Robust Skull Stripping of Clinical Glioblastoma Multiforme Data

William Speier, Juan E. Iglesias, Leila El-Kara, Zhuowen Tu, and Corey Arnold

University of California, Los Angeles, USA

Abstract. Skull stripping is the first step in many neuroimaging analyses and its success is critical to all subsequent processing. Methods exist to skull strip brain images without gross deformities, such as those affected by Alzheimer's and Huntington's disease. However, there are no techniques for extracting brains affected by diseases that significantly disturb normal anatomy. Glioblastoma multiforme (GBM) is such a disease, as afflicted individuals develop large tumors that often require surgical resection. In this paper, we extend the ROBEX skull stripping method to extract brains from GBM images. The proposed method uses a shape model trained on healthy brains to be relatively insensitive to lesions inside the brain. The brain boundary is then searched for potential resection cavities using adaptive thresholding and the Random Walker algorithm corrects for leakage into the ventricles. The results show significant improvement over three popular skull stripping algorithms (BET, BSE and HWA) in a dataset of 48 GBM cases.

1 Introduction

Automatic whole-brain extraction (known as skull stripping) from magnetic resonance images (MRI) is the first element of most neuroimaging pipelines. Therefore, its robustness is critical for the overall performance of the system. Many methods have been proposed to solve the problem and a good level of segmentation accuracy (overlap over 90% [1]) can usually be achieved with some parameter tuning. However, these methods are designed for MRI scans of brains that are either healthy or suffering from a disease that does not considerably alter the structure of the brain, such as Alzheimer's disease, Huntington's or multiple sclerosis. The segmentation accuracy drops substantially for scans of brains with tumors, contrast agents and/or resection cavities, which is typically the case in patients with glioblastoma multiforme (GBM).

The delineation of brains and resection cavities in volumes from patients with GBM is important in understanding the progression of the disease. Gross changes in morphology due to these processes have considerable effects on a patient's quality of life and ultimate outcome. Skull stripping ensures that non-brain voxels do not erroneously contribute to any subsequent quantitative analysis, and also improves registration accuracy to a template or related studies.

Most current methods require a voxel intensity distribution close to that of a normal brain. For example, the widely-used Brain Surface Extractor (BSE) [2] uses diffusion filtering, edge detection and a chain of morphological operations

to segment the brain. This approach falters in images containing cancer-related findings which have strong borders that mislead the edge detection. Another approach is the Hybrid Watershed Algorithm (HWA) [3], part of the Freesurfer package, which combines a watershed algorithm, deformable surface and probabilistic atlas. This method has trouble with GBM images because the watershed algorithm assumes white matter connectivity and the atlas assumes a healthy brain shape. The Brain Extraction Tool (BET) [4] uses a deformable model grown based on voxel intensity. Cancer-related deformities hinder the evolution of the mesh and false positive regions appear around the eyes and brain stem. This problem may be ameliorated using a two-pass scheme (BET*) where the preliminary mask is used to align the brain to an atlas for a second pass by BET.

To the best of our knowledge, no existing technique robustly extracts the brain from MR images of individuals afflicted with GBM. The contribution of this paper is a method of segmenting brains with tumors, contrast agents and resection cavities to characterize GBM patients in support of research and clinical care. We address this problem by extending a robust skull stripping algorithm (ROBEX) [5] that extracts the brain using a strong prior on normal anatomy. We augment this method using convexity analysis, constrained distance transform and the Random Walker algorithm [6] to accurately find the brain boundary in the presence of tumors and surgical resection cavities. The proposed technique outperforms BSE, BET and HWA on a set of 48 volumes of subjects with GBM.

2   Materials: The Labeled MRI Dataset

In this work, 113 heterogeneous T1-weighted MRI volumes from individuals afflicted by GBM were collected. Under the supervision of a neuroradiologist, the brain boundary (including the cerebellum) was manually delineated by the third author in 48 of these cases. Of the 48 cases, 27 contained contrast enhancement, 23 contained resection cavities and 28 contained active tumor. The remaining 65 scans were only used for tuning parameters by visiual inspection. The scans were acquired with several acquisition protocols with diverse voxel sizes. In-plane resolution (axial in all cases) ranges from 0.86 to 0.94mm. The slice separation was 3mm in 40 cases and 6mm in the other eight labelled cases.

Figure 1a shows a slice from a scan in our dataset that illustrates the difficulty of the extraction problem. At a glance, the resection cavity and left ventricle appear to be joined when the two are actually separated by the intact ventricular lining. Automatically delineating this resection-ventricle border is critical to understanding the subsequent morphological evolution after surgical resection.

3   Methods

The steps of the proposed method are illustrated in Figure 1. First, an initial brain boundary is obtained using ROBEX (Fig. 1b). Then, resection cavity candidates are found using adaptive thresholding (Fig. 1c). Each candidate is classified as either "cavity/undersegmentation" or "cavity and ventricle" using
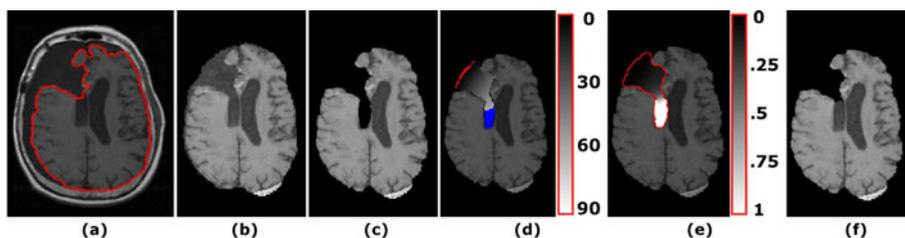
Fig. 1. Outline of the proposed algorithm. a) Axial slice of a T1 MRI with a resection cavity. b) Output from ROBEX, which leaves the cavity in the segmentation. c) Output from adaptive thresholding. d) Constrained distance transform within cavity/ventricle (in mm) and resulting seeds for Random Walker (positive in blue, negative in red). e) Probability map from Random Walker within the cavity/ventricle (marked in red). f) Final output, obtained by thresholding e).

convexity analysis. In the former case, the whole candidate region is removed from the brain mask. In the latter, a constrained distance transform (Fig. 1d) is used to seed Random Walker (Fig. 1e) which separates the region into cavity and ventricle (Fig. 1f). These steps are now discussed in detail.

3.1   Robust Brain Extraction

For the sake of completeness, a short description of ROBEX is included here. Designed to work with T1-weighted MRI, ROBEX applies signal intensity standardization and bias field correction to the input scan and then feeds it to a random forest classifier [7]. The random forest is trained to identify voxels on the brain surface. When applied, it generates a volume in which intensity represents the log-likelihood that a voxel is located on the boundary. Then, a shape model [8] of brain is fit to the output, maximizing the sum of the log-likelihood over the fit surface. This output is refined by allowing a small free deformation outside the model, which is optimized using the max-flow min-cut algorithm [9]. A sample output from ROBEX is displayed in Figure 1b.

ROBEX is very robust against tumors in the brain because it focuses on finding the brain boundary. However, contrast agents used in brain cancer imaging can bias the intensity standardization step as it is based on min/max normalization. We solve this problem by cropping the rightmost 5% of the histogram when computing the maximum. This value is high enough to remove all hyperintense tumor enhancement and residual skull, but low enough to retain some of the white matter. With regard to resection cavities, ROBEX will close the mask around them in most cases (unless they are very shallow) due to the fact that it is heavily based on of model of healthy brain. A solution to this problem is described in the following sections.

3.2   Adaptive Thresholding

Resection cavities correspond to low-intensity voxels in T1-weighted MRI because they are filled with cerebrospinal fluid (CSF). A simple thresholding

operator typically suffices to separate most of the CSF voxels from the rest of the brain matter, which is the principle behind some skull stripping methods [10,11]. However, algorithms of this type may fail in the presence of tumors and contrast agents that can adversely influence the global threshold. Instead, we use an adaptive scheme in which τ, the threshold value, is space-dependent:

$$\tau(x,y,z) = \frac{\sigma_{glob}^2}{\sigma_{loc}^2(x,y,z) + \sigma_{glob}^2}\tau_{glob} + \frac{\sigma_{loc}^2(x,y,z)}{\sigma_{loc}^2(x,y,z) + \sigma_{glob}^2}\tau_{loc}(x,y,z)$$

where σ_{loc}^2 is the local image intensity variance computed in a sphere of radius r_{loc} around each pixel, σ_{glob}^2 is the global variance across the scan within the preliminary brain mask, and τ_{glob} and τ_{loc} are a global and a local thresholds. The global threshold is the average of the median intensity in the preliminary brain mask and the minimum after cropping the leftmost 5% of the histogram. The local threshold is the mean intensity in the sphere around each pixel. Compared to using τ_{glob} alone, incorporating a local threshold produces a more accurate boundary, especially around resection cavities. After thresholding, the mask is smoothed by morphological closing with a spherical element (radius 5 mm), filling holes and removing islands. Sample output is displayed in Figure 1c.

3.3 Detecting Cavities

The next step in the pipeline is detecting resection cavities left in the segmentation by ROBEX. For each connected component of the negated mask, a shape convexity index $c \in (0,1]$ is computed as the ratio between its volume and that of its convex hull. The components for which $c > c_{min}$ are assumed to correspond to resection cavities, and the output mask is set to zero in that region. If $c < c_{min}$, leakage into the ventricles is suspected, and the candidate is fed to the next step of the pipeline. It is fairly easy to find a value of c_{min} that separates the two classes because the convexity drops considerably in the ventricles.

3.4 Separating Resection Cavities from Ventricles

Although a resection cavity and the ventricles are not connected in reality, it is often difficult to discern the border between the two due to the lack of contrast (both contain CSF), noise and limitations in resolution. Given a set of pre-labeled voxels (seeds), Random Walker can make this delineation by determining the probability of randomly reaching a seed from each unlabeled voxel.

Random Walker was conceived as a semi-automated segmentation method with manually placed seeds, but this work uses automatic seeding. A natural choice for the negative seed is the intersection of the cavity/ventricle candidate and the boundary of the mask of ROBEX, which is the most likely region to be part of the cavity. For the positive seed, we selected all points located more than d_{min} away from the negative seeds using a constrained distance transform [12]. The seed regions are overlaid in Figure 1d, the probability of resection cavity in the candidate region is shown in Figure 1e and the final segmentation (obtained by thresholding that probability at 0.5) is displayed in Figure 1f.
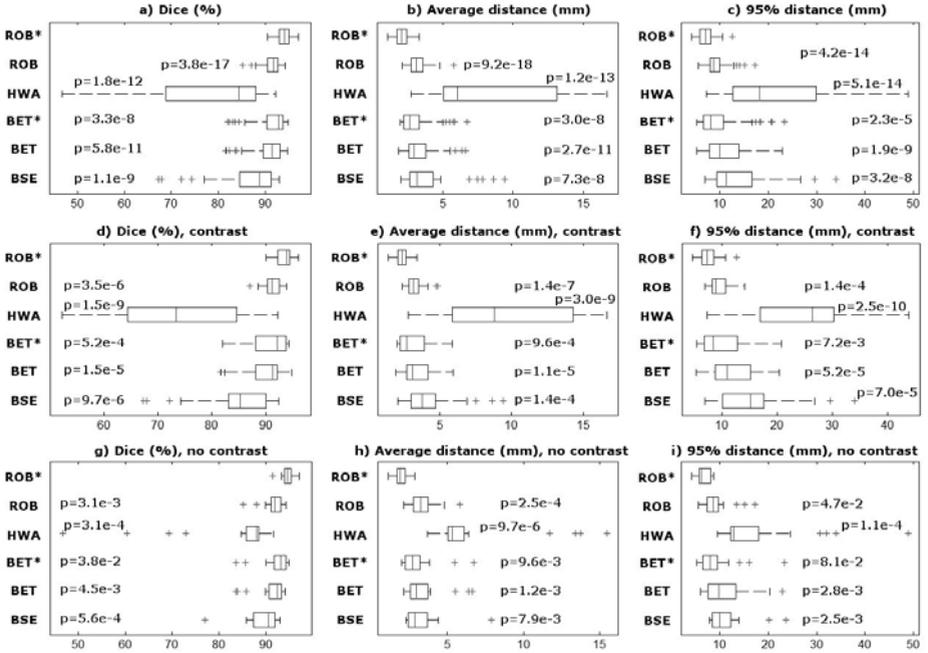
Fig. 2. Box plots and p-values for the different methods. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles and the whiskers extend to the most extreme data points not considered outliers i.e. within three standard deviations from the mean. The outliers are plotted individually. The p-values correspond to paired one tailed t-tests between our method and all of the others. ROB refers to ROBEX and ROB* to the extension proposed here.

4   Experiments and Results

4.1   Experimental Setup

As described in section 2, the 65 unlabeled scans were first used in pilot experiments to design the algorithm and tune parameters by visual inspection of the output. $r_{loc} = 4$mm, $c_{min} = 0.7$ and $d_{min} = 60$mm were found to be good values. The 48 scans for which ground truth was available were skull-stripped using BET v2.1 (one and two passes), BSE v09, HWA from FreeSurfer 5.0.0, ROBEX and the proposed extension. Default parameter values were used for BET, BSE and HWA. Three different metrics were used to measure the performance of the different methods: 1) the Dice overlap index $DICE(X,Y) = \frac{2|X \cap Y|}{|X|+|Y|}$; 2) the mean symmetric surface-to-surface distance; and 3) the 95% percentile of the symmetric surface-to-surface distance, a robust alternative to the Haussdorf (i.e., maximal) distance. Measure (3) is an effective complement to (1) and (2) as it captures fine boundary detail as opposed to gross overlap.
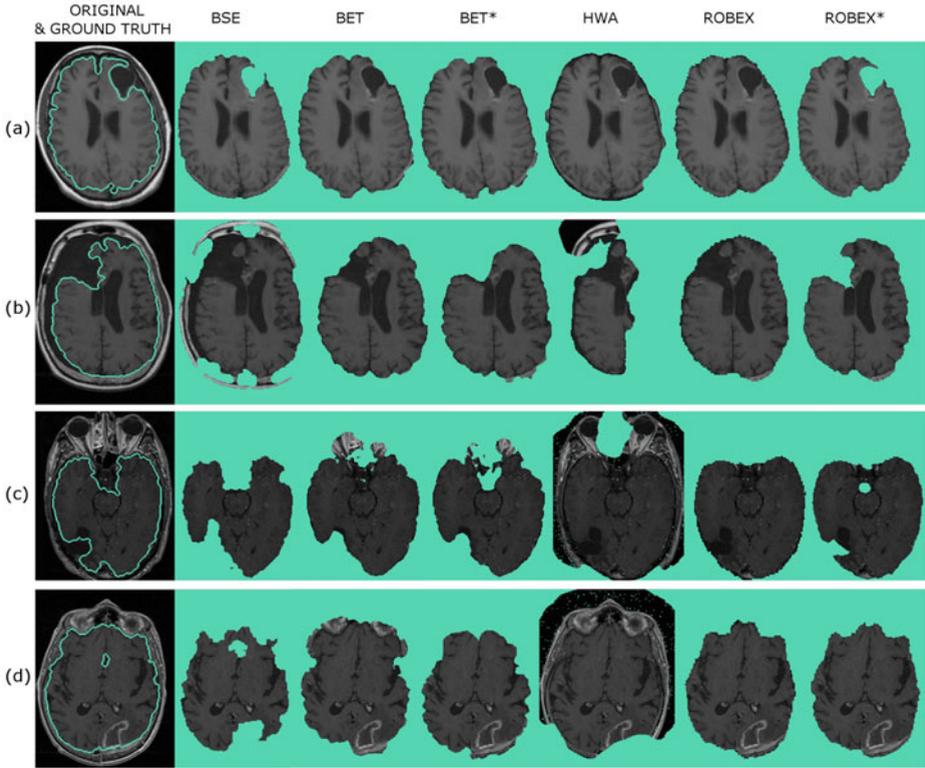
Fig. 3. Axial slices from sample outputs for the different extraction methods.
(a) Subject with resection cavity. (b-c) Subjects with resection cavities that interface
with vectricles. (d) Subject with an unresected tumor (note the hyperintense region
around the tumor due to contrast). ROBEX* represents the extension of ROBEX.

4.2   Results

Fig. 2 displays the box plots for each evaluation metric for all methods as well as
the p-values corresponding to one tailed t-tests comparing the proposed method
to the other approaches. When all test scans are considered (Fig. 2a-c), our
method significantly outperforms the selected comparative methods in terms
of the three metrics ($p \in [9.2e - 18, 2.3e - 5]$). Moreover, the robustness of
our method is apparent from the box plots. For example, the minimum DICE
overlap across the dataset is 67.3% for BSE, 81.6% for BET, 82.1% for two-pass
BET, 46.7% for HWA, 85.2% for ROBEX and 90.1% for the proposed algorithm.
When the scans with (Fig. 2d-f) and without contrast (Fig. 2g-i) are considered
separately, we observe that the improvement with respect to the other methods
is larger when contrast agents are present. However, in cases with no contrast,
the difference is still significant at $p = 0.05$ for all the methods and metrics
except for the 95% percentile of the distance of 2-pass BET.

Fig. 3 displays the output of the different methods for a few examples. BSE can reasonably handle resection cavities in some cases (Fig. 3a), but often fails to completely extract the brain boundary (Fig. 3b). BET consistently delineates the brain surface (Fig. 3b), but has difficulty identifying resection cavities as the deformable surface can: 1) grow beyond the cavity (Fig. 3a) or; 2) have insufficient flexibility to surround it (Fig. 3c). In addition, its known limitation of undersegmentation around the eyes is apparent (Fig. 3c-d). Two-pass BET improves the results from BET and has the second-highest performance after our method. HWA does not adapt well to the dataset as its watershed algorithm relies heavily on white matter intensity estimation, which is disturbed by tumors, contrast enhancement and resection cavities. Finally, the proposed method succeeds at separating resection cavities from brain while not removing unresected tumors. In addition, it often produces a good approximation of the cavity-ventricle interface (Fig. 3b-c), which the other methods seldom do.

5   Discussion and Conclusion

A skull stripping method that can handle scans from subjects with an aggressive cancer (GBM) has been presented in this paper. Experimental evaluation on 48 subjects shows that our method outperforms three other widely-used brain extraction algorithms: BET, BSE and HWA. While we acknowledge that these techniques were not designed for this specific problem, to the best of our knowledge they represent the best alternative options. Since the proposed postprocessing requires a brain-like initialization, it is applicable only to ROBEX.

The proposed method is specific to T1-weighted MRI. When several channels are acquired (e.g. T1, T2, proton density, pre- and post-contrast), skull stripping is typically performed on the T1 volume and the resulting mask is propagated to the other channels. In fact, most brain extraction methods in the literature (BET being a notable exception) are designed to work with T1 data. However, it would certainly be possible to take advantage of multiple modalities to improve the stripping results. This direction remains to be explored.

An interesting by-product of the proposed algorithm is the segmentation of resection cavities (see Fig. 1e). Finding non-ventricular CSF is a common processing step after skull stripping in neuro image analysis pipelines, but existing methods do not handle the case where CSF might be in resection cavities adjacent to the ventricles. It would be interesting to evaluate to what extent the segmentation given by our method is reliable, since it may be useful to quantify the evolution of the resection. Exploring this direction and evaluating the algorithm on more cases also remain as future work.

Finally, it is important do discuss the algorithms' computational complexity. BET and BSE run in a few seconds, whereas ROBEX and HWA require one or two minutes on a modern machine. The proposed post-processing only requires a few seconds, which is almost negligible next to the execution time of ROBEX.

References

1. Fennema-Notestine, C., Ozyurt, I., Clark, C., Morris, S., Bischoff-Grethe, A., Bondi, M., Jernigan, T., Fischl, B., Segonne, F., Shattuck, D., Leahy, R., Rex, D., Toga, A., Zou, K., BIRN, Brown, G.: Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: effects of diagnosis, bias correction, and slice location. Human Brain Mapping 27(2), 99–113 (2006)
2. Shattuck, D., Sandor-Leahy, S., Schaper, K., Rottenberg, D., Leahy, R.: Magnetic resonance image tissue classification using a partial volume model. NeuroImage 13(5), 856–876 (2001)
3. Segonne, F., Dale, A., Busa, E., Glessner, M., Salat, D., Hahn, H., Fischl, B.: A hybrid approach to the skull stripping problem in MRI. Neuroimage 22(3), 1060–1075 (2004)
4. Smith, S.: Fast robust automated brain extraction. Human Brain Mapping 17(3), 143–155 (2002)
5. Iglesias, J., Liu, C., Thompson, P., Tu, Z.: Robust brain extraction across datasets and comparison with publicly available methods. IEEE Transactions on Medical Imaging (in press, 2011)
6. Grady, L.: Random walks for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(11), 1768–1783 (2006)
7. Breiman, L.: Random forests. Machine Learning 45(1), 5–32 (2001)
8. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
9. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images-a graph-theoretic approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(1), 119–134 (2006)
10. Sadananthan, S., Zheng, W., Chee, M., Zagorodnov, V.: Skull stripping using graph cuts. NeuroImage 49(1), 225–239 (2010)
11. Mikheev, A., Nevsky, G., Govindan, S., Grossman, R., Rusinek, H.: Fully automatic segmentation of the brain from T1-weighted MRI using Bridge Burner algorithm. Journal of Magnetic Resonance Imaging 27(6), 1235–1241 (2008)
12. Piper, J., Granum, E.: Computing distance transformations in convex and non-convex domains. Pattern Recognition 20(6), 599–615 (1987)