Automatic Segmentation of Non-enhancing Brain Tumors in Magnetic Resonance Images

Lynn M. Fletcher-Heath, Lawrence O. Hall, Dmitry B. Goldgof and F. Reed Murtagh

Corresponding author: L.O. Hall, Dept. of CSE, ENB118, Univ. of South Florida, Tampa, Fl. 33620, .Ph. 813-974-4195; Fax: 813-974-5456

L. Fletcher-Heath, L. Hall and D. Goldgof are with the Computer Science Department and F. R. Murtagh is with the Department of Radiology of the University of South Florida, Tampa, Florida 33620. E-mail: lfletche,hall,goldgof@csee.usf.edu, murtagh@rad.usf.edu.

Abstract

Tumor segmentation from magnetic resonance (MR) images may aid in tumor treatment by tracking the progress of tumor growth and/or shrinkage. In this paper we present the first automatic segmentation method which separates non-enhancing brain tumors from healthy tissues in MR images to aid in the task of tracking tumor size over time. The MR feature images used for the segmentation consist of three weighted images (T1, T2 and proton density) for each axial slice through the head. An initial segmentation is computed using an unsupervised fuzzy clustering algorithm. Then, integrated domain knowledge and image processing techniques contribute to the final tumor segmentation. They are applied under the control of a knowledge-based system. The system knowledge was acquired by training on two patient volumes (14 images). Testing has shown successful tumor segmentations on four patient volumes (31 images). Our results show that we detected all six non-enhancing brain tumors, located tumor tissue in 35 of the 36 ground truth (radiologist labeled) slices containing tumor and successfully separated tumor regions from physically connected CSF regions in nine of nine slices. Quantitative measurements are promising as correspondence ratios between ground truth and segmented tumor regions ranged between 0.368 and 0.871 per volume, with percent match ranging between 0.530 and 0.909 per volume.

Keywords

MRI, non-enhancing brain tumors, image processing, automatic tissue classification, fuzzy, clustering

I. Introduction

Magnetic resonance imaging (MRI) is often the medical imaging method of choice when soft tissue delineation is necessary. This is especially true for any attempt to segment brain tissues, normal or abnormal. Image segmentation is a tool that has been applied to medical imaging modalities to differentiate tissue types for purposes of volume measurement and visualization [3], [9]. One such modality that has had a great deal of attention from those researching image segmentation techniques, is MRI [34], [35]. Although MR segmentation methods have been quite successful on normal tissues [23], [31], [24], [33], the actual methods of MR segmentation are still very much in the development stages for pathological tissues with some success recorded for specific disease processes [3], [6], [18]. Recent work has shown that segmentation of gadolinium enhanced glioblastoma-multiforme tumor tissues in magnetic resonance (MR) images is possible via clustering combined with knowledge-based (KB) techniques [6], [7], [5]. Here and in [6], [7], [5]

unsupervised clustering used three weighted images, T1, T2 and proton density (PD). Fuzzy c-means (FCM) clustering is a widely used clustering method, since it does not require training data and is therefore operator independent [3], [9], [24], [16], [36]. Also, it provides information on how well a pixel "fits" a cluster.

Other methods, such as k-nearest neighbors (k-NN) also group pixels based on their similarities in each feature image, however, manually selected training data from the various tissue types is required per slice per volume [9], [28]. Therefore, the accuracy of supervised segmentation techniques depends upon the accuracy and repeatability of the necessary operator intervention.

After clustering, a KB system was used to isolate and determine enhancing tumor volumes [6], [7], [5], better enabling a neurologist to quantitate tumor growth or shrinkage and therefore determine the stage and type of brain tumor. However, approximately 10% of all brain tumors do not brighten in MR images once a contrast agent has been administered, making separation of the various brain tissues via clustering techniques extremely difficult. Therefore, we define a non-enhancing brain tumor as one without any contrast induced enhancing regions. In addition, if a segmentation method was available that did not require enhancement agents for the planning of tumor treatment, or for analysis of current treatment protocols, this less invasive, cost effective imaging method would be preferred over others.

Responding to potential inconsistencies of supervised classification techniques, while using the universally available scan techniques, this paper describes a method which after training requires no further operator intervention. Data is grouped into fuzzy clusters with FCM and approximate shape characteristics and other measurements are applied to label clusters. The measurements utilized reflect expert domain knowledge resulting in a knowledge guided image processing system which isolates the non-enhancing tumor tissues, thereby allowing tumor volume measurement. Responding to potential inconsistencies of supervised classification techniques, while using the universally available scan techniques, this paper describes a method which after training requires no further operator intervention. It uses FCM clustering and knowledge guided image processing which isolates the non-enhancing tumor tissues, thereby allowing tumor volume measurement.

In the proceeding, there are sections on related work including fuzzy approaches to image analysis, images and our knowledge guided approach, experimental results and a discussion of them and a conclusions section.

II. Related work

Other researchers recent work has shown promise in the segmentation of non-enhancing pathology, however, not without the help of supervised classification techniques or scan techniques not available on all MR scanners. Vinitski et al. [37] demonstrate the use of k-NN while segmenting multiple sclerosis (MS) lesions and brain tumors from a limited number of patients. The method of Peck et al. [27] uses eigenimage analysis which nicely shows tumor segmentation on images without gadolinium, though operator intervention is required to select a region of interest (ROI) inside each tissue of a registered set of images per slice. Their algorithm then determines if any of the surrounding pixels should be included in the ROI. Although shown to be successful, these methods are susceptible to intra and inter-observer variabilities.

Image acquisition techniques currently available with non-standard imaging equipment such as echoplanar diffusion-weighted imaging (DWI) and perfusion-weighted imaging (PWI), which are both generally used for imaging potential stroke victims, have been used to analyze brain tumor physiology in [1], [32], [4], [20]. Cerebral Blood Volume (CBV) maps were calculated in [1] where progress was made in discerning between low and high-grade tumors. CBVs, however, were not sufficient in defining boundaries in four of eight non-enhancing tumors (six low-grade and two high-grade tumors) and in eight of eleven enhancing tumors (all high-grade) making tumor volume measurements difficult. Also, the measurement of CBVs requires perfusion imaging and therefore a contrast agent to perfuse across tissues in addition to many images per slice. Diffusion-weighted images are acquired on high-grade, enhancing tumors in [32] demonstrating the ability to differentiate between enhancing and non-enhancing components of high-grade tumors, however, evidence that non-enhancing tumor can be differentiated from white, grey and CSF tissues is essential to our task of non-enhancing segmentation. Also, a large contribution in [32] is the ability to separate necrosis from tumor tissue, which addresses only part of the segmentation problem. Good results in differentiating white matter, necrosis, cysts

and solid enhancing tumors are shown in [4] where they considered 40 patients, however, results regarding non-enhancing tumor delineation from grey matter were poor and CSF was not considered. Similar results in [20] concentrate on enhancing tumors, as only one of 28 patients had a non-enhancing brain tumor. Specialized equipment (echoplanar imaging) is required for the data acquisition in [1], [32], [4], [20] which is not available at all institutions or imaging centers.

A fast implementation of fluid-attenuated inversion-recovery (FLAIR) showed promise in [12] where the differentiation of enhancing and non-enhancing tumor parts was compared using fast FLAIR and fast spin echo (FSE) sequences. Visual delineation was possible in six of nine tumors given fast FLAIR and in four of nine tumors given FSE. This visual improvement implies that fast FLAIR, or even the original FLAIR sequence should be exploited in the future. Also, since the FLAIR sequence is designed to suppress the free water signal, CSF will be more easily segmented from edema and other bound water tissues. It may be added as a fourth feature in our work without equipment upgrades whereas PWI or DWI could be additional features given necessary equipment upgrades.

A. Utilizing Fuzzy Logic in Medical Imaging

Fuzzy logic has been used in a number of ways to add intelligence when analyzing images. Recently, a fuzzy logic based edge detector which makes use of human-like rules has been discussed in [2]. In earlier work on brain image segmentation, we have used a fuzzy edge detector [8] with reasonable success. Fuzzy edge detection has also been used in cardiac imaging [30]. Other work has used fuzzy logic to determine whether an object is above or below another in a scene [38]. A description of the use of fuzzy sets and logic in high-level image analysis is given in [19], [26].

In our work on using intelligence in the analysis of images [8], [6], [7], we have used fuzzy sets primarily in grouping objects into classes. Variations of the original fuzzy c-means clustering algorithm have been used [11], [21], [25], [14]. In particular, in this paper we use fuzzy c-means clustering to group voxels into homogeneous groups which can be identified. For medical imaging in particular, we have found that fuzzy clustering gives us consistent partitions. While sensitive to initialization, it is not nearly as sensitive as hard c-means in our experience [17].

The concept of fuzzy sets is used in our work when a search for a specific shape is carried out. In this paper, we search for the ventricles in axial magnetic resonance images of the human brain. In slices near the top of the head, the ventricles present themselves as an approximate or fuzzy X shape. Our system allows for elasticity of the image which the ventricles present and they are located unerringly.

For visualization, we have found that membership values of pixels in fuzzy clusters allow you to display colored regions which fade into adjacent regions [16]. This is a useful feature in generating ground truth images. A clustered image which provides a reasonable image partition with fuzzy boundaries can be presented to a physician who can then focus on correcting the boundary regions [10].

III. Images and knowledge guided algorithms

A. MR patient data

Six data sets from four patients with cerebral tumors were acquired on two 1.5 Tesla GE Signa MR scanners (General Electric Company, Milwaukee, WI) from two institutions. Patient volume 6 was acquired through the University of Texas, Galveston Medical Branch. The remaining five volumes were acquired through the University of South Florida, Department of Radiology. Axial, 2D, 5 mm thick slice images at the level of and superior to the lateral ventricles were acquired with a 256x192 acquisition matrix, field of view of either 240 mm (volumes 1 and 2) or 220 mm (volumes 3 - 6) and with one full excitation (NEX=1). Volumes 1-5 are contiguous slice data while volume 6 was collected with a slice gap of 2 mm. All T1-weighted images were acquired using a standard spin-echo (SE) sequence while PD and T2-weighted images were collected using a fast spin-echo (FSE) sequence. Table I defines the imaging parameters used for each volume. All patient images are of non-enhancing brain tumors; where none of the six presented with hyperintense tumor regions in T1, post-contrast images, however, all T1-weighted images used for classification were those collected before the contrast agent gadolinium was administered. Therefore, image misregistration, due to moving the patient table during an imaging session, is not an issue. Also, gentle head restraints were used during imaging, minimizing potential registration errors due to patient movement.

TABLE I

MR Image Parameters. (TR/TE)

Volume Number	Patient Number	T1-weighted (SE)	PD-weighted (FSE)	T2-weighted (FSE)	Thickness/Gap (mm/mm)	FOV (mm)
1	1	650/11	4000/17	4000/102	5/0	240
2	1	650/11	4000/17	4000/102	5/0	240
3	2	500/27	3000/17	3000/102.2	5/0	220
4	2	500/27	3000/17	3000/102.2	5/0	220
5	3	500/27	3000/17	3000/102.2	5/0	220
6	4	500/18	3500/16	3500/85	5/2	220

Specified in Table II are tumor type, the number of slices in each patient volume and a breakdown of how many slices within each volume contain tumor. Volumes are also identified as belonging to either training or testing sets. Two volumes from two different patients were chosen as training data. Within these volumes one tumor was accompanied by necrosis while being partially merged with the lateral ventricles, and the other was a very large tumor without any visible necrosis. Knowledge, such as spatial qualities and feature distributions of tumor tissues, was extracted from the training data while the testing data was used to verify the success of the method. The number of slices in the training data set was 14 while the testing data set contained 31 slices for a total of 45 slices.

B. Knowledge guided algorithms

The description of the knowledge guided approach for non-enhancing tumor classification is separated into two subsections. The first includes the unsupervised classification method FCM, and the second includes a description of image processing methods used and how the system uses acquired domain knowledge.

B.1 FCM clustering

Three weighted MR intensity images provide the features for FCM clustering. Each pixel is associated with a 3D feature vector of its T1, PD and T2 weighted values. An example axial slice through the lateral ventricles is represented by the three weighted images in Figure 1. The FCM clustering method is an unsupervised classification method since it does not require any hand labeled data. FCM groups pixels by iteratively calculating a set of C cluster centers and optimizing an objective function until a stopping criteria has been reached. Once clustering is completed, each image pixel belongs to each class with a fuzzy membership value between 0 and 1. The fuzzy membership values represent the fact that one pixel may partially belong to more than one class. When a membership value of a pixel is one within a particular class, it is believed that the pixel contains only that tissue class, whereas membership values of zero represent a pixel not containing that particular tissue class. The class with the maximum membership value is chosen as the class label for a given pixel. Fuzzy membership values are interesting in MR segmentation since a slice

through brain tissues represents a projection of the slice thickness onto a two-dimensional plane, contributing greatly to the partial volume effect (i.e. a pixel may contain more than one tissue). This effect could be reduced with thinner slices and a smaller field of view (FOV). Currently, with a 240 mm FOV and a 5 mm slice thickness, the pixel dimensions are approximately 0.94 mm x 0.94 mm x 5.0 mm. Obviously, we would wish to reduce the slice thickness since the partial volume effect would be significantly reduced. Thinner slices may be obtained in the future, at the possible cost of a lower signal to noise ratio.

Another attribute of FCM is that it tends to cluster data into equal sized clusters. Therefore, in order to obtain good segmentations (or clusters of primarily one tissue type), each clustering step should require that the number of classes be larger than the actual number of distinct tissue types [28]. Hence, our initial clustering step uses ten classes.

Extra-cranial tissues are removed after clustering using existing methods [6], [22] and we re-cluster the remaining data into seven classes. Figure 2 is a slice showing both the ten class segmentation and the seven class segmentation of the intracranial tissues. Note the difficulty in distinguishing between the pathology tissues and the CSF in the lateral ventricles as well as other tissues in and around the ventricles (e.g. corpus collosum, choroid plexus, fornix and septum pellucidum) based solely on class information. With this in mind, expert knowledge of shape and physical locations of normal brain tissues, in addition to class information, can aid in identifying normal tissues, such that all abnormal tissues may be identified as the remaining tissues occupying *feature space* in anticipated locations. For example, lateral ventricles are expected to 1) be located near the center of each slice containing them, 2) contain CSF, among other tissues, and 3) have a certain range of shapes depending on the location of the axial slice within the head. All locations of cluster centers, shapes and physical locations of tissue structures are anticipated due to observations of the training data. Acquiring this knowledge is a process called knowledge engineering used in an effort to train a KB system. Once the knowledge is extracted from the training data, it may be incorporated into the KB system in the form of heuristics. Heuristics are rules of thumb which may be expressed in relative or fuzzy terms such as "higher end of the PD spectrum," or "the three highest PD classes." Both expressions refer to the location of PD cluster centers in feature space.

TABLE II

DATA DESCRIPTION.

Patient Volume	Train/ Test	Tumor Type	Tumor Slices	Non-Tumor	Number Slices
1	Train	(Anaplastic) Astrocytoma	6	2	8
2	Test	(Anaplastic) Astrocytoma	6	2	8
3	Test	Astrocytoma	6	1	7
4	Test	Astrocytoma	5	0	5
5	Test	Astrocytoma	7	4	11
6	Train	Glioblastoma-Multiforme	6	0	6
Training data:			12	2	14
Testing data:			24	7	31
Total slices:			36	9	45

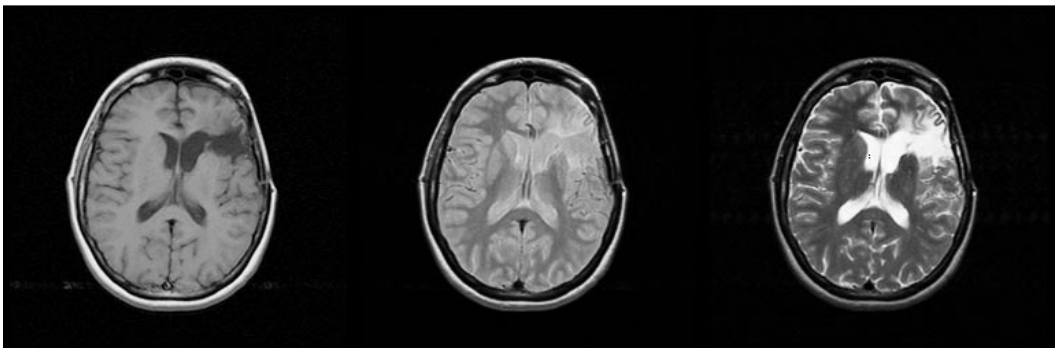

Fig. 1. Sample raw data from a patient volume used for FCM clustering. Images shown are T1, PD and T2-weighted from left to right.

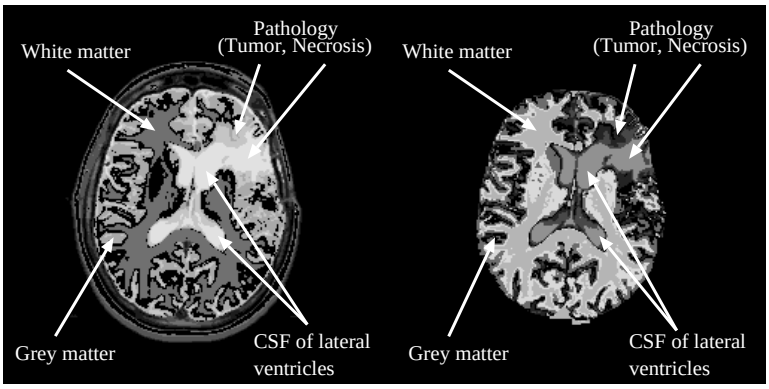

Fig. 2. Sample FCM clustered data. Images shown are class maps after image was clustered into ten classes (left), then re-clustered into seven classes after extra-cranial tissues were removed (right).

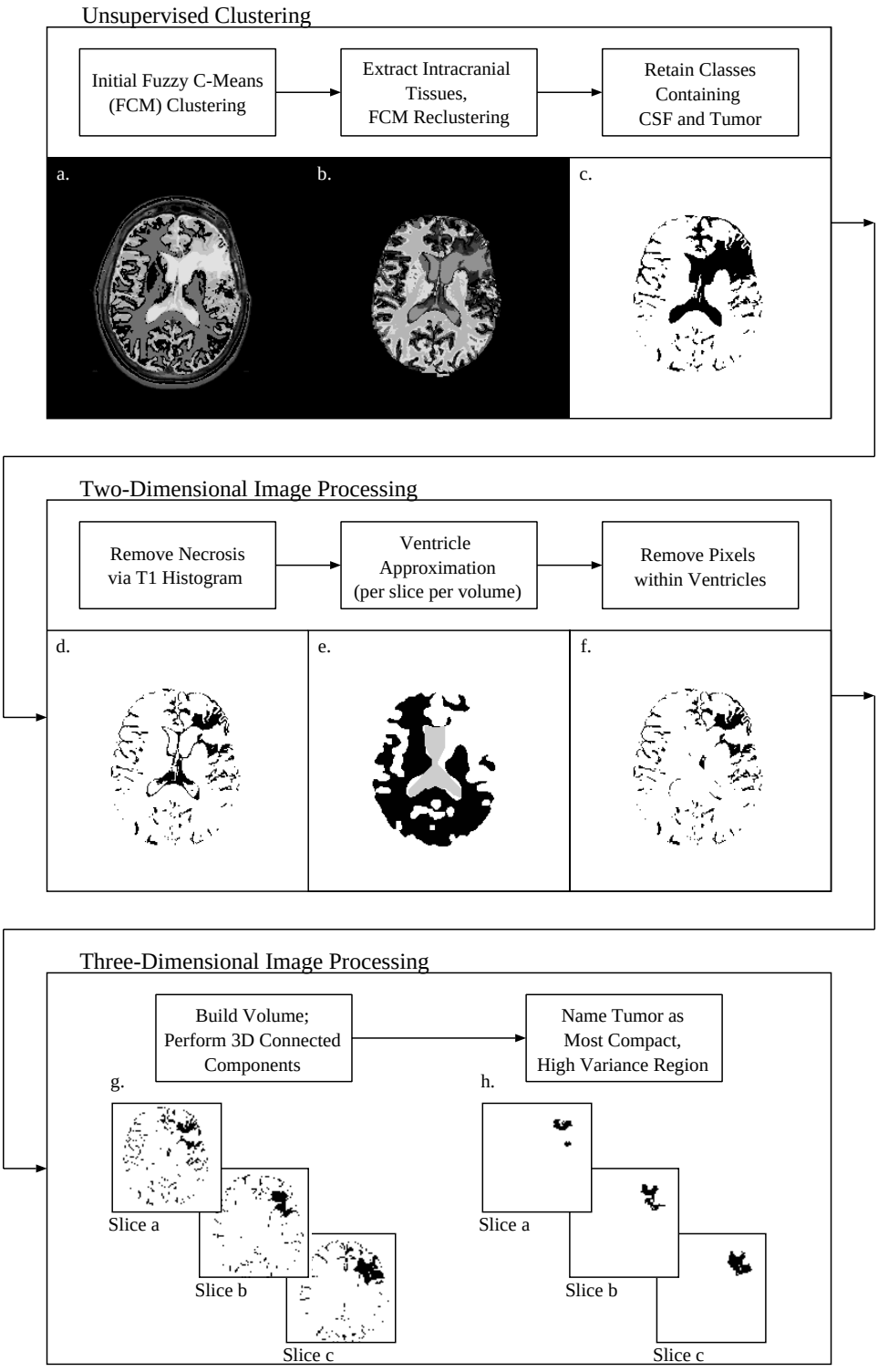

Fig. 3. An overview of the entire non-enhancing tumor classification method. The first phase of this three phase process begins with FCM clustering into ten classes (a), removal of extra-cranial tissues and re-clustering of the remaining tissues into seven classes (b) and removal of white and grey matter (c). The second phase uses histogram analysis in T1 feature space to remove any necrosis from the pathology image (d). Two-dimensional image processing techniques and knowledge of CSF shape and ventricle location eliminate large CSF regions from the image (e)-(f). Finally, three-dimensional connected components and a compactness measure aid in the final tumor detection (g)-(h).

Plotting the feature space of the training data, showed that in all slices, the tumor and CSF were separable from white and grey matter, but not from each other without misclassifying some tumor pixels. Since all pathology, as well as most of the CSF appear within the three PD clusters whose centroid values are highest, when the seven cluster centers are ranked in PD feature space, our system retains the CSF and pathology pixels by storing pixels belonging to these three clusters in a binary image to be used as a mask for further processing. The remaining classes, consisting mostly of white and grey matter tissues, are also stored in a binary image to aid in the ventricle approximation procedure discussed in Section III-B.2.b.

B.2 Knowledge-guided image processing steps

Figure 3 demonstrates how the image processing steps fit into the non-enhancing tumor classification system. The 2D and 3D image processing, when completed, aid in the decision of tumor location within the volume. This system operates under two assumptions. First, there is one tumor somewhere within the patient volume. Second, the tumor exists in at least three consecutive slices. The first assumption is fair since this system was designed to be added onto an existing tumor segmentation system for enhancing tumors [6], [7] and activated if deformation exists but a tumor is not located. Also, since this KB system is intended to be used to track tumor growth/shrinkage, the number of tumors existing in each volume will be known before volume measurements are calculated. If more than one tumor exists, that knowledge will be taken into consideration. The latter assumption, with present data of 5 mm thick slices, specifies that non-enhancing tumors must be at least 1.5 cm deep, perpendicular to the axial slice, or parallel to the Z direction, for reasons to be discussed within Section III-B.2.c. Knowledge obtained from the training data, such as the method for ventricle approximation, has been incorporated into a rule-base which is accessed by CLIPS [29], a rule-based expert system tool.

B.2.a Necrosis removal. After the binary CSF/pathology image is obtained by retaining the three classes with the highest relative proton density cluster centers of the seven class re-clustered image, we may remove any existing necrosis via the procedure shown pictorially in Figure 4. First, a histogram of the T1 space for each slice is generated. Each

histogram exhibits either uni-modal or bi-modal characteristics. However, since these histograms are of real data containing noise it is difficult to algorithmically determine if the histogram is uni or bi-modal. Therefore, the histogram is smoothed repeatedly with a Gaussian filter of sigma = 1.5, until there are two peaks or more, but never less than two peaks. Since the possibility exists for the filtering process to smooth a histogram with more than two peaks into a uni-modal shape, we anticipate a bi-modal shape in order to maximize the benefit of the necrosis removal procedure. If only one peak is detected, we instead use the previously smoothed histogram which by definition has more than two peaks.

Second, we locate the local minimum which separates the first peak from the second (or remaining peaks in the case of a multi-modal histogram) and use that value as a threshold. All pixels below the threshold are eliminated in the CSF/pathology image with one stipulation, the threshold must be less than half of the histogram pixel value range. If it is not then we assume that the true nature of the histogram is uni-modal and we do not remove any pixels from the CSF/pathology image. Since the necrotic tissues exhibit a very low intensity signal in the T1 space while tumor exhibits a range of high intensity signals, this successfully removes most pixels containing necrotic tissues from further processing. We refer to this new binary image as a CSF/tumor image.

As an example of rule construction, the necrosis removal procedure is represented by:

```
REPEAT
    IF (T1 histogram peaks > 2)
    THEN (Smooth histogram)
    ELSE IF (T1 histogram peaks = 2)
    THEN (Smoothing complete)
        (Break Repeat)
    ELSE (Retrieve previous T1 histogram)
        (Smoothing complete)
        (Break Repeat)
END REPEAT
```

IF (Smoothing complete)
THEN (Locate minimum between first 2 peaks)
(Remove pixels from image below minimum)

B.2.b Ventricle approximation and removal. Once a new binary CSF/tumor image has been obtained minus any necrosis, we attempt to remove the ventricles for each slice containing them. The first step in obtaining the ventricle approximation is to retrieve the white and grey matter classes from the previous clustering step. Since white and grey matter (or parenchymal tissues) should surround the lateral ventricles [15], we can estimate the shape and size of the ventricles by measuring the position of the white or grey matter pixels closest to the center of the slice in the binary white/grey matter image. This is done by emitting radial lines every 15 degrees, from the approximate center of the head. Figure 5 demonstrates this concept. Radial lines that intersect either the white or grey matter, mark the position of the closest white or grey matter pixel on their trajectory. When a radial line does not intersect a white or grey matter pixel, its mirror trajectory (given a vertical axis reflection plane) is examined next. If there was a detection along the mirror trajectory, then that point is mirrored back to the original trajectory. If neither line intersected white or grey matter, then the average length of their nearest neighboring emitter lines which intersected white or grey matter is used. Once all lines have been created, the end points of emitter lines are translated into Cartesian coordinates. The lengths of mirror trajectories are compared until each pair represents the same length by choosing the shortest length of the pair. Once all pair lengths have been determined, the new points are ordered by trajectory angle and a closed polygon is drawn and filled. Slices that do not contain ventricles result in the formation of very small polygons located near the center of the head. Usually, only a few pixels are contained within the assumed ventricles, hence only a few pixels may be removed. CSF, tumor, both or background pixels will make up the region of the polygon depending on the contents of the CSF/tumor image. The worst case would be if all pixels in the location of the small polygon are tumor pixels, which is possible if the tumor is located in the center of the slice. This problem however, has not been encountered.

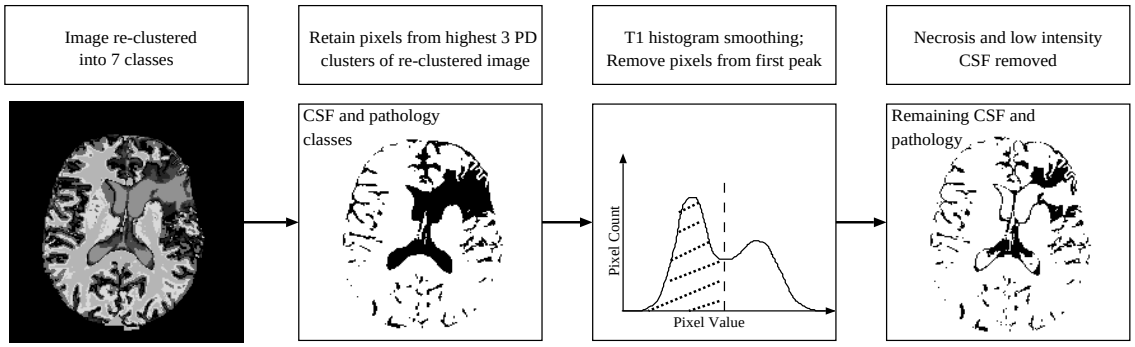

Fig. 4. Necrosis is removed by isolating the low intensity peak of the CSF/pathology T1 histogram. This histogram is first smoothed until only two peaks remain then the minimum between the two peaks is used as a threshold. The low intensity peak (marked with dashed lines) is removed, leaving tumor and some CSF.

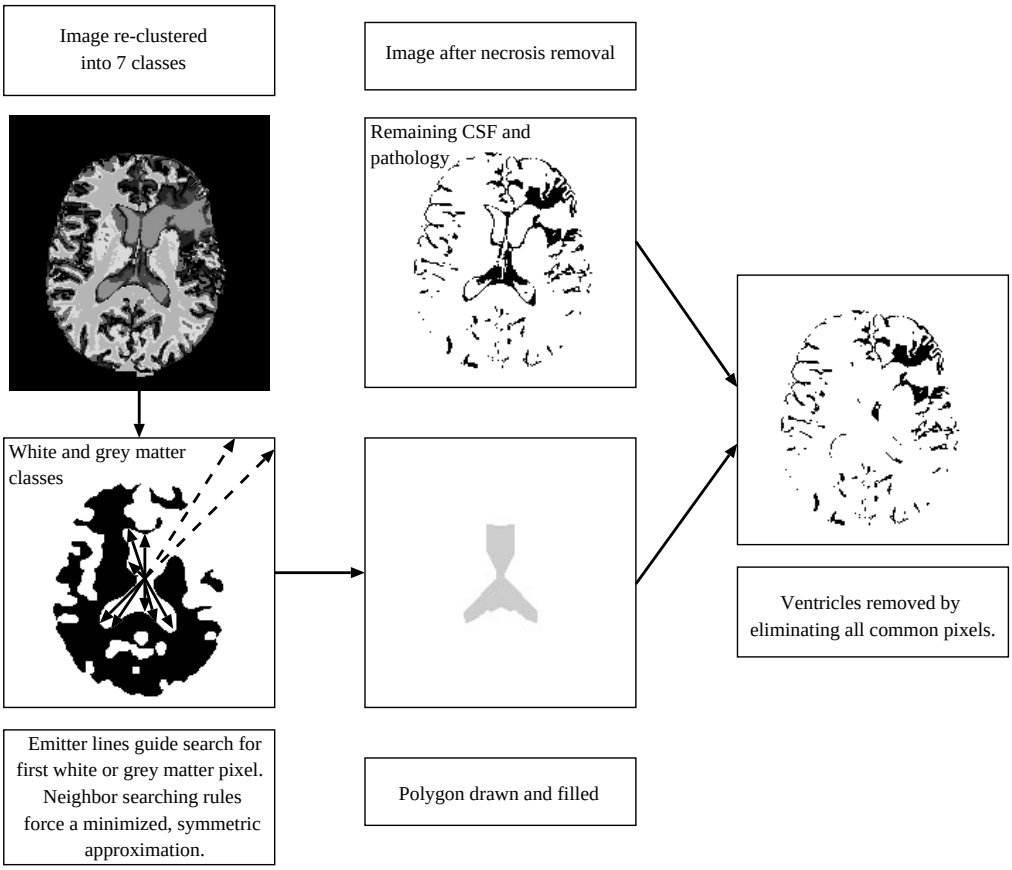

Fig. 5. Ventricle removal begins with collecting the white and grey matter classes from the seven cluster class-map. Emitted lines radially span the white/grey matter image and mark detected boundaries (solid radial lines). Undetected boundaries (dotted radial lines) are approximated via their "mirror" radial lines and close neighbors. The result is a list of dot locations which, when connected, approximate the shape of the lateral ventricles. Note the polygon is symmetrical along the vertical axis. Next, any common pixels in both the ventricle approximation and the image remaining after necrosis removal described in Figure 4 are removed.

We have deliberately generated a polygon which is symmetrical along the vertical axis for two reasons. First, since lateral ventricles in a healthy brain will be near symmetrical across the longitudinal fissure separating the left and right hemispheres, a good approximation of healthy ventricles should be symmetrical across the same axis. Obviously, approximation errors may be incurred when unhealthy ventricles are no longer symmetrically shaped. For example, asymmetric hydrocephalus (an abnormal enlargement of the ventricles which presents unevenly from the left to the right ventricle) is caused by atrophy of the cerebral cortex, and therefore, may accompany brain tumors. For this reason, we duplicate the smaller ventricle for removal in order to prevent the unintentional removal of tumor pixels. Second, for cases where it is difficult to detect a spatial boundary between ventricles and tumor due to their close proximity, we wish to approximate healthy ventricles so the removal of the lateral ventricles is possible. The ventricles are then removed by eliminating all foreground pixels common to the polygon and the CSF/tumor image.

B.2.c 3D processing. Next, a binary volume image is built by stacking consecutive slices of the remaining CSF/tumor pixels. Note, in Figure 4 and Figure 5, isolated pixels are distributed in each slice where CSF surrounds grey matter and sits within the sulci. A 3x3x3 morphological operator is used to convert pixels to background that do not have at least 17 of 26 neighboring pixels in the foreground (simulating two layers within the boundaries of the operator). This threshold is large enough to be effective in removing very small groups of pixels, but small enough to prevent the removal of all foreground pixels in a slice of the volume unless its neighboring slices do not contain foreground pixels in the same location. This helps reduce the number of non-tumor regions found during the following connected components operation. A 3D connected components algorithm labels groups of connected pixels in the volume as regions. Two statistical values are calculated for each region, a pixel count and the variance of the original intensity values in the PD weighted image. Next, a 3x3x3 erosion operator is used on the volume. This erosion operation leaves a pixel as foreground if all neighboring pixels are also foreground. In other words, a foreground pixel centered in the erosion window, remains as foreground only if all 26 pixels surrounding the center pixel are also foreground. Therefore, the tumor must be present in at least three consecutive slices due to the nature of the 3D erosion operator, or at least 1.5

cm in the Z direction, as was mentioned in Section III-B.2, and at least 119 mm^3 given a 240 mm FOV or 100 mm^3 given a 220 mm FOV, both with a 5 mm slice thickness. Pixels remaining within the boundaries of each original region after the erosion are counted and the percent of pixels retained governs which 3D region is chosen as the tumor. If there is one region with the highest percentage of pixels remaining, that region is chosen due to observations that tumor regions tend to be more spatially compact than CSF regions. Otherwise, the additional criteria is to choose the region with a higher variance. Statistical variance is pertinent since large regions remaining are either parts of ventricles which were missed during the ventricle removal stage or tumor tissues. The ventricle portions tend to be homogeneous due to the compartmental nature of brain/CSF boundaries whereas the tumor tissues are infiltrating other brain tissues, possibly exhibiting more partial volume effects.

IV. Results and Discussion

Our system was trained on two volumes and tested on four volumes previously unseen by the system. A total of 45 slices were processed. All tumor sizes in the volumes examined, exceed the previously stated lower bound of 100 mm^3 for a 220 mm FOV and 119 mm^3 for a 240 mm FOV. Again, smaller tumors would not be detected using the current implementation for the fixed slice thickness. Results are listed in Tables III and IV, where Table III qualitatively describes results on a per slice basis and Table IV quantitatively displays results on a pixel level per volume.

In Table III it is shown that of the 45 slices, 40 were labeled by the system as containing tumor. Thirty-five of the 40 slices were positively identified by a radiologist as containing tumor pixels, yielding five false positives, where one of the five was a slice with ground truth tumor pixels but none of those pixels were actually detected. Hence, all slices with tumor were detected. The slice that was both a false positive and a false negative corresponds to slice 20 of volume 1, where 40 pixels were believed to be tumor but none were shared by the 77 ground truth pixels. We chose to represent this slice as shared between the false positive and false negative category, both marked by the asterisk. Examining this slice further, however, we found that a convex hull measurement of each small "tumor" region overlaps with the other and the centroid of each is separated by only 11 pixels. Any false

positives may be caused by the 3D connected components algorithm since slices directly above and below the tumor may contain CSF pixels, which the algorithm connects to the tumor. Alternatively, any false negatives may be caused by the use of the morphological operators used during processing to remove small isolated groups of pixels.

Also seen in row three of Table III the ventricle approximation and removal phase contributed two positive results to the segmentation method. First and foremost, if the CSF ventricles are merged with the tumor, ventricles were removed from the image, hence, separating the two histologically distinct regions via the spatial domain. Secondly, by removing the only other large regions in all images, the most spatially compact region is the tumor based on the method of measurement. Occasionally, however, in lower slices when the ventricle approximation fails due to the presence of white or grey matter pixels near the center of the brain, the variance feature is needed to verify the correct tumor region. Note that all slices exhibiting the characteristic of the lateral ventricles merged with the tumor area were successfully separated from each other.

TABLE III

TUMOR CLASSIFICATION RESULTS. NUMBERS ARE LISTED AS: TOTAL (TRAIN/TEST)

	Ground Truth	Number Detected	True Positives	False Positives	False Negatives
Slices with Tumor	36 (12/24)	40 (13/27)	35 (11/24)	*5 (2/4)	*1 (1/0)
Slices without Tumor	9 (2/7)	5 (1/4)	5 (1/4)	0 (0/0)	4 (1/3)
Separation of Merged CSF and Tumor	9 (5/4)	9 (5/4)	9 (5/4)	0 (0/0)	0 (0/0)

Example worst and best case tumor segmentations are shown in Figure 6 which correspond to slice 14 of volume 2 and slice 20 of volume 5. Depicted are the raw data images beside the radiologist labeled ground truth image and the knowledge base segmented tumor.

Table IV displays the results for each volume on a pixel level. Ground truth was obtained from the boundary drawings of a radiologist, which were later looked at and verified by the same radiologist. It should be noted, however, that the task of drawing tumor boundaries is very difficult without the often unavailable histopathological findings required for absolute verification [9], [13]. This is compounded by the fact that all tumors in this study were non-enhancing. Also, Wells et al. [39] have shown that inter and intra-observer variability can be very high (inter: 17-21% among five experts, intra: 18-25% two ratings obtained with two methods) showing that the task of hand-labeling ground truth is very complex, and the correlation between ground truth and any automatic labeling system goes beyond pixel level comparisons and into the repeated trends of the tissue estimations.

The 3D tumor regions which were identified by the knowledge base were then compared to the pixel level ground truth on a per slice basis. Listed in Table IV are the number of true positives (ground truth tumor pixels found algorithmically), false positives (pixels isolated as tumor though not within ground truth boundaries) and false negatives (ground truth tumor pixels not found algorithmically).

From these three measures we chose to represent overall results with two additional calculations. The correspondence ratio (CR) shown in Equation 1 allows us to discuss the way in which the isolated tumor corresponds in size and location to the ground truth tumor while weighting the importance of false positives and false negatives. It is calculated by dividing the number of ground truth pixels (or all tumor pixels) into the difference of the true positives plus one half times the number of false positives. Although a correspondence ratio of one is ideal, any value can be very descriptive since it negatively weighs the value of the true positives with 50% of the false positives, indicating that a false positive is detrimental only to a certain degree. Specifically, a negative correspondence ratio indicates a greater than two to one rate of false positives with respect to true positives. Since we are interested in measuring tumor volume and location for the purpose of treatment, it

is intuitive that the immediate area around the tumor will also be treated, hence the weighting mechanism in our correspondence ratio. Also note that in Table IV we express the correspondence ratio as: 1.0 if the system identified no pixels as tumor where none existed in ground truth, and 0.0 if the system identified any pixels as tumor where none existed in ground truth.

$$CR = \frac{TruePositives - (0.5 * FalsePositives)}{GroundTruthTumor} \tag{1}$$

Percent match, shown in Equation 2, is calculated as the direct ratio of the true positives to the number of ground truth tumor pixels. An ideal percent match is 100%, with a value of zero indicating that there was a complete miss of any ground truth pixels for that slice. If the number of ground truth pixels is 0, then a percent match of 100% is shown in Table IV.

$$PercentMatch = \frac{TruePositives}{GroundTruthTumor} * 100\% \tag{2}$$

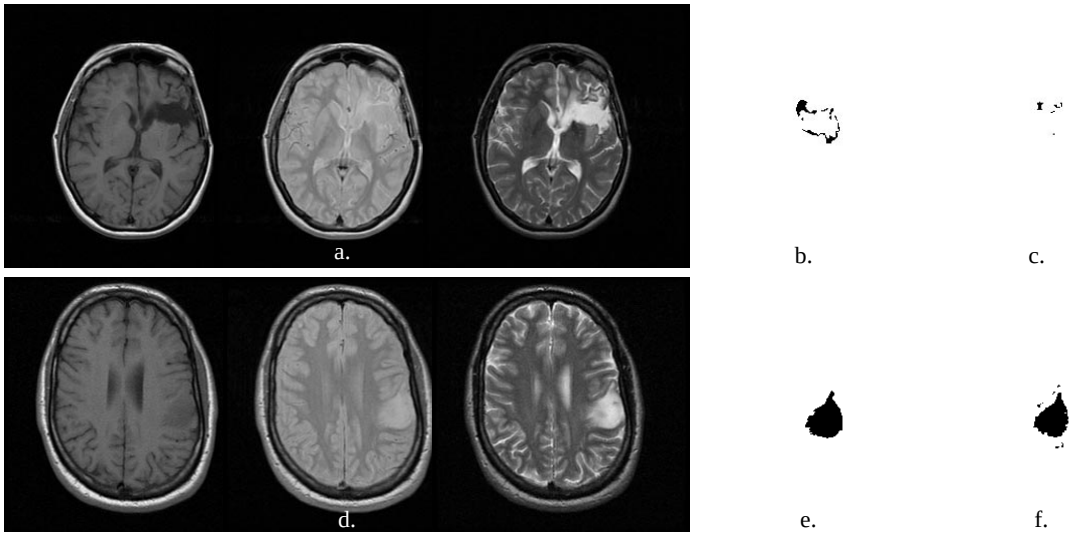

Fig. 6. Example results for two slices from worst case (top) and best case (bottom) volumes comparing ground truth to KB tumor segmentation. Poor performing slice from worst case volume 2 (a)-(c) and good performing slice from best case volume 5 (d)-(f). Shown are raw data for each (a) and (d), ground truth tumor (b) and (e), and KB tumor segmentation for slice shown (c) and (f). Slices are shown for ease of viewing, noting that full volumes were labeled by the KB system as tumor.

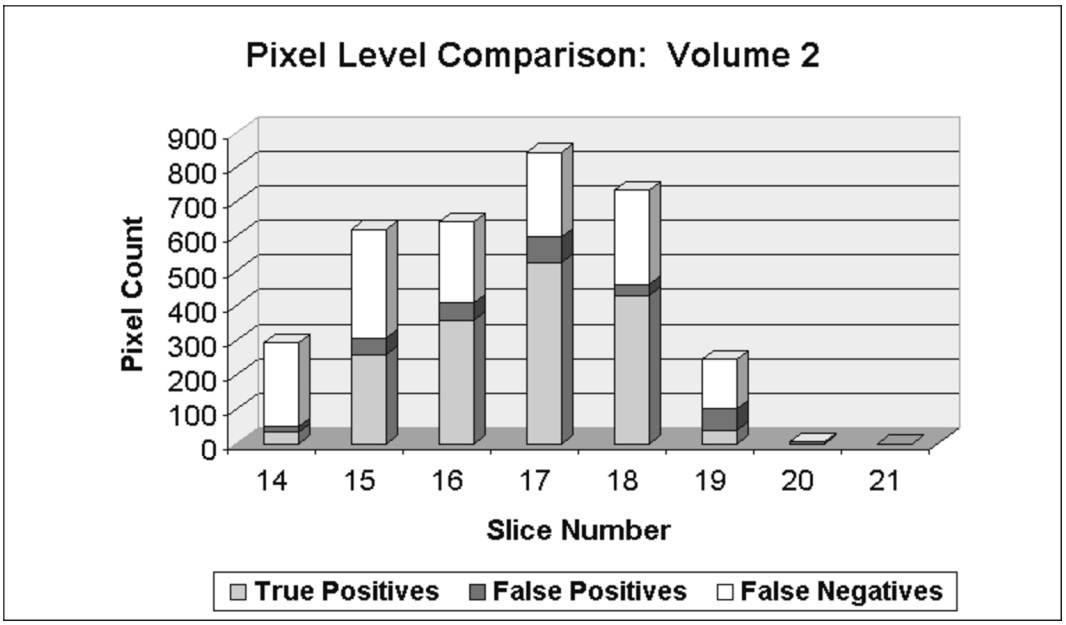

Fig. 7. Volume showing the worst results.

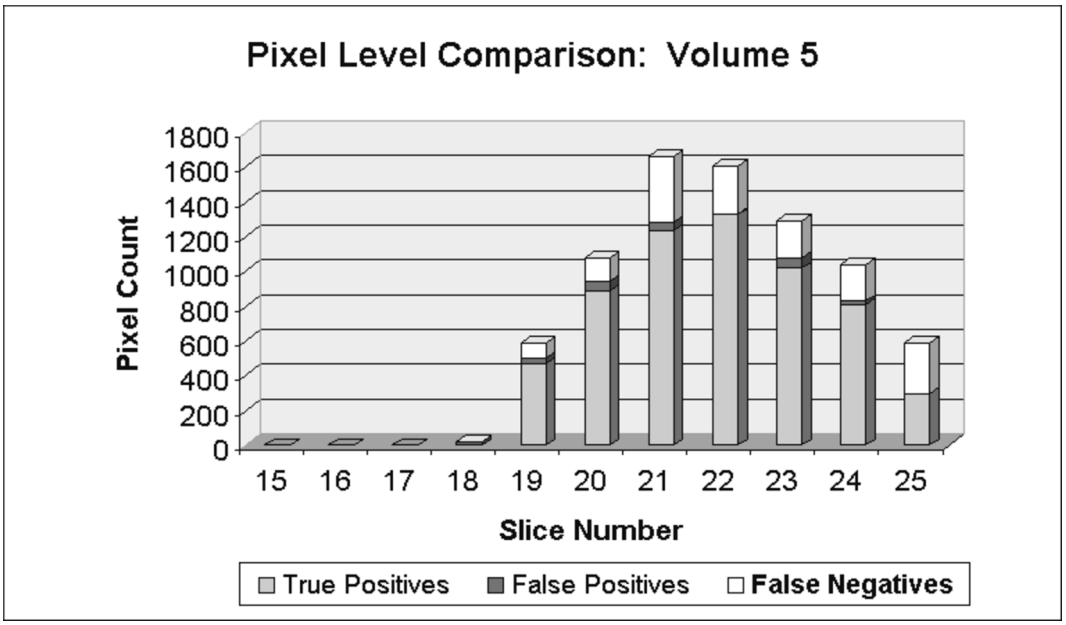

Fig. 8. Volume showing the best results.

Volumes 2 and 3 represent the worst of our testing results while volume 5 represents our best testing results. Although both volumes 2 and 3 have one slice each where no true positives are found, and both volumes have low correspondence ratios, (0.484 for volume 2 and 0.368 for volume 3) we chose to show volume 2 in Figure 7 as our current worst case results since volume 3 has a much higher percent match value (81.1%) than does volume 2 (53.0%). Volume 5 demonstrates our best results clearly seen by the high correspondence ratios for all but two slices of 11 and high percent match values for all but two slices. The overall correspondence ratio and percent match for the slice are 77.6% and 79.2% respectively. Slice number 18 is considered a complete false positive since 23 pixels were believed to be tumor where none existed in the ground truth image.

While these numerical values and calculations for volumes 2 and 5 fairly represent the data, Figures 7 and 8 graphically show true positive, false positive and false negative pixel counts for the two volumes mentioned above. Figure 7 represents some of our worst results, as seen by the high number of false negatives in each slice. Also, in slice 20, six pixels were believed to be tumor but none existed in the ground truth image, therefore, only false positives and no true positives exist for that slice. This error may be attributed to the usage of the 3D connected components algorithm which tends to connect any pixels in the slices above and below the tumor before the 3D erosion operator is used. Although the correspondence ratio and percent match values seem low, success of our method can also be measured as having correctly detected tumor pixels in six of all six slices containing tumor and correctly identifying one slice of two that contain no tumor pixels in the ground truth image.

Figure 8 shows some of our best results. Slices 15 - 17 each have full correspondence and percent match since, no tumor existed in these slices and none was found. In slice 18, we again detected pixels, admittedly only 23, where none exist in the ground truth image. This again would be due to the usage of the 3D connected components algorithm. Tumor regions within slices 19 - 25 were all detected with correspondence ratios above 0.75 in all but slice 25, and percent match values were above 75% in all but slice 25. The pixels missed in slice 25 were all towards the exterior of brain, some of which were removed in very early stages of processing during the extra-cranial removal stage and we may consider

ways to recover such pixels in the future. The volume correspondence ratio is equally high at 0.776 with a percent match of 79.2%.

TABLE IV

Pixel Level Comparisons

Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
1	15	52	38	349	401	90	0.082	13.0
Train	16	295	52	266	561	347	0.480	52.6
	17	421	85	249	670	506	0.565	62.8
	18	498	21	254	752	519	0.648	66.2
	19	284	76	164	448	360	0.549	63.4
	20	0	40	77	77	40	-0.260	0.0
	21	0	15	0	0	15	0.000	100.0
	22	0	0	0	0	0	1.000	100.0
Volume Totals:		1550	327	1359	2909	1877	0.477	53.3

Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
2	14	33	16	242	275	49	0.091	12.0
Test	15	257	48	313	570	305	0.409	45.1
	16	357	49	235	592	406	0.562	60.3
	17	524	72	245	769	596	0.635	68.1
	18	428	31	273	701	459	0.588	61.1
	19	39	61	143	182	100	0.047	21.4
	20	0	6	0	0	6	0.000	100.0
	21	0	0	0	0	0	1.000	100.0
Volume Totals:		1638	283	1451	3089	1921	0.484	53.0

Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
3	15	252	254	148	400	506	0.313	63.0
Test	16	194	686	83	277	880	-0.538	70.0
	17	622	313	58	680	935	0.685	91.5
	18	549	215	81	630	764	0.701	87.1
	19	145	267	30	175	412	0.066	82.9
	20	24	203	15	39	227	-1.987	61.5
	21	0	13	0	0	13	0.000	100.0
Volume Totals:		1786	1951	415	2201	3737	0.368	81.1

Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
4	16	687	319	123	810	1006	0.651	84.8
Test	17	636	861	29	665	1497	0.309	95.6
	18	1053	533	137	1190	1586	0.661	88.5
	19	774	481	37	811	1255	0.658	95.4
	20	429	48	73	502	477	0.807	85.5
Volume Totals:		3579	2242	399	3978	5821	0.618	90.0

Pixel Level Comparisons, continued								
Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
5	15	0	0	0	0	0	1.000	100.0
Test	16	0	0	0	0	0	1.000	100.0
	17	0	0	0	0	0	1.000	100.0
	18	0	23	0	0	23	0.000	100.0
	19	470	34	87	557	504	0.813	84.4
	20	890	52	137	1027	942	0.841	86.7
	21	1238	43	381	1619	1281	0.751	76.5
	22	1327	6	275	1602	1333	0.826	82.8
	23	1019	59	210	1229	1078	0.805	82.9
	24	811	18	209	1020	829	0.786	79.5
	25	295	0	294	589	295	0.501	50.1
Volume Totals:		6050	235	1593	7643	6285	0.776	79.2

Patient Volume	Slice	True Positives	False Positives	False Negatives	Ground Truth	Total Tumor Isolated	Correspondence Ratio	Percent Match
6	4	412	511	144	556	923	0.281	74.1
Train	5	1716	338	85	1801	2054	0.859	95.3
	6	3125	141	185	3310	3266	0.923	94.4
	7	3502	22	331	3833	3524	0.911	91.4
	8	3040	137	241	3281	3177	0.906	92.7
	9	2383	35	427	2810	2418	0.842	84.8
Volume Totals:		14178	1184	1413	15591	15362	0.871	90.9

V. Conclusions

Our results show that this method has great potential for solving the difficult problem of segmenting, localizing and measuring volumes of non-enhancing tumors. We automatically detected all six non-enhancing brain tumors, and located them in 35 of the 36 ground truth slices containing tumor. We also successfully separated CSF regions merged with tumor in nine of nine slices. The correspondence ratio calculation which negatively weights half the number of false positives against the true positives ranged from 0.368 - 0.871 per volume indicating that a fair level of correspondence exists between ground truth measurements and tumor pixels isolated via our segmentation method. Generally high percent match calculations, ranging from 53.0% to 90.9% per volume, indicate that high percentages of ground truth pixels were identified as tumor pixels by our segmentation method.

False positives and false negatives may be attributed to the complexity of defining the ground truth tumor regions as well as the difficulty in determining tumor boundaries algorithmically. Partial volume effects of course limit the accuracy of any method. Thinner slices will improve results greatly in terms of partial volume restrictions and the use of the 3D erosion operator in the tumor detection process, since the erosion operator reduces the tumor depth by two slice thicknesses. Therefore, if a tumor does not exist in at least three consecutive slices (regardless of the slice thickness) that tumor region will have no chance of passing the compactness criteria due to the effect of the 3D erosion operation.

Additional improvements are needed to refine the tumor approximation method as there is potential to miss the ventricles if a white or grey matter pixel is near the center of the head. Eventually, we will approximate the size and shape of the lateral ventricles only in the slices which contain them. Once the consistency of our method is established, we will have completed the important first step in successful tumor volume tracking over time.

Acknowledgments

The authors would like to thank G.R. Hillman of the University of Texas, Galveston Medical Branch and L.P. Clarke and R.P. Velthuizen of the University of South Florida Department of Radiology for supplying all patient volumes.

References

[1] H.J. Aronen, I.E. Gazit, D.N. Louis, B.R. Buchbinder, F.S. Pardo, R.M. Weisskoff, G.R. Harsh, G.R. Cosgrove, E.F. Halpern, F.H. Hochberg, and B.R. Rosen. Cerebral blood volume maps of gliomas: Comparison with tumor grade and histologic findings. *Radiology*, 191(1):41–51, 1994.

[2] J.C. Bezdek, R. Chandrasekhar, and Y. Attikiouzel. Geometric approach to edge detection. *IEEE Transactions on Fuzzy Systems*, 6(1):52–75, 1999.

[3] J.C. Bezdek, L.O. Hall, and L.P. Clarke. Review of MR image segmentation techniques using pattern recognition. *Medical Physics*, 20(4):1033–1048, 1993.

[4] J.A. Brunberg, T.L. Chenevert, P.E. McKeever, D.A. Ross, L.R. Junck, K.M. Muraszko, R. Dauser, J.G. Pipe, and A.T. Betley. In vivo MR determination of water diffusion coefficients and diffusion anisotropy: Correlation with structural alteration in gliomas of the cerebral hemispheres. *Am. J. Neuroradiol.*, 16:361–371, 1995.

[5] M. C. Clark. *Knowledge-Guided Processing of Magnetic Resonance Images of the Brain*. PhD thesis, University of South Florida, 1997.

[6] M.C. Clark, L.O. Hall, and D.B. Goldgof. MRI segmentation using fuzzy clustering techniques: Integrating knowledge. *IEEE Engineering in Medicine and Biology*, 13(5):730–742, 1994.

[7] M.C. Clark, L.O. Hall, D.B. Goldgof, R. Velthuizen, F.R. Murtagh, and M.S. Silbiger. Automatic tumor segmentation using knowledge-based techniques. *IEEE Transactions on Medical Imaging*, 17(2):187–201, 1998.

[8] M.C. Clark, L.O. Hall, D.B. Goldgof, R. Velthuizen, R. Murtagh, and M.S. Silbiger. Unsupervised brain tumor segmentation using knowledge-based fuzzy techniques. In H-N Teodorescu, A. Kandel, and L.C. Jain, editors, *Fuzzy and Neuro-Fuzzy Systems in Medicine*. CRC Press, 1998.

[9] L.P. Clarke, R.P. Velthuizen, M. Camacho, J. Heine, M. Vaydianathan, L.O. Hall, R. Thatcher, and M. Silbiger. MRI segmentation: Methods and applications. *Magnetic Resonance Imaging*, 13(3):343–368, 1995.

[10] L.P. Clarke, R.P. Velthuizen, M. Clark, G. Gaviria, L.O. Hall, D.B. Goldgof, F.R. Murtagh, S. Phuphanich, and S. Brem. Mri measurement of brain tumor response: Comparison of visual metric and automatic segmentation. *Magnetic Resonance Imaging*, 16(3):271–279, 1998.

[11] R.N. Dave and K. Bhaswan. Adaptive fuzzy c-shells clustering and detection of ellipses. *IEEE Trans Neural Networks*, 3(5):643–662, 1992.

[12] M. Essig, H. Hawighorst, S.O. Schoenberg, R. Engenhart-Cabillic, M.V. Knopp, and G. van Kaick. Fast fluid-attenuated inversion-recovery (FLAIR) MRI in the assessment of intraaxial brain tumors. *JMRI*, 8(4):789–798, 1998.

[13] R. Galloway, R. Maciunas, and A. Failinger. Factors affecting perceived tumor volumes in magnetic resonance imaging. *Ann. Biomed. Eng.*, 21:367–375, 1993.

[14] A.B. Geva and D.H. Kerem. Forecasting generalized epileptic seizures from the eeg signal by wavelet analysis and dynamic unsupervised fuzzy clustering. *IEEE Transactions on Biomedical Engineering*, 45(10):1205–1216, 1998.

[15] D.E. Haines. *Neuroanatomy: An Atlas of Structures, Sections, and Systems*. Williams and Wilkins, fourth edition, 1995.

[16] L.O. Hall, A.M. Bensaid, L.P. Clarke, R. Velthuizen, M. Silbiger, and J. Bezdek. A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. *IEEE Transactions on Neural Networks*, 3(5):672–682, 1992.

[17] L.O. Hall, I.B. Ozyurt, and J.C. Bezdek. Clustering with a genetically optimized approach. *IEEE Transactions on Evolutionary Computation*, 3(2):103–112, 1999.

[18] B. Johnson, M. Atkins, B. Mackiewich, and M. Anderson. Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. *IEEE TMI*, 15(2):154–169, 1996.

[19] J.M. Keller. Fuzzy logic rules in low and mid level computer vision tasks. In *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS 1996*, pages 19–22, 1996.

[20] K. Krabbe, P. Gideon, P. Wagn, U. Hansen, C. Thomsen, and F. Madsen. MR diffusion imaging of human intracranial tumours. *Neuroradiology*, 39:483–489, 1997.

[21] R. Krishnapuram and J.M. Keller. Possibilistic c-means algorithm: insights and recommendations. *IEEE Transactions on Fuzzy Systems*, 4, 1996.

[22] C. Li, D.B. Goldgof, and L.O. Hall. Automatic segmentation and tissue labeling of MR brain images. *IEEE TMI*, 12(4):740–750, December 1993.

[23] K.O. Lim and A. Pfferbaum. Segmentation of MR brain images into cerebrospinal fluid spaces, white, and gray matter. *Journal of Computer Assisted Tomography*, 13:588–593, 1989.

[24] A. Lundervold and G. Storvik. Segmentation of brain parenchyma and cerebrospinal fluid in multispectral magnetic resonance images. *IEEE TMI*, 14(2):339–349, June 1995.

[25] F. Masulli and A. Schenone. Fuzzy clustering based segmentation system as support to diagnosis in medical imaging. *Artificial Intelligence in Medicine*, 16(2):129–147, 1999.

[26] S. Medasani, R. Krishnapuram, and J.M. Keller. Are fuzzy definitions of basic attributes of image objects really useful? *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, 29(4):378–386, 1999.

[27] D. Peck, J. Windham, L. Emery, H. Soltanian-Zadeh, D. Hearshen, and T. Mikkelsen. Cerebral tumor volume calculations using planimetric and eigenimage analysis. *Medical Physics*, 23(12):2035–2042, 1996.

[28] W. Phillips, S. Phuphanich, R. Velthuizen, and M. Silbiger. Automatic magnetic resonance tissue characterization for three-dimensional magnetic resonance imaging of the brain. *American Society for Neuroimaging*, 5:171–177, 1995.

[29] G. Riley. Version 4.3 CLIPS reference manual. Technical Report JSC-22948, Artificial Intelligence Section, Lyndon B. Johnson Space Center, 1989.

[30] S.K. Setarehdan and J.J. Soraghan. Automatic cardiac lv boundary detection and tracking using hybrid fuzzy temporal and fuzzy multiscale edge detection. *IEEE Transactions on Biomedical Engineering*, 46(11):1364–1378, 1999.

[31] T. Taxt and A. Lundervold. Multispectral analysis of the brain using magnetic resonance imaging. *IEEE TMI*, 13(3):470–481, September 1994.

[32] R.D. Tien, G.J. Felsberg, H. Friedman, M. Brown, and J. MacFall. MR imaging of high-grade cerebral gliomas: Value of diffusion-weighted echoplanar pulse sequences. *AJR*, 162:671–677, 1994.

[33] C. Tsai, B.S. Manjunath, and R. Jagadeesan. Automated segmentation of brain MR images. *Pattern Recognition*, 28(12):1825–1837, 1995.

[34] M.W. Vannier, R.L. Butterfield, D. Jordan, W. Murphy, R. Levitt, and M. Gado. Multispectral analysis of magnetic resonance images. *Radiology*, 154(1):221–224, January 1985.

[35] M.W. Vannier, C.M. Speidel, and D.L. Rickmans. Magnetic resonance imaging multispectral tissue classification. *News Physiol Sci*, 3:148–154, August 1988.

[36] R.P. Velthuizen, S. Phuphanich, L.P. Clarke, L.O. Hall, A. Bensaid, J. Arrington, H. Greenberg, and M. Sil-

biger. Unsupervised tumor volume measurement using magnetic resonance brain images. *JMRI*, 5(5):594–605, 1995.

[37] S. Vinitski, C. Gonzalez, F. Mohamed, T. Iwanaga, R. Knobler, K. Khalili, and J. Mack. Improved intracranial lesion characterization by tissue segmentation based on a 3d feature map. *Magnetic Resonance in Medicine*, 37:457–469, 1997.

[38] X. Wang, J.M. Keller, and P. Gader. Using spatial relationship as features in object recognition. In *Annual Conference of the North American Fuzzy Processing Society - NAFIPS 1997*, pages 160–165, 1997.

[39] W.M. Wells, W.E.L. Grimson, R. Kikinis, and F.A. Jolesz. Adaptive segmentation of MRI data. *IEEE TMI*, 15(4):429–442, 1996.