

Article

Improved Brain Tumor Segmentation via Registration-Based Brain Extraction

Maxwell Uhlich ^{1,*}, Russell Greiner ¹ , Bret Hoehn ¹, Melissa Woghiren ¹, Idanis Diaz ², Tatiana Ivanova ³ and Albert Murtha ⁴

¹ Department of Computing Science, University of Alberta, Edmonton, AB T6G 2E8, Canada; rgreiner@ualberta.ca (R.G.); bhoehn@ualberta.ca (B.H.); woghiren@ualberta.ca (M.W.)

² Engineering Faculty, University of Magdalena, Santa Marta 470003, Colombia; idiaz@unimagdalena.edu.co

³ Management Department, Alumni Association, Concordia University of Edmonton, Edmonton, AB T5B 4E4, Canada; taivanovapa@shaw.ca

⁴ Cross Cancer Institute of Alberta, Edmonton, AB T6G 1Z2, Canada; Albert.Murtha@albertahealthservices.ca

* Correspondence: uhlich@ualberta.ca

Received: 9 July 2018; Accepted: 5 September 2018; Published: 12 September 2018



Abstract: Automated brain tumor segmenters typically run a “skull-stripping” pre-process to extract the brain from the 3D image, before segmenting the area of interest within the extracted volume. We demonstrate that an effective existing segmenter can be improved by replacing its skull-stripper component with one that instead uses a registration-based approach. In particular, we compare our automated brain segmentation system with the original system as well as three other approaches that differ only by using a different skull-stripper—BET, HWA, and ROBEX: (1) Over scans of 120 patients with brain tumors, our system’s segmentation accuracy (Dice score with respect to expert segmentation) is 8.6% (resp. 2.7%) better than the original system on gross tumor volumes (resp. edema); (2) Over 103 scans of controls, the new system found 92.9% (resp. 57.8%) fewer false positives on T1C (resp. FLAIR) volumes. (The other three methods were significantly worse on both tasks). Finally, the new registration-based approach is over 15% faster than the original, requiring on average only 178 CPU seconds per volume.

Keywords: tumor segmentation; skull extraction; brain extraction; registration; BET; HWA; ROBEX

1. Introduction

Segmenting brain tumors within Magnetic Resonance Imaging (MRI) is an important challenge in computer vision; segmentation is required for treatment planning and other tasks and is currently performed by clinicians in a time-consuming and subjective manner [1]. The challenges for segmenters arise from the inhomogeneity of tumors, as well as natural variance within brains and noise within MRI [2].

Here, we describe a system for automatically segmenting cerebral abnormalities (specifically high grade astrocytoma) using the two standard MRI modalities used in a modern clinic: as is practiced in the clinic, edema (swelling within and around the tumor) is segmented primarily using FLAIR (Fluid Attenuated Inversion Recovery) images and the gross tumor volume (GTV) primarily using T1C (T1-weighted after injection of the contrast agent gadolinium) images; see Figure 1. Accurate segmentation of these two regions is useful to treatment and diagnosis [3].

Automatic segmenters are approaching the level of manual segmentation by clinicians for many cases [4,5], but are not used clinically due to sizable errors in infrequent cases. These segmentation errors often occur due to errors in *brain extraction* (also known as “skull-stripping”), a key preprocessing step of many segmenters that removes all non-brain tissue (skull, skin, eyes, eye sockets, sinuses, etc.)

from consideration before segmenting the abnormal region. Skull, skin, and other non-brain tissues can have image intensity profiles that are similar to brain tumors, which is why it is a common practice to remove all non-brain tissue from consideration, as this will simplify the challenge of tumor segmentation. It should be noted that any segmenter that uses skull-stripping as a preprocessing step risks having any extracerebral tumor extensions removed; we do not consider such tumors here.

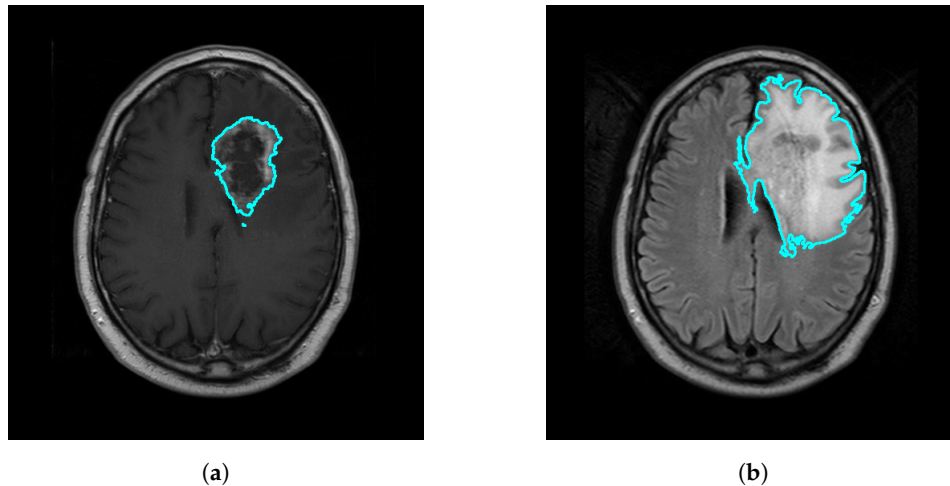


Figure 1. (a) GTV outlined on T1C image and (b) Edema outlined on FLAIR image.

We build upon an earlier, very effective Automated Brain Segmentation system, *ABS*(CBE, AS); here, we view the overall segmentation system in terms of just two components: a skull extraction module (CBE) and an automatic segmenter which acts upon the skull-stripped patient image (AS). CBE is an active contour-based brain extraction component, which is run as a pre-processing step, and AS (which identifies abnormal areas in the skull-stripped volume using thresholds derived from histograms of the intensities present) is run afterward to find the GTV and edema [4]. (Appendix A provides a short summary of this algorithm). Here, we replace that CBE with a registration-based brain extractor, **RBE**, and show that the resulting *ABS*(RBE, AS) system produces significantly better segmentations.

In addition to those results [4], we also evaluated the segmenter AS on the skull-stripped volumes that were used for the 2012 Brain Tumor Segmentation challenge (BRATS 2012) [6] held by the Medical Image Computing and Computer Assisted Intervention Society, where its results were comparable to the winners of the competition (BRATS ranked competitors by Dice score between segmentations generated and expert segmentations). While simple registration-to-atlas skull-stripping approaches have previously been suggested [7,8], an exhaustive literature search did not unearth any study that quantitatively compares this to other skull-stripping methods in the context of segmentation of unhealthy brains, which is especially challenging as standard rigid registration is problematic due to abnormalities within such unhealthy brains.

Furthermore, while many other skull-stripping pipelines have been suggested [9–13], none provide a comprehensive quantitative evaluation involving the considerable challenge of brain tumors. Some only consider healthy brains [9,12], or rely on artificial lesions or artificial datasets [9,10]; others evaluate on very small datasets [10,12] or fail to address the mass effect of brain tumors [7,9–13], which is the principle challenge to a brain extraction pipeline used in our domain. The most similar study to this one is by Chaddad and Tanougast [14], who show skull-stripping results for two popular skull-strippers in comparison to the authors' method (Chaddad–Tanougast), and show tumor segmentation results using only the Chaddad–Tanougast method; the skull-stripping results show the Chaddad–Tanougast method achieving equivalent results to the BET method that we compare to here, and superior results to the brain surface extractor (BSE) [15] method that we have omitted. Our work

rectifies the shortcomings of the previous studies by testing on two large datasets of real patients, a set of brain tumor patients as well as a control set with tumor-free brains.

We consider $ABS(e, AS)$ for five different brain extraction tools e : the RBE and CBE mentioned above, as well as three well-known skull-strippers, Brain Extraction Tool (BET) [16], Hybrid Watershed Algorithm (HWA) [17], and Robust Brain Extraction (ROBEX) [18].

CBE [4] strips the skull using two active contours applied after thresholding and morphological operations. The first contour is initialized to the borders of the image and then propagates inward, settling on the boundary between the skull and the background; in effect shrink-wrapping the head. The second contour starts where the first one stopped, and once again propagates inward, this time stopping in the interior of the skull/skin. Under certain circumstances, these contours can be led astray, leading to erroneous skull-stripping and poor quality downstream tumor segmentations (see Figure 2).

BET strips the skull by iteratively evolving a surface mesh that grows to model the brain, using an estimated brain/non-brain threshold, starting with a sphere localized around the center of gravity of the head. BET works on multiple modalities, so it is applied to both T1C and FLAIR separately in this study.

ROBEX strips the skull using a shape model trained on healthy brains to fit a triangular mesh to the probabilistic output of a Random Forest brain boundary classifier. It then uses graph cuts to refine the resulting contour.

Finally, HWA uses a watershed algorithm to create an initial model of the white matter and then uses this model in deforming a smooth surface to fit the brain. As HWA is optimized for T1-weighted MR images, it is only applied to the T1C volumes; co-registration is used to obtain the FLAIR results.

Although previous work has shown the use of computed tomography (CT) scans in combination with MRI to be beneficial in both skull-stripping and in tumor segmentation [19], we consider only MRI here as CTs were not available for all patients, and both the existing segmenter AS and BRATS segmenters only make use of standard MRI images.

Section 2 introduces our skull-stripping method, RBE, and Section 3 evaluates the resulting systems on two large datasets: the first dataset contains T1C and FLAIR images of 120 high-grade tumor patients, while the second dataset contains T1C and FLAIR images of 103 *brain-tumor-free* individuals.

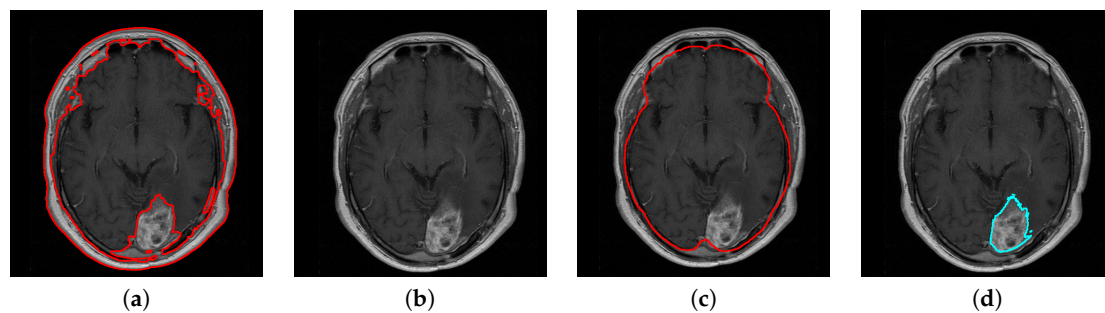


Figure 2. (a) an erroneous skull-stripping by CBE; (b) resulting tumor segmentation, by $ABS(CBE, AS)$ —i.e., no tumor found; (c) RBE skull-stripping of same patient; (d) resulting segmentation of $ABS(RBE, AS)$.

2. Materials and Methods

This section summarizes our RBE approach. RBE uses a brain atlas created by the Montreal Neurological Institute, called MNI-152, which was created by nonlinearly aligning 152 T1 brain volumes and their accompanying T2 volumes in a way that prevents bias toward any particular brain [20,21]. It is essentially a spatial average of 152 brains into which we can transform the MRI volumes of a new patient. Crucially, the MNI provides a corresponding *brain mask*, $Atlas_{BrainMask}$, that identifies brain tissue within the volume.

Figure 3 illustrates how RBE identifies brain tissue within the T1C and FLAIR images. Given a volume of T1C images and a corresponding volume of FLAIR images of a patient's brain, RBE begins by registering the T1C volume to the FLAIR. These two modalities are acquired on the same date and are of the same brain, so rigid registration is sufficient, producing the transformation denoted $\text{Map}_{\text{T1C} \rightarrow \text{FLAIR}}$. After this co-registration step, the volumes are in *FLAIR-space*.

The FLAIR volume is then registered to the T2-weighted MNI-152 atlas to obtain a transform from the FLAIR-space to atlas-space, for which a brain mask is provided. As the patient brain may not have the same size and shape as the atlas, rigid and affine registration are used to achieve the best possible fit while preserving easy invertibility of the resulting transform, $\text{Map}_{\text{FLAIR} \rightarrow \text{Atlas}}$.

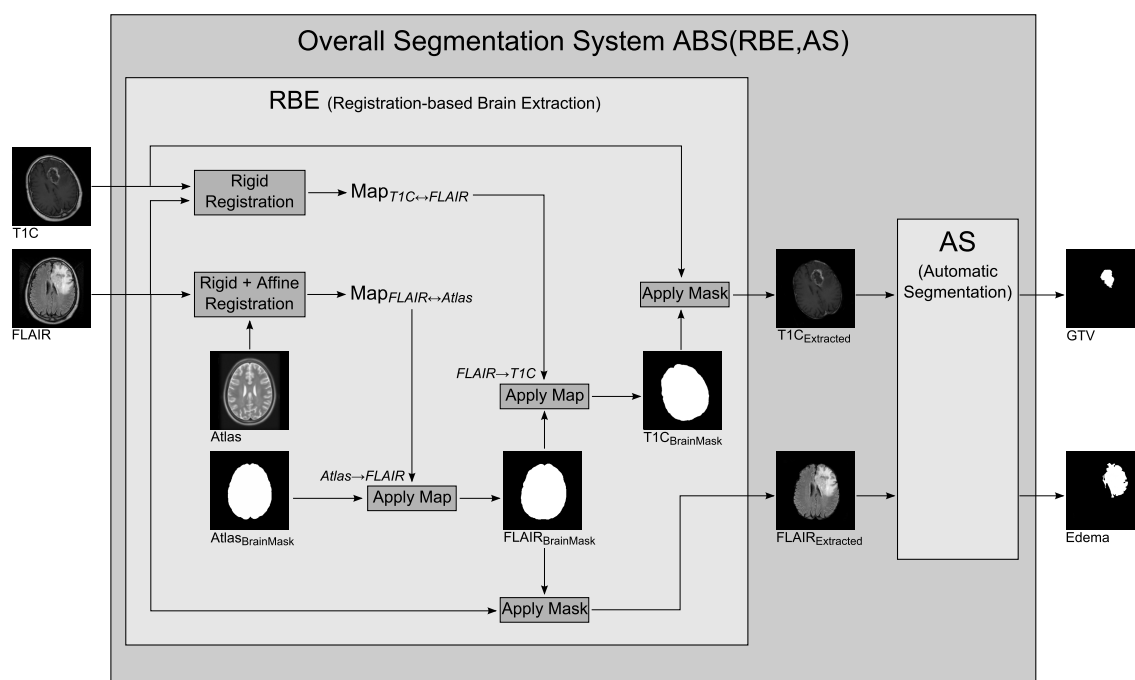


Figure 3. *ABS(RBE, AS)* system for automatic segmentation. Input are T1C and FLAIR volumes: RBE co-registers the T1C to the FLAIR, and registers the FLAIR to an atlas for which a brain mask (indicating what is brain tissue and what is not) is known. The inverse transformations are applied to the atlas brain mask to obtain brain tissue masks for the input T1C and FLAIR. The resulting skull-stripped T1C and FLAIR volumes are then segmented by the Automatic Segmenter, AS.

We considered a more general class of deformable transformations but found it did not work as well: as tumor and skull had similar pixel intensity, general models tended to be confused by local features such as tumors, which resulted in larger misclassification error.

The segmenter AS has been optimized for unmodified MR images, so the atlas brain mask is transformed into the FLAIR-space to be applied to the original images; this is achieved by inverting $\text{Map}_{\text{FLAIR} \rightarrow \text{Atlas}}$, and applying this inverse transformation ($\text{Map}_{\text{Atlas} \rightarrow \text{FLAIR}}$) to the atlas brain mask. The resulting FLAIR brain mask is then transformed by the inverse of $\text{Map}_{\text{T1C} \rightarrow \text{FLAIR}}$ to produce a brain mask for the original unregistered T1C volume. The two brain masks, $\text{FLAIR}_{\text{BrainMask}}$ and $\text{T1C}_{\text{BrainMask}}$, are applied to the original FLAIR and T1C volumes to provide skull-stripped brains for AS to segment.

Our empirical studies found that registration of patient FLAIR volumes to the T2 atlas resulted in fewer registration errors than T1C to T1, which is why RBE co-registers the patient's T1C to the FLAIR, and then registers the FLAIR to the atlas. All registration steps are performed using the freely available BRAINSFit and BRAINSResample command line modules provided with the 3D Slicer application (<http://www.slicer.org>) [22].

All methods are evaluated over two datasets. The **Tumor Dataset** contains the T1C and FLAIR brain volumes of 120 unique patients with high grade astrocytoma, as well as corresponding ground truths for edema and GTV (provided by a human expert) for each patient. The **Control Dataset** contains the T1C and FLAIR brain volumes of 103 control subjects with no apparent brain tumors (these subjects had cancer in other parts of the body and had brain MRIs to rule out metastasis). All scans were done with a 1.5 T magnet, at a spatial resolution of $0.45 \text{ mm} \times 0.45 \text{ mm} \times 5 \text{ mm}$. FLAIR scans had an echo time (TE) of 110 ms and a repetition time (TR) of 9525 ms; the T1C scans had TE ranging from 11 ms to 12 ms and TR ranging from 460 ms to 528 ms. All data was used with Institutional Review Board approval.

3. Results

Each of the skull-stripping methods considered here $e \in \{\text{RBE}, \text{CBE}, \text{BET}, \text{HWA}, \text{ROBEX}\}$ is used as the first step of the $ABS(e, AS)$ segmentation system, all using the same segmenter, AS. Let $V_e(p)$ be the segmented volume (say GTV) returned when running the $ABS(e, AS)$ on patient p , and $V^*(p)$ be the true segmentation (manually created by an expert). To evaluate the brain extractor e with respect to a patient p , we use the Dice value:

$$D(e, p) = \frac{2 |V_e(p) \cap V^*(p)|}{|V_e(p)| + |V^*(p)|}.$$

We will then use the average value over a population of patients to evaluate the extractor e . A Dice value of 1 indicates that two regions are identical, while a value of 0 indicates that there is no overlap. To put these scores in perspective, we found the mean Dice agreement achieved between our two experts (MDs) over 19 cases was 0.85 for GTV and 0.79 for edema. This agrees with average inter-operator variability observed in previous studies [1].

3.1. Qualitative Analysis

Infrequently, CBE fails when the active contour meant to shrink to the inner border of the skull does not stop and proceeds into the interior of the brain, as seen in Figure 2a,b; in this case, the region between the contours will be removed and a large amount of tumor tissue is unavailable for the segmenter to find. RBE performs well, much more consistently than CBE. Of the 223 FLAIR volumes considered, only two were poor quality registrations: one in the Tumor Dataset and one in the Control Dataset. These two cases failed to strip parts of the eye sockets, which were then erroneously segmented as tumors.

As previously shown [23], HWA often removes brain tissue along with the skull, but rarely fails to strip non-brain tissue. This helps the performance of $ABS(\text{HWA}, AS)$ on the Control dataset but hurts its performance on the Tumor dataset. It occasionally avoids large tumors altogether, only extracting the brain tissue of one hemisphere of the brain due to the presence of a large abnormality in the other. Conversely, BET is much less sensitive, often including the eyes and other non-brain tissues in the stripped image that is provided to the segmenter. This is greatly penalized on the Control dataset, since these areas can then be mistakenly segmented as tumors. ROBEX is also more likely to include non-brain tissue than to strip brain tissue, which means it will often mistakenly include eye sockets and other non-brain areas. Our empirical analysis confirmed that BET, HWA, and ROBEX are problematic within the context of a tumor segmenter.

Figure 4 shows some results of these algorithms, in order to illustrate examples of where each failed.

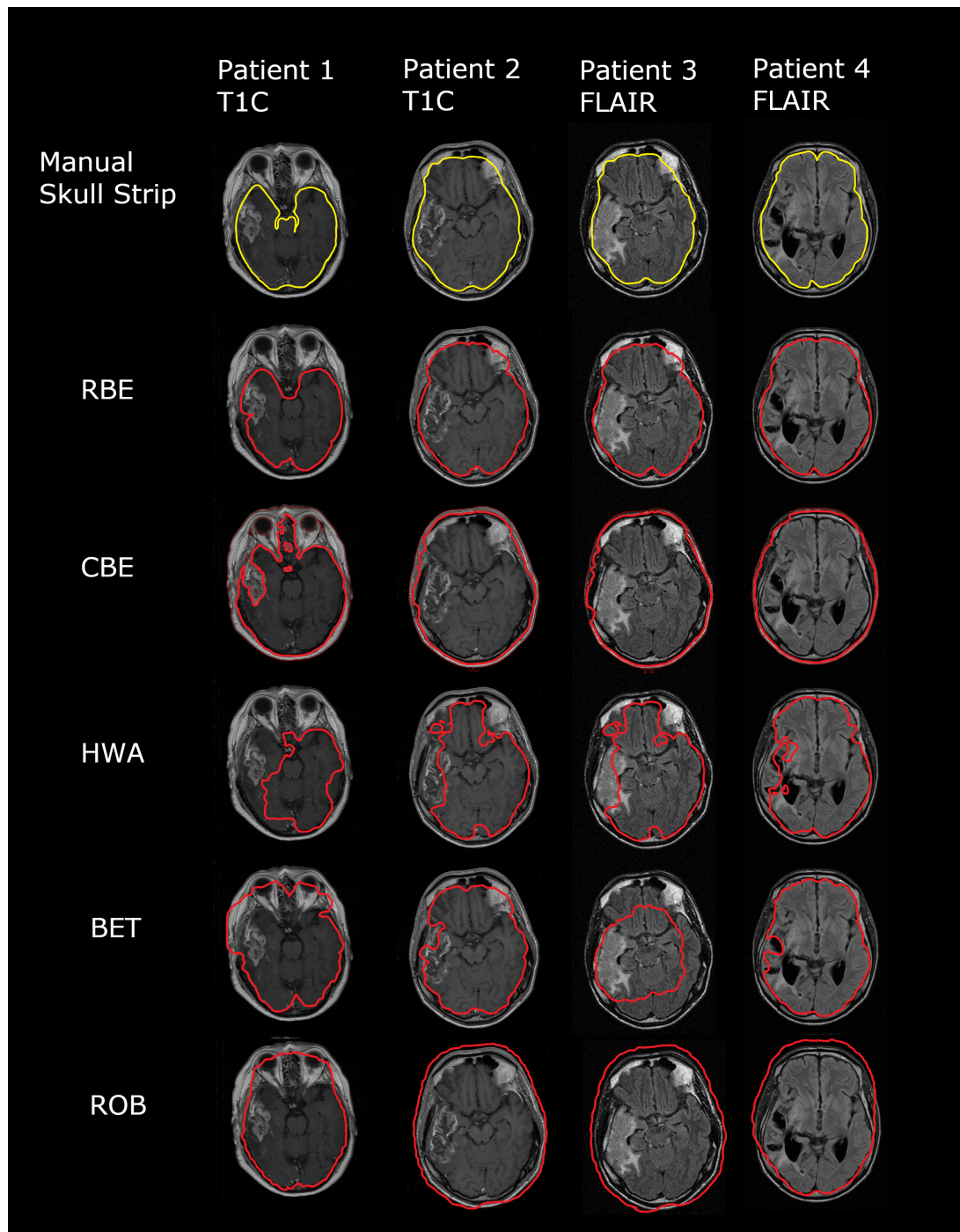


Figure 4. Comparison of our RBE approach to other skull-stripping methods. RBE often improved over CBE by correctly identifying the outer border of the brain rather than the inner border of the skull. This resulted in a more refined skull strip; see, especially, Patient 1, as well as outer contours of other three. We see that HWA and BET sometimes failed to define accurate borders between brain and non-brain tissues, especially in the presence on cerebral abnormalities; see all four patients. Although performance of ROBEX was similar to RBE's throughout our experiments, RBE was still superior overall, as it more accurately and reliably excluded ocular tissue and the meninges; see especially Patient 1, as well as outer contours of other three.

3.2. Tumor Dataset

On the Tumor Dataset, segmentation errors can occur either due to the skull-stripper or the segmenter; a false positive could occur because a portion of skull that the skull-stripper failed to remove, is mistaken for a tumor, or because the segmenter mis-labels some healthy brain tissue as tumor. A false negative also indicates either a failure of the skull-stripper (when it excludes an area of brain that would otherwise have been segmented as tumor) or the segmenter (when it fails to detect some part of a tumor in the brain tissue).

Figure 5 displays the Dice value results for each of the 120 volumes in the Tumor Dataset. The cases in each chart are sorted such that the Dice value of the RBE method is increasing; RBE is compared separately against CBE, BET, HWA and ROBEX to avoid clutter. In Figure 5a (Top), every point where the dark line is above the light line is an instance where $ABS(RBE, AS)$ outperforms $ABS(CBE, AS)$; conversely, each point where the light line is above the dark line is a case where $ABS(CBE, AS)$ outperforms $ABS(RBE, AS)$. To simplify notation in the following discussion, we use e as a shorthand for $ABS(e, AS)$. For the T1C volumes, there are 91 cases where both RBE and CBE produce roughly equivalent segmentations (Dice value within 0.02), four cases where CBE slightly outperforms RBE (mean difference in Dice of 0.04 with a max of 0.06), and 25 cases where RBE performs significantly better than CBE (mean difference in Dice of 0.20 with a max of 0.57). This is easy to see, by noting that there are only a few “bumps” where the light green line is above the dark line, but many places where that green line is far below the dark one. On the T1C volumes overall, RBE achieves a mean Dice value of 0.4983 and CBE achieves 0.4587 (see Figure 6). This difference is significant in a paired t -test at $p < 0.01$. We observed similar results for the FLAIR volumes, with RBE achieving a mean Dice value of 0.6985 and CBE achieving 0.6803, which is significant at $p < 0.05$.

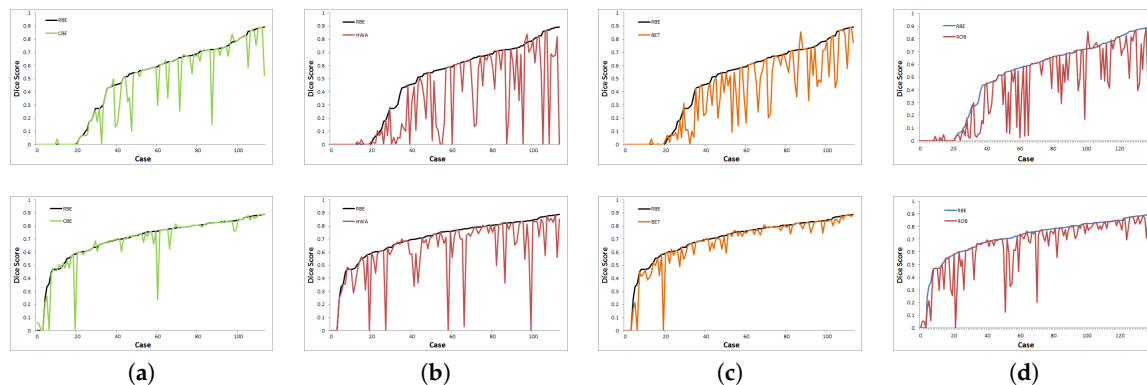


Figure 5. Segmentation results for a tumor dataset. The top row is with respect to T1C volume and second row is wrt FLAIR volumes. The cases in each graph are sorted in an ascending RBE Dice value: (a) RBE vs. CBE; (b) RBE vs. HWA; (c) RBE vs. BET; (d) RBE vs. ROBEX.

With default parameter settings, six patients could not be processed by one or both of BET and HWA (either on the T1C or the FLAIR volume), but a comparison was possible over the remaining 114 patients. For the “process-able” T1C volumes, HWA achieves a mean dice value of 0.3367 and BET achieves 0.3908. Over the corresponding FLAIR volumes, HWA achieves a mean dice value of 0.6112 and BET achieves 0.6669. ROBEX achieves a mean Dice value of 0.4046 over all 120 T1C volumes and a mean Dice value of 0.6396 over the corresponding FLAIR volumes; see Figure 6. In paired t -tests with RBE, each of HWA, BET and ROBEX are significantly worse at $p < 0.01$, indicating that BET, HWA, and ROBEX are not suitable skull-strippers within the context of a tumor segmenter.

These results were based on the default setting of the various algorithms. We then explored whether they would produce better results, with other settings. We could not change ROBEX, as it was a black box that did not permit any variation in parameters. For BET, we re-ran our study while varying the f parameter for both the Robust and Standard versions of the algorithm, following [3]. For HWA,

we adjusted the two parameters: whether or not to use the AtlasCorrection feature, and whether to use Less Restriction or More Restriction. In every case, we found that these modifications produce identical results—that is, they converged to the exact same segmentation (although some modifications led to slightly shorter runtimes, and others, slightly longer). We think this may be because the available BET and HWA parameters are basically designed to reduce the inclusion of extraneous tissue such as neck and shoulder regions (which are almost never present in our datasets), but not to help avoid the skull.

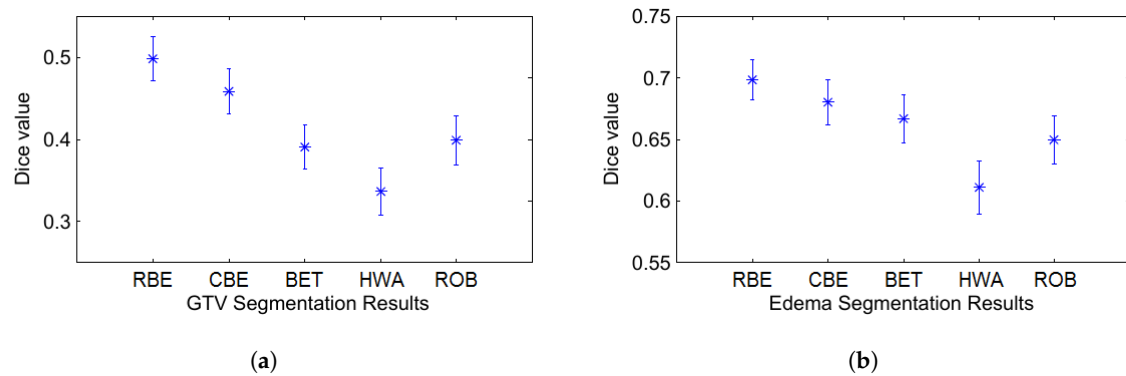


Figure 6. (a) Mean Dice values and associated standard error bars on GTV results; (b) Mean Dice values and associated standard error bars on edema results.

3.3. Control Dataset

While our primary goal is to increase performance of the segmenter on the Tumor Dataset, it is useful to compare the segmenters on a dataset of control patients; on these patients, an ideal tumor segmenter should detect nothing—i.e., 0 voxels of GTV and 0 of edema. In practice, a segmenter with perfect skull-stripping may occasionally find abnormalities due to natural variance in the brain or noise in the MR images. However, with poor skull removal, a segmenter may find abnormalities more frequently.

Over 103 control subject T1C volumes, CBE identified “tumors” in 28 cases, while RBE incorrectly identified tumors in just two cases—one was a failure of the skull-stripper (parts of the eyes were not stripped and were identified as abnormal brain tissue) and the other a failure of the segmenter (some dura mater was segmented). Errors were more common in FLAIR volumes, with CBE falsely identifying 45 cases of edema and RBE falsely identifying 19. For both T1C and FLAIR volumes, the results were significantly different at $p < 0.01$.

The alternative skull-strippers, BET and HWA, ran to completion on 102 of the 103 cases in the Control Dataset. Over those 102 T1C volumes, the BET-based segmenter produced 85 false positives and the HWA-based segmenter produced 39. Over the corresponding 102 FLAIR volumes, BET resulted in 87 false positives and HWA in 46. ROBEX ran to completion in all 103 cases, and produced 32 false positives over the T1C volumes and 41 false positives over the corresponding FLAIR volumes. See Figure 7.

Once again, varying the parameters did not change the BET and HWA results—they again converged to the same final answer.

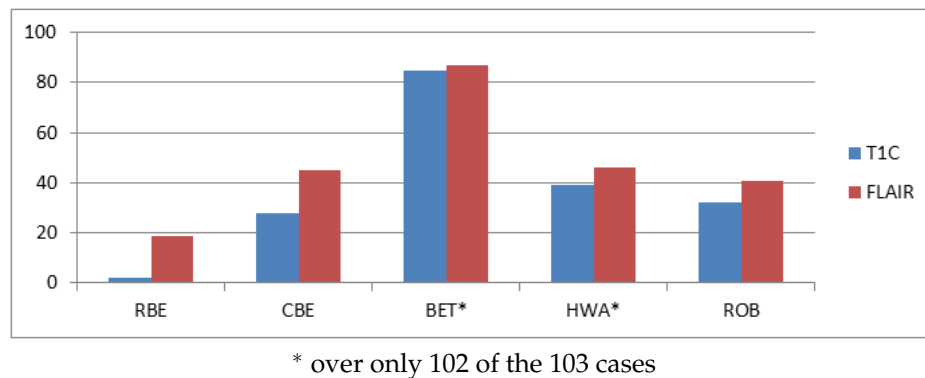


Figure 7. Number of false positives, over the dataset of 103 s.

3.4. Runtime

It is critical that our segmenter be accurate—i.e., find the correct skull regions. It is also important that it finds these regions quickly; a primary reason for developing strong automatic segmenters is to save clinicians and researchers the time they spend doing manual segmentation; previous studies have shown that experts require an average of 7 to 12 min to segment a single patient [24]. While computer processing time is certainly less valuable than a clinician’s time, it is still important for these methods to be reasonably fast if they are to be used in practice.

We considered the runtime over only the 216 cases that all skull-strippers could process to completion. The mean runtime of the skull-stripping component RBE over these 216 cases (both Tumor and Control datasets) was 98.84 s for each case, compared to 131.16 for CBE (All experiments were performed on an Intel Xeon dual core 2 GHz processor with 8 GB RAM, running Ubuntu). ROBEX was the slowest, taking 154 s on average to complete each skull strip. While BET and HWA are very fast, at 13.74 s and 47.10 s, respectively, the results of the previous sections indicate that they should not be used for tumor segmentation, so we only compare RBE and CBE in the remaining discussion.

Table 1 shows the time consumption of the major steps of each algorithm; operations common to each, such as co-registration of FLAIR and T1C, are not included. The time taken by the segmentation component AS (acting upon the skull-stripped volumes) is relatively low, requiring only 15.38 s on average. In total (including steps common to both algorithms as well as file conversions and other tasks), *ABS*(RBE, AS) requires 178 s to skull-strip and segment a tumor patient, while *ABS*(CBE, AS) requires 215 s, meaning that our registration-based RBE is faster than the active contour approach CBE, as well as producing more accurate segmentation results.

Table 1. Runtime breakdown of RBE and CBE in seconds.

RBE (Total 98.84)	
Step 1: Register FLAIR to Atlas	93.07
Step 2: Back transform to FLAIR	3.20
Step 3: Back transform to T1C	2.57
CBE (Total: 131.16)	
Step 1: Get FLAIR features	0.48
Step 2: Get T1C features	0.60
Step 3: Compute initial threshold	28.29
Step 4: Active Contour propagation	98.92
Step 5: Back transform to T1C	2.88

4. Conclusions

Effective skull-stripping is a key component of many automatic tumor segmenters. The experiments presented here demonstrate that a simple atlas-based approach to skull-stripping can achieve segmentation

results that are significantly more accurate than an active-contour-based approach. Although the use of rigid registration suggests that the atlas and brain mask may not be perfectly matched to the patient, our two studies show the registration-based approach to be very robust within the context of segmenting tumors in glioma patients, and greatly reduces the spurious segmentation of nonexistent tumors in controls. Moreover, the registration-based approach runs faster than the active contour approach. We anticipate that this system will be a useful tool in the treatment and study of brain tumors.

Author Contributions: Conceptualization, M.U. and R.G.; Funding acquisition, R.G.; Investigation, B.H. and M.W.; Methodology, M.U., B.H. and I.D.; Project administration, R.G. and A.M.; Resources, R.G. and I.D.; Supervision, R.G. and A.M.; Validation, M.U., M.W., T.I. and A.M.; Visualization, M.U. and T.I.; Writing—Original draft, M.U., R.G., B.H. and M.W.

Acknowledgments: This work was supported, in part, by grants from NSERC (Canada's Natural Science and Engineering Research Council), and Amii (Alberta Machine Intelligence Institute). We are also grateful for assistance from Prof. Pierre Boulanger and members of the Brain Tumor Analysis Project.

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

Appendix A. Summary of ABS(CBE, AS)

The ABS(CBE, AS) process relies on intensity thresholds that are automatically identified from a histogram of the intensity values present in the image (which allows the algorithm to run without modification on images from different scanners or on MRIs generated with different parameters, which may have very different intensity profiles); the thresholds identified help mask out background voxels, as well as identify likely candidates for skull, GTV and edema.

The brain extraction component, CBE, localizes the skull by evolving a pair of active contours from outside the brain. The first contour begins on the outside of the head and wraps around the patient's skull (delineating the border between the bright skull and the dark background outside the skull). The second contour is initialized where the first contour stops, propagates through the skull wall and terminates on the interior of the skull (delineating the boundary between the bright skull/non-brain tissue and the lower-intensity brain tissue). These two contours form the boundaries of the mask that are used to strip the skull.

The segmentor component, AS, applies the automatically identified thresholds to identify areas of interest within the skull-stripped brains, in particular identifying small seed regions that are very likely GTV/edema, and surrounding candidate regions (geodesic mask). Geodesic dilation is used to grow the seed regions within the geodesic mask, in order to identify contiguous volumes representing GTV in the T1C volume and Edema in the FLAIR volume.

References

1. Mazzara, G.; Velthuizen, R.; Pearlman, J.; Greenberg, H.; Wagner, H. Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. *Int. J. Radiat. Oncol. Biol. Phys.* **2004**, *59*, 300–312. [[CrossRef](#)] [[PubMed](#)]
2. Pham, D.; Xu, C.; Prince, J. Current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 315–337. [[CrossRef](#)] [[PubMed](#)]
3. Popescu, V.; Battaglini, M.; Hoogstrate, W.; Verfaillie, S.; Sluimer, I.; van Schijndel, R.; van Dijk, B.; Cover, K.; Knol, D.; Jenkinson, M.; et al. Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. *NeuroImage* **2012**, *61*, 1484–1494. [[CrossRef](#)] [[PubMed](#)]
4. Ben Salah, M.; Diaz, I.; Greiner, R.; Boulanger, P.; Hoehn, B.; Murtha, A. Fully Automated Brain Tumor Segmentation using two MRI Modalities. In *International Symposium on Visual Computing*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8033, pp. 30–39.
5. Diaz, I.; Boulanger, P.; Greiner, R.; Hoehn, B.; Rowe, L.; Murtha, A. An Automatic Brain Tumor Segmenter Tool. In *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 3–7 July 2013.

6. Menze, B.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans. Med. Imaging* **2014**, *34*, 1993–2024. [[CrossRef](#)] [[PubMed](#)]
7. Doshi, J.; Erus, G.; Ou, Y.; Gaonkar, B.; Davatzikos, C. Multi-Atlas Skull Stripping. *Acad. Radiol.* **2013**, *20*, 1566–1576. [[CrossRef](#)] [[PubMed](#)]
8. Bauer, S.; Fejes, T.; Reyes, M. A Skull-stripping Filter for ITK. *Insight J.* **2012**, 70–78. [[CrossRef](#)]
9. Hahn, H.; Peitgen, H. The Skull Stripping Problem in MRI Solved by a Single 3D Watershed Transform. In *Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1935, pp. 134–143.
10. Acosta-Cabronero, J.; Williams, G.; Pereira, J.; Pengas, G.; Nestor, P. The Impact of Skull-Stripping and Radio-Frequency Bias Correction on Grey-Matter Segmentation for Voxel-Based Morphometry. *NeuroImage* **2008**, *39*, 1654–1665. [[CrossRef](#)] [[PubMed](#)]
11. Prasad, G.; Joshi, A.; Feng, A.; Toga, A.; Thompson, P.; Terzopoulos, D. Skull-stripping with machine learning deformable organisms. *J. Neurosci. Methods* **2014**, *236*, 114–124. [[CrossRef](#)] [[PubMed](#)]
12. Carass, A.; Wheeler, M.; Cuzzocreo, J.; Bazin, P.; Bassett, S.; Prince, J. A joint registration and segmentation approach to skull stripping. In Proceedings of the 4th IEEE International Symposium on Biomedical Imaging, Arlington, VA, USA, 12–15 April 2007; pp. 656–659.
13. Leung, K.; Barnes, J.; Modat, M.; Ridgway, G.; Bartlett, J.; Fox, N.C.; Ourselin, S. Automated brain extraction using multi-atlas propagation and segmentation (MAPS). In Proceedings of the IEEE International Symposium on Biomedical Imaging, Chicago, IL, USA, 30 March–2 April 2011; pp. 2053–2056.
14. Chaddad, A.; Tanougast, C. Quantitative evaluation of robust skull stripping and tumor detection applied to axial MR images. *Brain Inform.* **2016**, *3*, 53–61. [[CrossRef](#)] [[PubMed](#)]
15. Shattuck, D.W.; Sandor-Leahy, S.R.; Schaper, K.A.; Rottenberg, D.A.; Leahy, R.M. Magnetic Resonance Image Tissue Classification Using a Partial Volume Model. *NeuroImage* **2001**, *13*, 856–876. [[CrossRef](#)] [[PubMed](#)]
16. Smith, S. Fast robust automated brain extraction. *Hum. Brain Mapp.* **2002**, *17*, 143–155. [[CrossRef](#)] [[PubMed](#)]
17. 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* **2004**, *22*, 1060–1075. [[CrossRef](#)] [[PubMed](#)]
18. Speier, W.; Iglesias, J.; El-Kara, L.; Tu, Z.; Arnold, C. Robust Skull Stripping of Clinical Glioblastoma Multiforme Data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 659–666.
19. Weltens, C.; Menten, J.; Feron, M.; Bellon, E.; Demaerel, P.; Maes, F.; den Bogaert, W.V.; van der Schueren, E. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiother. Oncol.* **2001**, *60*, 49–59. [[CrossRef](#)]
20. Fonov, V.; Evans, A.; Botteron, K.; Almli, C.; McKinstry, R.; Collins, D. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* **2011**, *54*, 313–327. [[CrossRef](#)] [[PubMed](#)]
21. Fonov, V.; Evans, A.; McKinstry, R.; Almli, C.; Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **2009**, *47*, S102. [[CrossRef](#)]
22. Fedorov, A.; Beichel, R.; Kalpathy-Cramer, J.; Finet, J.; Fillion-Robin, J.C.; Pujol, S.; Bauer, C.; Jennings, D.; Fennessy, F.; Sonka, M.; et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **2012**, *30*, 1323–1341. [[CrossRef](#)] [[PubMed](#)]
23. Fennema-Notestine, C.; Ozyurt, I.; Clark, C.; Morris, S.; Bischoff-Grethe, A.; Bondi, M.; Jernigan, T.; Fischl, B.; Segonne, F.; Shattuck, D.; et al. Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location. *Hum. Brain Mapp.* **2006**, *27*, 99–113. [[CrossRef](#)] [[PubMed](#)]
24. Birkbeck, N.; Cobzas, D.; Jagersand, M.; Murtha, A.; Kesztyues, T. An interactive graph-cut method for brain tumour segmentation. In Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 7–8 December 2009.

