., the whole skull) from the original image  $I_o$  by taking an *AND* between the  $Blob_{mask}$  and  $I_o$ . This results in removing annotation labels and over-exposures from the input image. The resulting image  $I_{crop}$  is then log-transformed using equation 1.

$$I_{log} = \log(I_{crop} + \epsilon) \quad (1)$$

The value for  $\epsilon$  in equation 1 is determined using equation 2 which basically computes the average of standard deviations calculated at neighborhoods  $S_{xy}$  of size  $m \times n$  around each pixel  $(x,y)$  with mean  $\bar{p}$ . The values of  $m$  and  $n$  are computed as  $0.5 * M$  and  $0.5 * N$ , respectively, where  $M$  and  $N$  are the number of rows and columns in the input image, respectively. The value is then normalized by dividing it with the maximum intensity in the image.

$$\epsilon = \left( \frac{1}{MN} \sum_{x=0}^M \sum_{y=0}^N \sqrt{\frac{\sum_{p \in S_{xy}} (p - \bar{p})^2}{mn}} \right) / \max(I_{crop}) \quad (2)$$

The resulting image  $I_{log}$  is then applied with Contrast Limited adaptive histogram equalization using the method proposed in [21] by setting the tile grid size to  $5 \times 5$ , the distribution to

'uniform', number of bins to 256 and clip-limit threshold to 0.005. This is followed by contrast stretching to form the processed image  $I_{processed}$ .

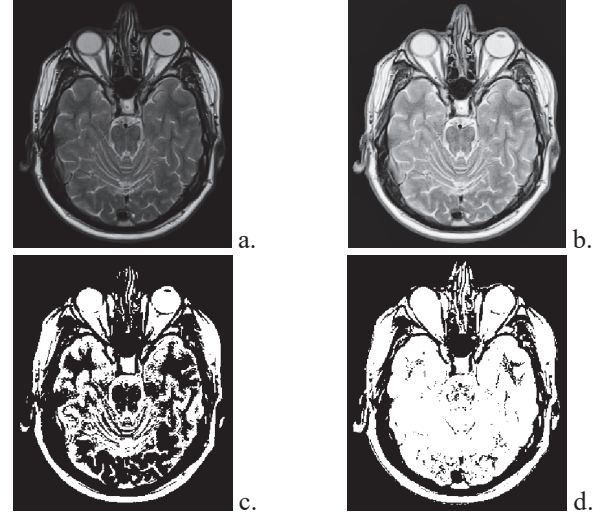


Fig. 2. Illustration of skull Thresholding. a) Original image cropped to  $Blob_{mask}$ . b)  $I_{log}$  c) Thresholding of  $I_{log}$  at high threshold  $T_{hard}$  d) Thresholding of  $I_{log}$  at low threshold  $T_{soft}$ .

### 1) Skull Thresholding

To extract the skull, the  $I_{processed}$  image needs to be binarized at an optimal threshold value such that the skull and the brain tissues are clearly separated from one another. However, finding an optimal threshold is a challenging task as not all MRI scans have the same intensity distribution. The proposed method uses a heuristic to find a threshold  $T_{raw}$  by taking the average value of the second smallest and second largest intensities in the input image  $I_{processed}$ . The smallest and the largest intensities are omitted to account for the background and white noise. The  $T_{raw}$  threshold is then used to calculate a soft threshold  $T_{soft}$  using equation 3 where  $\mu$  is a manual offset taking values between [0.9 to 1.5] with a step of 0.1 to adjust or fine tune the results in case some brain part is shown as skull part or vice versa.

$$T_{soft} = T_{raw} \times \mu \quad (3)$$

Binarization at the threshold value  $T_{soft}$  does not always guarantee a separation between the skull and the brain, especially in images with anomalies like tumors etc. In order to deal with this problem, the proposed method uses a second stricter threshold as given in equation 4. It uses an offset of 0.1 from  $T_{soft}$  to ensure that the skull and the brain part are separated by sacrificing connectivity within the skull and the brain regions respectively.

$$T_{hard} = T_{soft} + 0.1 \quad (4)$$

The offset value of 0.1 is chosen to be the smallest positive value at which the reconstruction of the connected components from  $T_{hard}$  using the  $Outline_{mask}$  generates a hole at the center of the reconstruction which simply represents the skull. The existence of the enclosed hole is detected by taking a difference from a morphologically hole-filled version of the same. Binarization at the threshold value  $T_{hard}$  guarantees a separation between the skull and the brain but does not guarantee the protection of the complete portions of the regions of interest. The proposed method uses both thresholds  $T_{soft}$  and

$T_{hard}$  together to identify the skull mask from the input image  $I_{Processed}$  as discussed in the following section.

## 2) Removing the soft brain tissue

Figure 4 illustrates the use of the  $T_{hard}$  and  $T_{soft}$  thresholds for binarization of the  $I_{Processed}$  image. The binarized images resulting from the two different thresholds are referred to as  $I_{hard}$  and  $I_{soft}$ , respectively. In order to remove the soft brain tissue, the proposed method uses a *cross-mask* as shown in figure 3a wherein the horizontal and the vertical bars have lengths equal to 70% of the width and height of the  $Blob_{mask}$  respectively and the inner filled circle has a radius equal to 40% of the  $minimum(height, width)$  of the  $Blob_{mask}$ . The cross-mark is centered at the center of the  $Blob_{mask}$ . The fractions 70% and 40% for designing the *cross-mask* shape are selected such that when it is overlapped on  $I_{hard}$  or  $I_{soft}$  versions of an axial scan, the *cross-mask* won't overlap the eye sockets and the skull even when the axial scans could be in lateral or collateral position. This is based on the observation that for axial scans, the skull area lies at a distance of more than 70% of the height and width of the  $Blob_{mask}$  respectively from the centroid of the  $Blob_{mask}$  along the two axis. Moreover, the binary components representing the eye sockets (at their extreme intrusion inside the skull) lie at a distance of more than 40% of the  $minimum(height, width)$  of the  $Blob_{mask}$  from the centroid of the  $Blob_{mask}$  along the two axis.

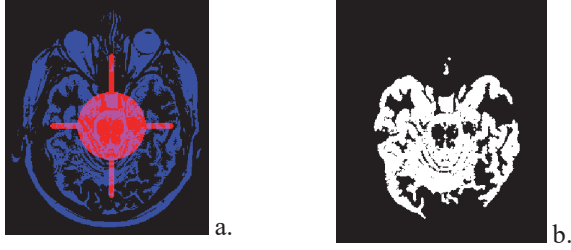


Fig. 3. Illustration of brain tissue removal. a) Overlap of *cross-mask* on  $I_{hard}$ . b) Connected components removed from  $I_{hard}$  after *cross-mask* reconstruction.

Initially, an *AND* operation is performed between *cross-mask* image and the  $I_{hard}$  image to extract the brain seeds. These seeds are used to reconstruct all the connected components from the  $I_{hard}$  image as shown in figure 3b. The connected components thus identified are subtracted from the  $I_{soft}$  image to get a new image  $I_{soft\_pruned}$ , as shown in figure 4a. This process aims to break any connections between the skull and the brain in the  $I_{soft}$  image, especially in images with anomalies like tumors. Next, the same process is repeated for the *cross-mask* image and the  $I_{soft\_pruned}$  image to remove major portions of the brain from  $I_{soft\_pruned}$  image, as shown in figure 4 and form a new image referred to as  $I_{skull\_raw}$ . Now, the *Outline\_mask* as discussed in section A is used to extract skull seeds from the  $I_{skull\_raw}$  image and the  $I_{hard}$  image, as shown in equation 5.

$$S_{seeds} = Outline_{mask} \wedge (I_{skull\_raw} | I_{hard}) \quad (5)$$

$$I_{skull\_base} = reconstruct(S_{seeds}, (I_{skull\_raw} | I_{hard})) \quad (6)$$

## 3) Generation of Skull-Mask

The seeds image  $S_{seeds}$  is used to reconstruct the minimal skull referred as  $I_{skull\_base}$  image from the *OR* of  $I_{skull\_raw}$  and  $I_{hard}$  images, as shown in equation 6 and figure 4d. The  $I_{hard}$  image is again used in this process just to be sure that at least a basic representation of the skull is protected as it could be

removed from the  $I_{soft}$  image while removing the soft brain tissue due to stubborn connections between the skull and the brain.

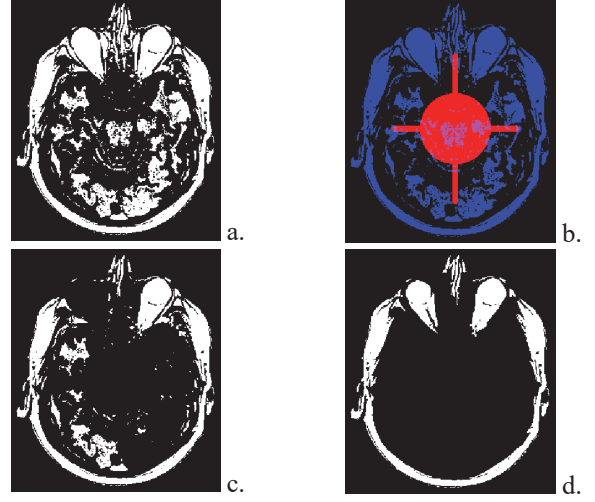


Fig. 4. Illustration of skull protection. a)  $I_{soft\_pruned}$  b) Overlap of *cross-mask* and  $I_{soft\_pruned}$  c)  $I_{skull\_raw}$  d)  $I_{skull\_base}$  obtained by reconstruction using *Outline\_mask* on the *OR* of ( $I_{skull\_raw}$  and  $I_{hard}$ ).

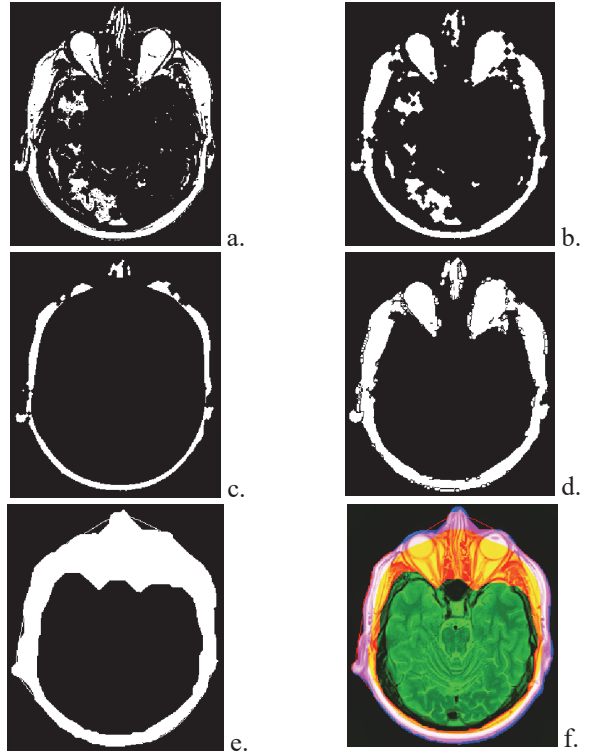


Fig. 5. Illustration of skull mask generation. a)  $I_{skull}$  b) morphological opening and closing on  $I_{skull}$ . c) seed for maximal skull obtained by the *AND* of *Outline\_mask* and  $I_{skull}$  d) reconstruction of  $I_{skull}$  using (c). e)  $I_{skull\_mask}$ . f) overlap of original image  $I_{skull\_mask}$  and *Outline\_mask*.

The  $I_{skull\_base}$  image is now *Ored* with  $I_{skull\_raw}$  followed by morphological opening and closing to get image  $I_{skull}$ , as shown in figures 5a and 5b. The structuring element used here involves a disk mask of size  $m \times n$  where  $m = n = 1\%$  of the  $minimum(M, N)$ . Now, to extract the seeds representing the maximal skull, as shown in figure 5c, an *AND* operation is performed between the *Outline\_mask* and  $I_{skull}$ . These seeds are then used to reconstruct the maximal skull from  $I_{skull}$ , as shown

in figure 5d. The extracted mask image is then morphologically processed by opening, skeletonization, and pruning, followed by thickening operations to represent the final skull mask  $I_{skull\_mask}$  as shown in figures 5e and 5f.

### C. Skull Subtraction and Brain Extraction

In order to extract the brain, the skull needs to be removed from image  $I_{processed}$  generated in section B. To do so, the outer area of the  $I_{skull\_mask}$  is filled and the result is complemented to form a mask referred to as  $I_{brain\_raw}$  representing the inner soft tissue containing brain and some cerebral fluid, as shown in figure 6. The grayscale brain image  $I_{brain}$  is then extracted using equation 7, which simply performs an *AND* operation over the  $I_{brain\_raw}$  and the  $I_{processed}$  images, as illustrated in figure 7a.

$$I_{brain} = I_{brain\_raw} \wedge I_{processed} \quad (7)$$

The  $I_{brain}$  image often contains the cerebral fluid in its outer area which needs to be trimmed. This refinement is performed as follows.



Fig. 6. Illustration of Skull subtraction. a) External fill of  $I_{skull\_mask}$ . b)  $I_{brain\_raw}$  obtained by complement of (a).

#### 1) Brain Thresholding

To extract the proper brain-mask, the  $I_{brain}$  image needs to be binarized at an optimal threshold value such that the cerebral fluid and the brain tissue are separated. However, it is difficult to always find an optimum threshold due to intensity variations in different types of MRI scans. The proposed method refines the brain tissue as follows. The image  $I_{brain}$  is transformed with contrast limited adaptive histogram equalization using the method proposed in [21] by setting the tile grid size to  $10 \times 10$ , the distribution to 'rayleigh', number of bins to 11 and clip-limit threshold to 0. The small number of bins is used to simplify the threshold selection. Further, the resulting image is contrast stretched to form the  $I_{brain\_processed}$  image. The  $I_{brain\_processed}$  is binarized to  $I_{brain\_binary}$  using the threshold  $T_\epsilon$  which is computed using equation 8, where  $mid_{brain}$  is the average of the second smallest and the largest intensities in the  $I_{brain\_processed}$  image and  $std_{brain}$  is the standard deviation of the non-zero pixels in the  $I_{brain\_processed}$  image normalized by the largest intensity in the same image. The smallest intensity is omitted to account for the background. The parameter  $\delta$  is a manual parameter used to fine tune the brain segment area. The value for  $\delta$  ranges between  $[0.2 \text{ to } 2]$  with a step of 0.2 due to the intensity variations in the brain regions of the multimodal scans. The illustration of this process is shown in figure 7b.

$$T_\epsilon = mid_{brain} + std_{brain} * \delta \quad (8)$$

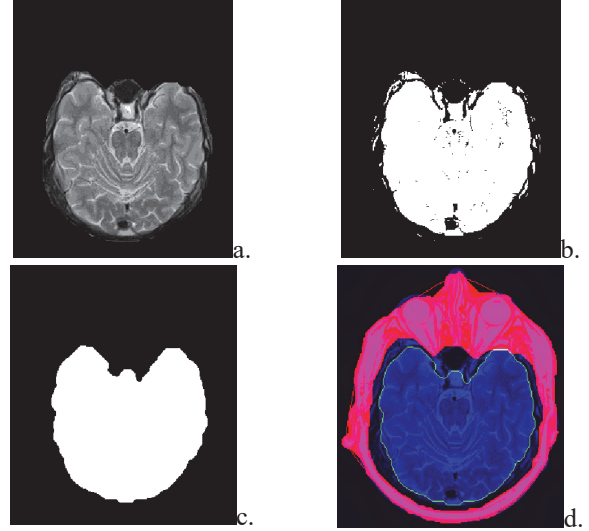


Fig. 7. Illustration of Brain extraction. a)  $I_{brain}$  b)  $I_{brain\_binary}$  c)  $I_{brain\_mask}$  d) Final segmentation representing the brain and skull masks overlapped on the original image.

#### 2) Morphological Processing

The binarized brain image  $I_{brain\_binary}$  is passed through the following sequence of morphological operation to extract the final brain mask  $I_{brain\_mask}$ .

1. Opening: using the structuring element disk of size  $m \times n$  where  $m = n = 1\%$  of the  $minimum(M, N)$ .
2. Closing: using the structuring element disk of size  $m \times n$  where  $m = n = 1\%$  of the  $minimum(M, N)$ .
3. Hole filling: to fill in empty regions in the brain induced due to binarization.
4. Opening: using the structuring element disk of size  $m \times n$  where  $m = n = 2\%$  of the  $minimum(M, N)$ .

This process is illustrated in figure 7c.

## IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

To evaluate the proposed method, this paper uses four publicly available MRI datasets as listed in table II. The results of the proposed method are compared with nine state-of-the-art skull stripping methods on different datasets. The characteristics of the different methods considered here are summarized in table III. Although, the proposed method works with both individual MRI slices as well as full 3D volumes, the other methods considered here only work with 3D volumes. In this regard the lgg-segmentation [22] is converted from multiple individual slices to 3D volumes by simply stacking

TABLE II  
DATASETS USED

Name	SOURCE	MRI Modalities included
LGG-Segmentation	Pedano et al. [22]	Pre-contrast T1, Post-contrast T1, FLAIR
MSSEG	Commowick et al. [23]	T1, T1GADO, DP, T2, FLAIR
NFBS	Puccio et al. [24]	T1
Tumor	Cheng [30]	T1

TABLE III  
NON-FUNCTIONAL CHARACTERISTICS OF THE DIFFERENT METHODS

Name	Ease of Use	Modalities for which acceptable results are generated	Speed (Seconds)
ROBEX [17]	Easy	T1, T1GADO, DP	
SwissSkullStripper [12]	Easy	T1, T2, FLAIR, DP, T1GADO	
SkullStripper [29]	Few parameters to adjust	T1	
BEAST [26]	Too many parameters to adjust	T1	
Freesurfer [13]	Few parameters to adjust	T1	
AFNI [15]	Moderate	T1, T2, DP	
FSL-BET [27]	Easy	T1	
HD-BET [2]	Too many parameters to adjust	T1, FLAIR, T2, DP, T1GADO	
Brainsuite [3]	Few parameters to adjust	T1, pre-contrast, FLAIR, T1GADO, DP	
Proposed Method	Two parameters to adjust, which can be done visually.	T1, T2, pre/post-contrast T1, FLAIR, DP, T1GADO	

Fig. 8. Running time comparison between various methods.

them together. The rest of the datasets include only 3D volumes. The colors are used to annotate the brain masks detected by the various methods through the experimental results, as shown in the legends for plots in table III. All the methods were run on a workstation with eight cores and 16 GB RAM. The proposed method is implemented using MATLAB and the code is made available on the GitHub repository [here](https://github.com/bhatsajid/SkullStripping2023.git)<sup>1</sup>.

### A. Results on LGG-Segmentation Dataset

The LGG-Segmentation dataset consists of three types of scans viz pre-contrast T1, post-contrast T1 and FLAIR. Results are presented for a single case in each of the categories present in the dataset.

#### 1) Pre-contrast T1

The results generated for the pre-contrast T1 scan are shown in figure 9. Only seven out of the other nine methods generated

results on this input successfully. Methods SkullStripper and BEAST failed to generate any results here. It can be observed from figures 9a and 9b that the proposed method provided the best results for this input compared to the other methods by generating acceptable brain and skull masks.

#### 2) Post-contrast T1

The results generated for the post-contrast T1 scan are shown in figure 10. Only six out of the other nine methods generated results on this input successfully. Methods SkullStripper, Brainsuite, and BEAST failed to generate any results here. It can be observed from figures 10a and 10b that the proposed method generated the best results for this input compared to the other methods by generating acceptable brain and skull masks.

<sup>1</sup> <https://github.com/bhatsajid/SkullStripping2023.git>



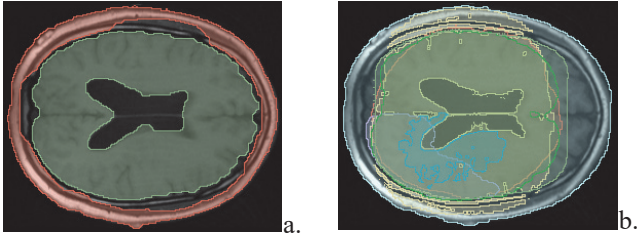


Fig. 9. Result on pre-contrast T1 scan [case id TCGA\_CS\_5395\_19981004] a) Result generated by proposed method b) Result generated by other methods.

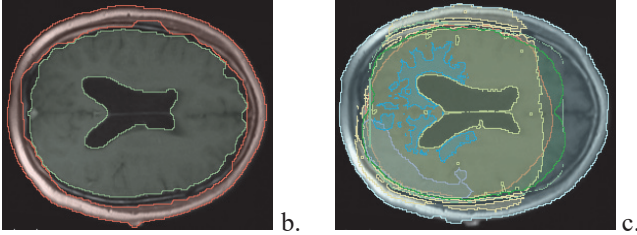


Fig. 10. Results on post-contrast T1 scan [case id TCGA\_CS\_5395\_19981004] a) Result generated by proposed method b) Results generated by other methods.

### 3) FLAIR

The results generated for the FLAIR scan are shown in figure 11. Only five out of the other nine methods generated results on this input successfully. Methods SkullStripper, FSL, Freesurfer, and BEAST failed to generate any results here. It can be observed from figures 11a and 11b that the proposed method along with Brainsuite generated the best results for this input compared to the other methods.

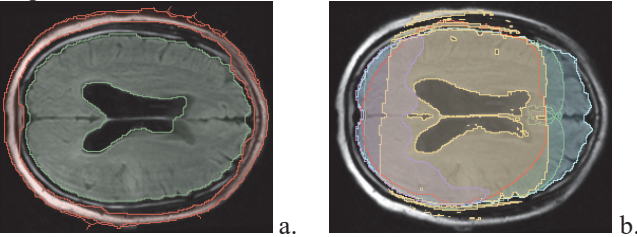


Fig. 11. Results on FLAIR scan [case id TCGA\_CS\_5395\_19981004] a) Result generated by proposed method b) Results generated by other methods.

## B. Results on MSSEG Dataset

The MSSEG dataset consists of five types of scans for each subject viz T1, T2, FLAIR, T1GADO, and DP. Results are presented for a single case in each of the categories presented in the dataset.

### 1) T1

The results generated for the T1 scan are shown in figure 12. Only seven out of the other nine methods generated results on this input successfully. Methods including AFNI and Freesurfer failed to generate any results here. It can be observed from figures 12a and 12b that all the methods in question except ROBEX generated acceptable results.

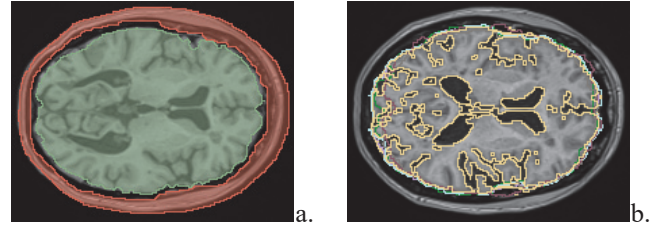


Fig. 12. Results on T1 scan [case id 07003SATH] a) Result generated by proposed method b) Results generated by other methods.

### 2) T1GADO

The results generated for the T1GADO scan are shown in figure 13. Only seven out of the other nine methods generated results on this input successfully. Methods including Freesurfer and BEAST failed to generate any results here. It can be observed from figures 13a and 13b that only the proposed method, SwissSkullStripper, HD-BET, and Brainsuite generated acceptable results for this input.

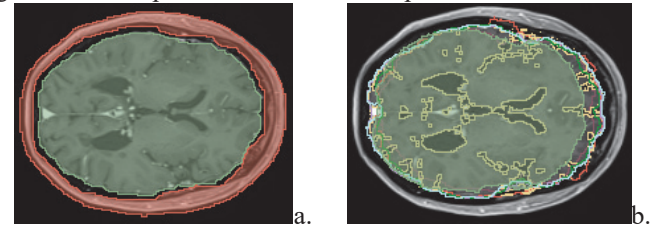


Fig. 13. Results on T1GADO scan [case id 07003SATH] a) Result generated by proposed method b) Results generated by other methods.

### 3) DP

The results generated for the DP scan are shown in figure 14. Only six out of the other nine methods generated results on this input successfully. Methods including Freesurfer, FSL and BEAST failed to generate any results here. It can be observed from figures 14a and 14b that all the methods in question except ROBEX generated acceptable results for this input.

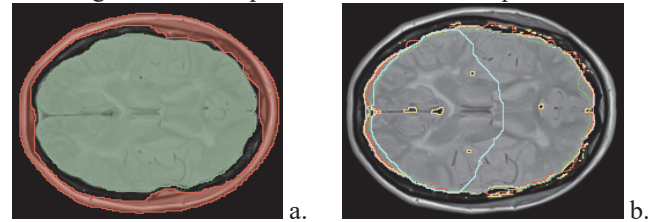


Fig. 14. Results on DP scan [case id 07003SATH] a) Result generated by proposed method b) Results generated by other methods.

### 4) T2

The results generated for the T2 scan are shown in figure 15. Only five out of the other nine methods generated results on this input successfully. Methods including Freesurfer and SkullStripper, FSL-BET, and Brainsuite failed to generate any results here. It can be observed from figures 15a and 15b that only the proposed method, SwissSkullStripper, AFNI and HD-BET generated acceptable results for this input.

### 5) FLAIR

The results generated for the FLAIR scan are shown in figure 16. Only six out of the other nine methods generated results on this input successfully. Methods including Freesurfer, FSL-BET, and ROBEX failed to generate any results here. It can be



observed from figures 16a and 16b that only the proposed method, Brainsuite, SwissSkullStripper and HD-BET generated acceptable results for this input.

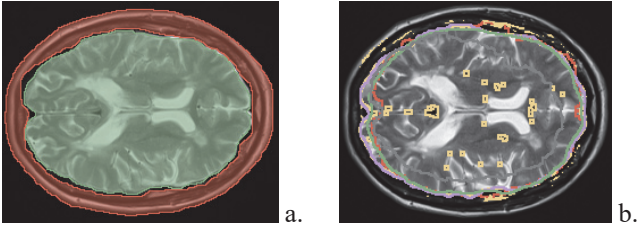


Fig. 15. Results on T2 scan [case id 07003SATH a) Result generated by proposed method b) Results generated by other methods.

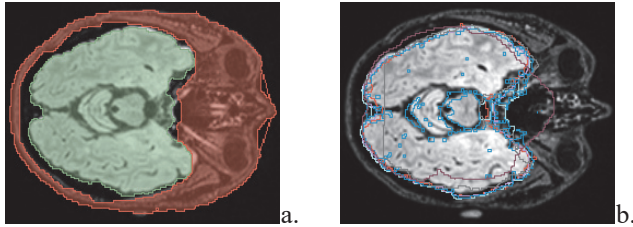


Fig. 16. Results on FLAIR scan [case id 07003SATH a) Result generated by proposed method b) Results generated by other methods.

### C. Results on NFBS Dataset

Sample result for a single case in this dataset is illustrated in figure 17. It can be observed from figures 17a and 17b that all methods except for FSL-BET and SkullStripper generate acceptable results for this input.

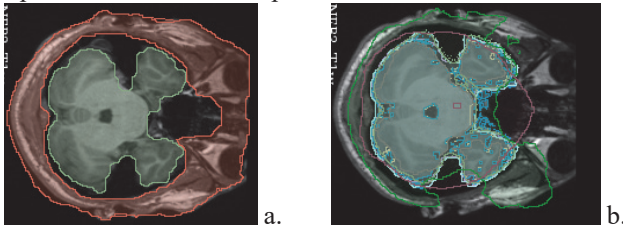


Fig. 17. Results on T1 scan [case id A00037848 a) Result generated by proposed method b) Results generated by other methods.

### A. Consolidated Results

To compare the overall performance of all the methods in question on all the MRI modalities, we compare the resulting brain masks of each of the methods to the ground truth brain masks. However, among the three datasets used in this paper, only one of the datasets i.e. the NFBS dataset consists of ground truth for T1 scans. For all other modes belonging to the datasets in question, we select 2 volumes from each mode at random and select 5 axial slices from each of the selected volumes at percentiles including 30, 40, 50, 60, and 70. This results in a total of  $9 \times 5 \times 5 = 225$  slices out of which 25 slices (from NFBS dataset) already have ground truth and 200 slices are segmented by 3 experts using the NFBS ground truth as reference, BEaST[26] as the baseline and 3D Slicer [40] as the segmentation editing tool. The final ground truth masks are then extracted from the intersection of the segmentations of the three experts for each slice. The performance of each of the methods on these selected slices is evaluated using various metrics as illustrated in table IV. The illustration of the evaluation approach used here is shown in figure 18. The scores for each method are averaged over all the 225 axial slices of all the MRI modalities to reflect the overall performance of the methods. It

can be observed from tables III and IV that the proposed method outperforms all the other methods in question when considering the overall performance of the methods on all the MRI modalities. Moreover, it can be observed from table III that the proposed method is among the very few methods including Brainsuite, SwissSkullStripper, and HD-BET that generated acceptable results on multiple categories of MRI modalities with only the proposed method generating acceptable results on all modalities. In terms of speed of generation of results, HD-BET is relatively slow and Brainsuite is the fastest to generate the results. However, the speed of the proposed method is comparable with the fastest methods in question that generated acceptable results.

### A. Results on Tumor Dataset

TABLE IV  
CONSOLIDATED PERFORMANCE OF THE METHODS

Name	Accuracy	Sensitivity	FMeasure	Precision	Specificity
ROBEX [17]	0.925	0.903	0.884	0.813	0.910
SwissSkullStripper [12]	0.938	0.840	0.835	0.850	<b>0.962</b>
SkullStripper [29]	0.820	0.801	0.766	0.719	0.819
BEaST [26]	0.868	0.836	0.652	0.618	0.850
Freesurfer [13]	0.673	0.665	0.691	0.622	0.614
AFNI [15]	0.920	0.923	0.907	0.891	0.941
FSL-BET [27]	0.898	0.854	0.709	0.649	0.903
HD-BET [2]	0.940	0.930	0.912	0.873	0.921
BrainSuite [3]	0.982	0.956	0.900	0.844	0.912
Proposed Method	<b>0.984</b>	<b>0.962</b>	<b>0.917</b>	<b>0.896</b>	0.960

The brain tumor dataset [30] consists of T1-weighted contrast-enhanced images with three kinds of brain tumor: *meningioma*, *glioma*, and *pituitary tumor*. Since the dataset consists of individual images and not volumes, the results here are shown only for the proposed method as the other methods in question work with only volume data. Results are shown for a single case of each type of tumor in figure 19.

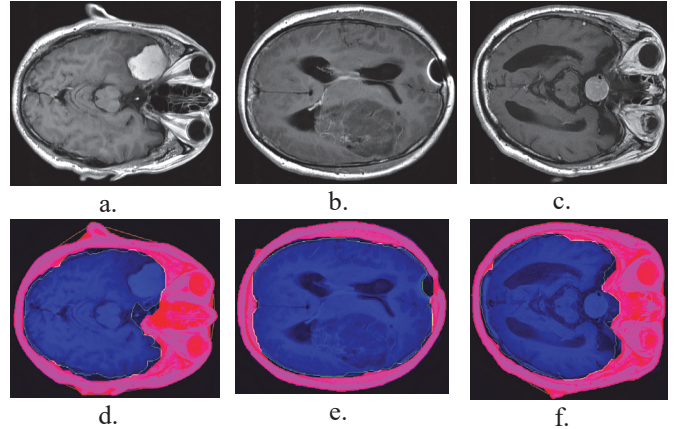


Fig. 19. Results on the brain tumor dataset images. First row represents original images while second row represents the skull and brain extraction results using the proposed method.

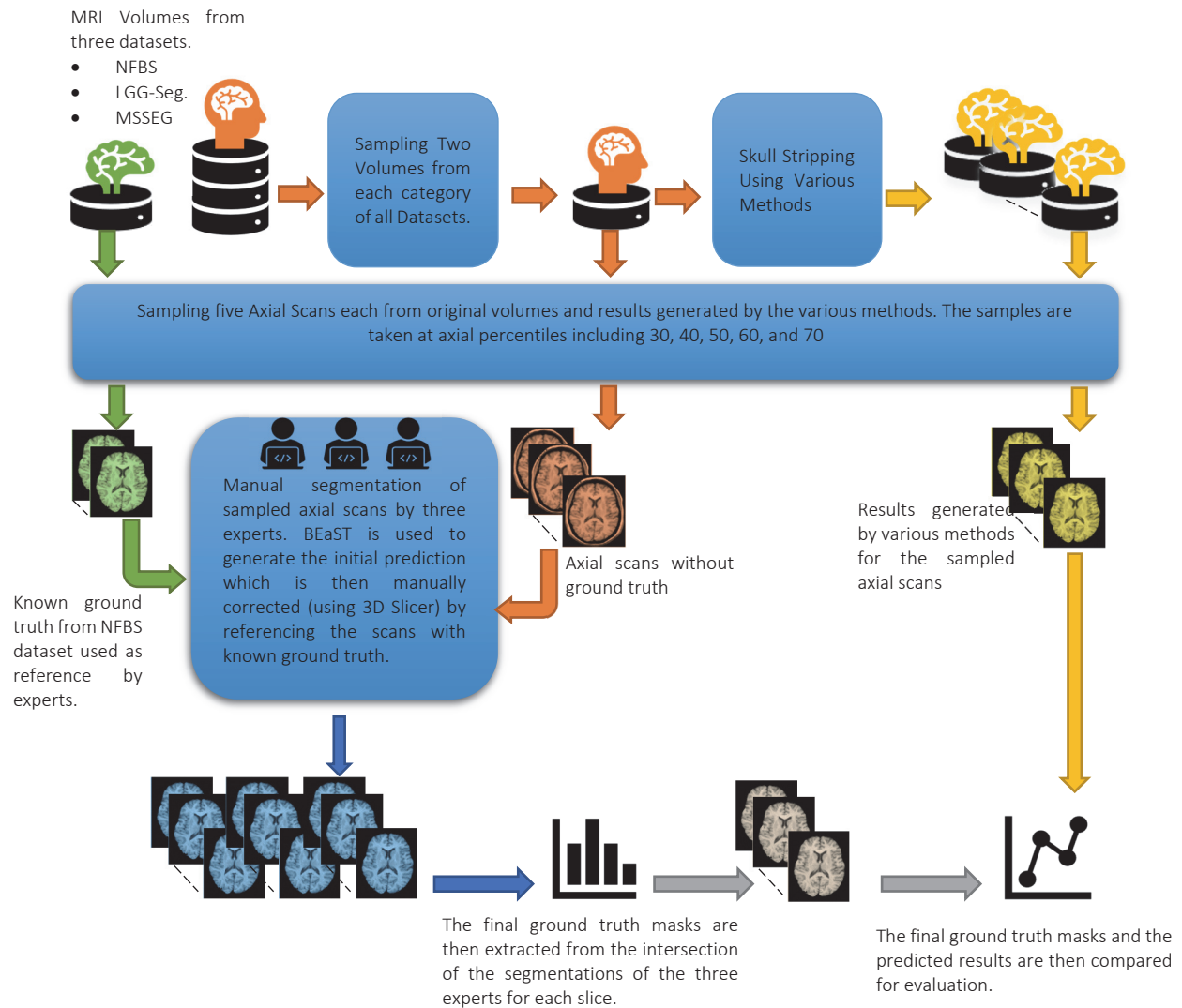


Fig. 18: Illustration of the evaluation approach.

In figure 19, the first column is an image with *meningioma*, the second column image has *glioma*, and the third column image has *pituitary tumor*. It can be seen from figure 18 that the proposed method identifies the skull and the brain masks very well even in presence of abnormalities like tumors.

## V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel skull stripping method for different types of MRI modalities. The proposed method is relatively easy to use because it requires adjustment of a single fine-tuning parameter. From the experimental results considering all datasets, it can be observed that in general only the proposed method, Brainsuite, SwissSkullStripper, and HD-BET generated acceptable results on multiple categories of MRI modalities. In terms of speed, HD-BET is relatively slow and Brainsuite is the fastest to generate the results. It can be argued that the proposed method is comparable to the best methods among the state-of-the-arts in skull stripping domain. However, unlike other methods, it works well on different types of MRI modalities. Moreover, it generates the skull mask along with the brain mask that can be used to study various skull pathologies.

Along the future direction, the authors aim to apply the proposed method on the whole LGG-Segmentation dataset [22] and use the resulting dataset to train supervised machine-learning models for detection of tumors using the *skull-stripped* T1 and FLAIR brain modalities.

## REFERENCES

- [1] Lavini, C. (2001). Do different MRI scanners produce different fMRI results?. In *Proceedings of the 9th Annual Meeting of ISMRM, Glasgow*.
- [2] Isensee, F., Schell, M., Pflueger, I., Brugnara, G., Bonekamp, D., Neuberger, U., ... & Kickingereder, P. (2019). Automated brain extraction of multisequence MRI using artificial neural networks. *Human brain mapping, 40*(17), 4952-4964.
- [3] Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A., & Leahy, R. M. (2001). Magnetic resonance image tissue classification using a partial volume model. *NeuroImage, 13*(5), 856-876.
- [4] Swiebocka-Wiek, J. (2016, September). Skull stripping for MRI images using morphological operators. In *IFIP International Conference on Computer Information Systems and Industrial Management* (pp. 172-182). Springer, Cham.
- [5] Ezhilarasan, K., Praveenkumar, S., Somasundaram, K., Kalaiselvi, T., Magesh, S., Kiruthika, S., & Jeevarekha, A. (2021). Automatic brain extraction from MRI of human head scans using Helmholtz free energy principle and morphological operations. *Biomedical Signal Processing and Control, 64*, 102270.

- [6] Fatima, A., Shahid, A. R., Raza, B., Madni, T. M., & Janjua, U. I. (2020). State-of-the-Art Traditional to the Machine-and Deep-Learning-Based Skull Stripping Techniques, Models, and Algorithms. *Journal of Digital Imaging*, 33(6), 1443-1464.
- [7] Abd El-Kader, S., Morse, M., Abo-Elsoud, M. E. A., & Mokhtar, R. (2016). Improvement Skull Stripping Algorithm of MRI Brain Images Based on Fuzzy Morphological Operation. *Int. J. Sci. Res*, 5(1), 13-17.
- [8] Somasundaram, K., & Kalaiselvi, T. (2011). Automatic brain extraction methods for T1 magnetic resonance images using region labeling and morphological operations. *Computers in biology and medicine*, 41(8), 716-725.
- [9] Muhammet Üsame, Ö. Z. İ. Ç., ÖZBAY, Y., EKMEKÇİ, A. H., & KIVRAK, A. S. Skull Stripping on Brain MR Images Using Prior Knowledge.
- [10] Abd El-Kader, S., Morse, M., Abo-Elsoud, M. E. A., & Mokhtar, R. (2016). Improvement Skull Stripping Algorithm of MRI Brain Images Based on Fuzzy Morphological Operation. *Int. J. Sci. Res*, 5(1), 13-17.
- [11] Tehoketch Kebir, S., Mekaoui, S., & Bouhedda, M. (2019). A fully automatic methodology for MRI brain tumour detection and segmentation. *The Imaging Science Journal*, 67(1), 42-62.
- [12] Tao, X., & Chang, M. C. (2010, March). A skull stripping method using deformable surface and tissue classification. In *Medical Imaging 2010: Image Processing* (Vol. 7623, p. 76233L). International Society for Optics and Photonics.
- [13] Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *Neuroimage*, 22(3), 1060-1075.
- [14] Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, 17(3), 143-155.
- [15] Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3), 162-173.
- [16] Leung, K. K., Barnes, J., Modat, M., Ridgway, G. R., Bartlett, J. W., Fox, N. C., ... & Alzheimer's Disease Neuroimaging Initiative. (2011). Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. *Neuroimage*, 55(3), 1091-1108.
- [17] Iglesias, J. E., Liu, C. Y., Thompson, P. M., & Tu, Z. (2011). Robust brain extraction across datasets and comparison with publicly available methods. *IEEE transactions on medical imaging*, 30(9), 1617-1634.
- [18] Kobashi, S., Fujimoto, Y., Ogawa, M., Ando, K., Ishikura, R., Kondo, K., ... & Hata, Y. (2007, November). Fuzzy-ASM based automated skull stripping method from infantile brain MR images. In *2007 IEEE International Conference on Granular Computing (GRC 2007)* (pp. 632-632). IEEE.
- [19] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- [20] Bradley, D., & Roth, G. (2007). Adaptive thresholding using the integral image. *Journal of graphics tools*, 12(2), 13-21.
- [21] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. *Graphics gems*, 474-485.
- [22] Pedano, N., Flanders, A. E., Scarpone, L., Mikkelsen, T., Eschbacher, J. M., Hermes, B., & Ostrom, Q. (2016). Radiology data from the cancer genome atlas low grade glioma [TCGA-LGG] collection. *Cancer Imaging Arch*, 2.
- [23] Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., ... & Barillot, C. (2018). Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Scientific reports*, 8(1), 1-17.
- [24] Puccio, B., Pooley, J. P., Pellman, J. S., Taverna, E. C., & Craddock, R. C. (2016). The preprocessed connectomes project repository of manually corrected skull-stripped T1-weighted anatomical MRI data. *Gigascience*, 5(1), s13742-016.
- [25] Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9(2), 179-194.
- [26] Eskildsen, S. F., Coupé, P., Fonov, V., Manjón, J. V., Leung, K. K., Guizard, N., ... & Alzheimer's Disease Neuroimaging Initiative. (2012). BEaST: brain extraction based on nonlocal segmentation technique. *NeuroImage*, 59(3), 2362-2373.
- [27] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208-S219.
- [28] Bauer S., Fejes T., Reyes M., January-December. (2012). A Skull-Stripping Filter for ITK". *Insight Journal*.
- [29] Xiaodong Tao, Ming-ching Chang. (2010). A Skull Stripping Method Using Deformable Surface and Tissue Classification. *SPIE Medical Imaging, San Diego, CA*
- [30] Cheng, J. (2017): brain tumor dataset. *figshare*. Dataset. <https://doi.org/10.6084/m9.figshare.1512427.v5>
- [31] Kalavathi, P., & Prasath, V. S. (2016). Methods on skull stripping of MRI head scan images—a review. *Journal of digital imaging*, 29(3), 365-379.
- [32] Roslan, R., Jamil, N., & Mahmud, R. (2010, November). Skull stripping of MRI brain images using mathematical morphology. In *2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)* (pp. 26-31). IEEE.
- [33] Somasundaram, K., Gayathri, S. P., Rajeswaran, R., & Manjiri, D. (2018). Fetal brain extraction from magnetic resonance image (MRI) of human fetus. *The Imaging Science Journal*, 66(3), 133-138.
- [34] Hoopes, A., Mora, J. S., Dalca, A. V., Fischl, B., & Hoffmann, M. (2022). SynthStrip: skull-stripping for any brain image. *NeuroImage*, 260, 119474.
- [35] Lucena, O., Souza, R., Rittner, L., Frayne, R., & Lotufo, R. (2019). Convolutional neural networks for skull-stripping in brain MR imaging using silver standard masks. *Artificial intelligence in medicine*, 98, 48-58.
- [36] Thakur, S. P., Doshi, J., Pati, S., Ha, S. M., Sako, C., Talbar, S., ... & Bakas, S. (2020, May). Skull-stripping of glioblastoma MRI scans using 3D deep learning. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I* (pp. 57-68). Cham: Springer International Publishing.
- [37] Hwang, H., Rehman, H. Z. U., & Lee, S. (2019). 3D U-Net for skull stripping in brain MRI. *Applied Sciences*, 9(3), 569.
- [38] Rehman, H. Z. U., Hwang, H., & Lee, S. (2020). Conventional and deep learning methods for skull stripping in brain MRI. *Applied Sciences*, 10(5), 1773.
- [39] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.
- [40] Pieper, S., Halle, M., & Kikinis, R. (2004, April). 3D Slicer. In *2004 2nd IEEE international symposium on biomedical imaging: nano to macro (IEEE Cat No. 04EX821)* (pp. 632-635). IEEE.
- [41] Walczak, M., Nalepa, J., Kawulok, M., Dudzik, W., & Smolka, B. (2018, May). Evolutionary cortical surface segmentation. In *Real-Time Image and Video Processing 2018* (Vol. 10670, pp. 104-114). SPIE.