A Survey of MRI-based Medical Image Analysis for Brain Tumor Studies

Stefan Bauer[1]‡,Roland Wiest[2], Lutz-P. Nolte[1], and Mauricio Reyes[1]

[1] Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland
[2] SCAN, University Institute of Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, Switzerland

Abstract. MRI-based medical image analysis for brain tumor studies is gaining attention in recent times due to an increased need for efficient and objective evaluation of large amounts of data. While the pioneering approaches applying automated methods for analysis of brain tumor images date back almost two decades, the current methods are becoming more mature and coming closer to routine clinical application. This review article aims at providing a comprehensive overview by giving a brief introduction to brain tumors and imaging of brain tumors first. Then we review the state of the art in segmentation, registration and modeling related to tumor-bearing brain images with a focus on gliomas. The objective in segmentation is outlining the tumor including its sub-compartments and surrounding tissues, while the main challenge in registration and modeling is the handling of morphological changes caused by the tumor. The qualities of different approaches are discussed with a focus on methods that can be applied on standard clinical imaging protocols. Finally, a critical assessment of the current state is performed and future developments and trends are addressed, giving special attention to recent developments in radiological tumor assessment guidelines.

‡ To whom correspondence should be addressed (stefan.bauer@istb.unibe.ch)

1. Introduction

This review is intended to give an overview of the state of the art in MRI-based medical image analysis for brain tumor studies. It also provides a brief background on brain tumors in general and non-invasive imaging of brain tumors in order to give a comprehensive insight into the field.

1.1. Brain Tumors

With a prevalence of less than 1‰ in the western population, brain tumors are not very common, however they are among the most fatal cancers (DeAngelis (2001)). A recent study estimated the US incidence rate for primary tumors of the brain or nervous system to be around 25 per 100,000 adults with approximately one-third of the tumors being malignant and the rest either benign or borderline malignant (Kohler et al. (2011)). The word "tumor" is of latin origin and means swelling. Today, it is frequently associated with a neoplasm, which is caused by uncontrolled cell proliferation.

Brain tumors can be classified according to their origin or degree of aggressiveness. Primary brain tumors arise in the brain, while metastatic brain tumors frequently originate from other parts of the body. The most widely used grading scheme today has been introduced by the World Health Organization (WHO) (Kleihues et al. (1993)). It classifies brain tumors into grades I to IV with increasing aggressiveness.

Gliomas are the most frequent primary brain tumors in adults and account for 70% of adult malignant primary brain tumors. They arise from glial cells and they can be classified into the four WHO grades, where high-grade gliomas are of grade III or IV. Grade I and grade II tumors may be considered as semi-malignant tumors that carry a better prognosis, while grade III and grade IV tumors are malignant tumors that almost certainly lead to a subject's death. The WHO grade IV gliomas are also called glioblastoma multiforme (GBM). They are the most common malignant primary brain tumors in humans, exhibiting very rapid growth (Deimling (2009)). These tumors are infiltrative and spread mostly along the white matter fiber tracts (Giese et al. (2003)). They form abnormal vessels and exhibit a necrotic core. This and the surrounding edema leads to a mass-effect, which they exert on the healthy tissues of the brain. Regarding tumor angiogenesis, current understanding is based on the fact that low-grade gliomas are moderately vascularized tumors, whereas high-grade gliomas show prominent microvascular proliferations and areas of high vascular density (Plate and Risau (1995)). Average survival time for GBM is one year (Krex et al. (2007)). Treatment options for gliomas include surgery, radiation therapy or chemotherapy, including combinations of all these (DeAngelis (2001)).

Meningiomas are the most common extra-axial intracranial neoplasms. They account for 15-20% of intracranial neoplasms. They are non-glial neoplasms that originate from the arachnoid cap cells of the meninges. Treatment is usually performed by surgery and/or radiation therapy (Greenberg et al. (1999)).

Primary brain tumors must be differentiated from metastatic brain tumors that

originate most frequently from lung, breast, melanoma, renal and colon cancers and account for approximately 40% of intracranial neoplasms.

Due to clinical relevance and the amount of prior work in MRI-based medical image processing, the focus of this article will be on gliomas.

1.2. Imaging of Brain Tumors

The standard technique for brain tumor diagnosis is magnetic resonance imaging (MRI) (DeAngelis (2001); Wen et al. (2010)). MRI is a non-invasive technique, which provides good soft tissue contrast (Liang and Lauterbur (2000)) and is widely available in clinics. It is used in combination with other imaging modalities, such as computed tomography (CT), Positron Emission Tomography (PET) and Magnetic Resonance Spectroscopy (MRS) to provide the most exact information about tumor morphology and metabolism. Especially PET imaging can provide additional information (Chen (2007)), however so far MRI remains the accepted standard and therefore we will focus on MRI-based methods.

MRI makes it possible to produce markedly different types of tissue contrast by varying excitation and repetition times, which makes it a very versatile tool for imaging different structures of interest. Due to the nature and appearance of brain tumors, one MRI sequence is not sufficient to fully segment the tumor including all its subregions. In current clinical routine, different MRI sequences are employed for diagnosis and delineation of tumor compartments (Drevelegas and Papanikolaou (2011)). These sequences include T_1-weighted MRI (T_1), T_1-weighted MRI with contrast enhancement (T_{1c}), T_2-weighted MRI (T_2) and T_2-weighted MRI with fluid attenuated inversion recovery (T_{2FLAIR}), however acquisition parameters of these modalities are not standardized. Patients with gliomas are usually examined by the previously described MR imaging protocols according to the RANO guidelines with a slice thickness of $\leq$ 5mm without a gap between the slices (Pope and Hessel (2011)). For volumetry, high-resolution 3D volume images are performed, including at least contrast-enhanced T_1-weighted images with isotropic resolution. Figure 1 shows an axial slice of the four standard sequences for a glioma patient including manually drawn tumor regions.

T_1-weighting is the most commonly used sequence for structural analysis, it also allows for an easy annotation of the healthy tissues. In T_1-weighted contrast-enhanced images (Gadolinium-DTPA), the tumor borders appear brighter because the contrast agent accumulates there due to the disruption of the blood-brain barrier in the proliferative tumor region. In this sequence, the necrotic and the active tumor region can be distinguished easily. In T_2-weighted MRI, the edema region, which surrounds the tumor appears bright. T_{2FLAIR} (FLAIR) is a special sequence, which helps in separating the edema region from the cerebrospinal fluid (CSF) because the free water signal is suppressed. The radiological definition of the tumor margins in the clinical context are often manually determined by the radiologist on the T_2 and post-gadolinium T_1
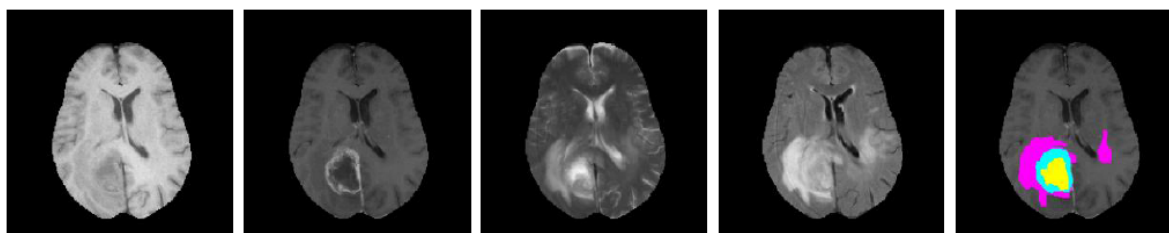
Figure 1. One axial slice of an MR image of a high-grade glioma patient. From left to right: T_1-weighted image, T_1-weighted image with contrast enhancement, T_2-weighted image, T_{2FLAIR}-weighted image and manual segmentation into necrotic (yellow), active (green), edema (pink) tumor compartments. Necrosis and active tumor region were segmented based on the T_1-weighted image with contrast enhancement, whereas the edema region was segmented based on the registered T_{2FLAIR}-weighted image.

images by thresholding boundaries between T_2 hyperintense / T_1 contrast-enhanced lesions and the surrounding healthy tissue to define the outer margins of a tumor. Clinical measurements of the tumor size traditionally incorporate either the product of the major and minor axis (2D measures) or of the three main axes of a tumor (3D measures).

Despite all the advantages that non-invasive imaging offers, it should be noted that a final diagnosis can only be made after biopsy and histology. It must also be emphasized that in gliomas, in contrast to metastatic brain tumors, the imageable component of the tumor is only part of the complete extent of the tumor. The imageable component and physiological information of the tumor depends on the tumor cell distribution and the image modality (Kelly et al. (1987), Tovi (1993)).

1.3. Image Analysis for Brain Tumors

This review is focused on image analysis for brain tumor studies. Image analysis deals with automatic or semi-automatic methods to help interpret the acquired images. Due to the large amount of data, that is currently being generated in the clinics, it is not possible to manually annotate and segment the data in a reasonable time. Usually, the domain of medical image analysis is divided into segmentation and registration, as well as into several further areas like enhancement, visualization, quantification and modeling (Bankman (2008)). For brain tumor studies, segmentation, registration and modeling appear to be the most important and the most challenging, so the remainder of this article focuses on these aspects.

Image segmentation (Pham et al. (2000)) aims at partitioning an image into several segments. These segments can be chosen according to structures of interest, tissue types, functional areas, etc. Balafar et al. (2010) presented a recent review targeted at brain segmentation specifically. Image registration (Maintz and Viergever (1998)) aims at aligning two different images in a common reference space. This is especially important for aligning images taken at different points in time during longitudinal studies or for aligning images from different modalities taken from the same patient, but also for brain

morphometry. Klein et al. (2009) compared non-linear algorithms for brain registration specifically. Finally, image-based modeling and simulation (Neal and Kerckhoffs (2010)) can help in disease understanding, treatment planning and decision making.

The primary use of MRI-based medical image analysis for brain tumor studies is in diagnosis, patient monitoring and treatment planning, but it could also be useful in clinical trials. Segmentation is crucial for monitoring of tumor growth or shrinkage in patients during therapy, for tumor volume measurements and it also plays an important role in surgical planning or radiotherapy planning, where not only the tumor has to be outlined, but also surrounding healthy structures are of interest. In current clinical practice, segmentation is usually still done manually, which is time-consuming and tedious for the radiologists and is also of limited use for an objective quantitative analysis. Concrete application areas for registration include multi-modal and longitudinal alignment of intra-patient images, but also aligning pre-, intra- and post-operative images as well as deformable registration for atlas-based segmentation or position mapping of gliomas for statistical analysis.

Although high-grade gliomas are not the most common cancers, they are among the most deadly. Novel therapies have improved patients' prognosis and novel clinical trial designs have gained in importance to further investigate the response of gliomas to different treatment regimens. A most accurate description of changes in tumor size and physiology is therefore mandatory. For many years now, treatment response of gliomas has been evaluated by MR and CT imaging; and a set of guidelines, usually referred to as the Macdonald criteria (Macdonald et al. (1990)) is available to monitor the assessment of tumor response after treatment. These criteria include changes in tumor size and new or increasing enhancement. However, this approach carries some important limitations (Clarke and Chang (2009)). Although the Macdonald criteria represented an important step in neuro-oncology research, it is now clear that evaluation of gadolinium enhancement alone is not adequate to characterize tumor growth or response. The RANO (Response Assessment in Neuro-Oncology) working group has defined novel criteria for tumor progression dependent on measurable or non-measurable lesions in at least two directions (Wen et al. (2010)). Automatic 3-D volumetric assessment of gliomas would offer the next important step forward to better understand tumor dynamics and response to treatment.

A previous review of the field was done by Angelini et al. (2007). In the meantime however, there has been significant progress, with the most important methods having been developed after 2007, mostly due to the rapid advancements in machine learning (Wang and Summers (2012)). Therefore, this review article aims at providing an update on the state of the art in image analysis for brain tumor studies.

The methods presented here have been collected mostly from Pubmed and Google scholar. We included more than 150 journal papers and conference papers from the most important conferences in the field (MedIA, IEEE TMI, NeuroImage, MRI, MRM, JMRI, PMB; MICCAI, IPMI, IEEE ISBI; and others). For the time before 2007, we included only the historically most relevant contributions, for the rest we refer to Angelini et al.

(2007).

The remainder of this review article is structured as follows: in the methods & results section (section 2), we present the different approaches and briefly summarize their results. The section is subdivided into methods for evaluation of accuracy and validation (2.1), segmentation (2.2), registration (2.3) and integrated approaches (2.4). Within each subsection we tried to group related methods together. Finally, in the discussion & outlook section (section 3) we review the state of the art in general, compare it to the clinical requirements and conclude about the applicability of the current methods, before we give an outlook on possible future directions and trends.

2. Methods & Results

2.1. Evaluation and Validation Methods

Validation and comparison against the state of the art is crucial for any newly developed method. Therefore, before diving into the presentation of the different approaches we would like to briefly cover the possibilities and challenges for evaluating and validating methods in brain tumor image analysis.

It would be optimal to compare any method against the real case. However, this is a big challenge in this field, if not impossible. In lack of a well-accepted ground-truth, the current gold standard for evaluation is to compare with manual segmentations by an expert. However, this is an extremely time-consuming and tedious task, additionally it is not objective. Mazzara et al. (2004) reported intra-rater volume variabilities of 20%±15% and inter-rater volume variabilities of 28%±12% for manual segmentations of brain tumor images. Weltens et al. (2001) found even larger values for inter-observer variability. One way to overcome this problem would be to use an algorithm that combines several expert segmentations in an optimal way like STAPLE (Warfield et al. (2004)). This method was applied for the evaluation of brain tumor segmentations by Archip et al. (2007b), however in most cases there are not enough expert segmentations available for using that method. Recently, Xu et al. (2012) suggested to use web-based collaborative manual tumor labeling by a large number of briefly trained non-experts to address this problem and they reported encouraging results. A more subjective, but occasionally used way is to employ a semi-automatic segmentation with a different well-accepted method as an intermediate result (e.g. Kikinis and Pieper (2011); Gao et al. (2012); Egger et al. (2013)), which is manually corrected by a human expert where necessary. This approach still lacks in objectivity, but it alleviates the burden from the clinician to spend a large amount of time for segmentation.

Another possibility for a first sanity check is to assess results on a synthetic dataset including ground-truth. Although synthetic data lacks important characteristics of real images, it has been used by many groups for initially assessing both segmentation and registration methods on healthy datasets with the BrainWeb phantoms (Cocosco et al. (1997)). For brain tumor studies, Prastawa et al. (2009) made a similar attempt to

provide simulated multi-sequence tumor image data including an objective ground truth, which was also based on the BrainWeb phantom and combined with a well-defined and accepted tumor growth model.

The most common way to quantitatively evaluate segmentation or registration results is to calculate the overlap with the ground truth. Usually, Dice similarity coefficient (DSC) or Jaccard coefficient (JC) are used (Crum et al. (2006)). They can range from 0 to 1 with 0 indicating no overlap and 1 indicating perfect overlap. Other possibilities for registration evaluation include manual landmark definition and calculation of landmark errors or surface distances. For a more qualitative assessment, checkerboard images can be shown or the outline of a structure can be analyzed visually. Zou et al. (2004) compared the three different validation metrics: area under the receiver operating characteristics (ROC) curve, mutual information (MI) and Dice similarity coefficient (DSC) for probabilistic brain tumor segmentation. They concluded that for overall classification accuracy the area under the ROC curve should be used, when interested in sensitivity to changes in tumor size MI is the metric of choice and for spatial alignment evaluation the Dice coefficient is best.

In the lack of a brain tumor database with ground-truth segmentations, that is available to a broad community of clinicians and researchers, so far most authors validated their algorithms on a limited number of cases from their own data. This makes it difficult to compare the performance of different methods against each other in an unbiased way. Therefore, and due to the different metrics used, the accuracy and speed of the individual methods, which have been collected from the respective publications, can not be directly compared with each other. Until recently, the only data available, which could serve such a purpose of general comparison, was the synthetic data of Prastawa et al. (2009), but so far only few groups tested their methods on these images. For the related topic of MS lesion segmentation from brain MRI (García-Lorenzo et al. (2013)), an open database has been available for some time (Styner et al. (2008)). Such an open image database has long been missing for an objective comparison of brain tumor segmentation algorithms, however an effort in this direction has finally been undertaken by the BraTS challenge at MICCAI 2012 §.

2.2. Segmentation of Brain Tumor Images

Segmentation of tumor-bearing brain images is a challenging task for several reasons. Firstly, high-grade gliomas usually exhibit unclear and irregular boundaries with discontinuities. Contrast uptake and image acquisition time after contrast injection can vary, which changes tumor appearance significantly and it is debatable if and how the non-imageable component of the tumor should be handled by segmentation algorithms. Additionally, tumor subregions can only be separated when several modalities are combined, which requires accurate registration as a pre-processing step. Finally, clinical datasets are usually acquired in a highly anisotropic way, leading to a much higher intra-
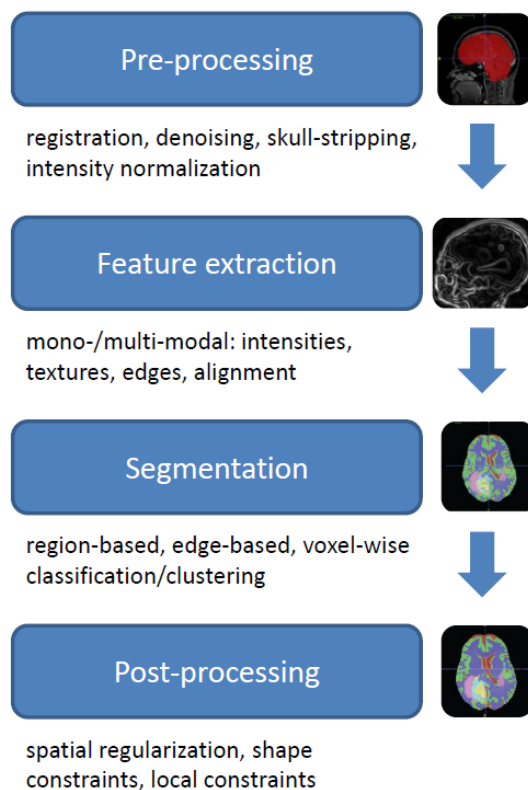
§ http://www2.imm.dtu.dk/projects/BRATS2012/

Figure 2. Illustration of the main blocks used for building up the segmentation pipeline of most algorithms included in this review.

slice resolution than inter-slice resolution. This causes problems with partial-volume effects for segmentation, but also registration and resampling of the different modalities in a common space of reference. A diagram illustrating the major steps during the segmentation pipeline of most algorithms is shown in figure 2. In the following, the different algorithms are discussed according to the major blocks of this pipeline.

2.2.1. Image Pre-processing Most algorithms rely on some kind of preprocessing for image preparation and image enhancement. Image denoising is a standard pre-processing task for MRI. Many approaches have been suggested, the most popular ones being based on anisotropic diffusion filtering (Weickert (1998)). Diaz et al. analyzed different denoising algorithms for the specific task of brain tumor segmentation (Diaz et al. (2011)). They concluded that, although image noise was reduced, many algorithms introduced artefacts which had a negative effect on segmentation.

Intensity normalization (Nyúl and Udupa (1999)) is a very critical task for MRI, especially when classification methods are used for segmentation. However, for tumor-bearing images, this is more challenging than for healthy images due to the confounding effects caused by the differences in tumor appearance. Ekin (2011) developed an intensity-normalization technique for MRI, which they claimed was robust

to pathologies. Most approaches employ a bias-field correction (Vovk et al. (2007)) before the segmentation is applied, in order to compensate for the effect of magnetic field inhomogeneities during image acquisition.

When operating on multi-modal images, pre-processing always includes registration of all modalities in a common space of reference. In most cases, this is performed using a linear transformation model with mutual information similarity metric (Mang et al. (2008)) and resampling in order to ensure voxel-to-voxel correspondence across all modalities.

Most approaches for brain tumor segmentation rely on a skull-stripping step (Fennema-Notestine et al. (2006)) before the actual segmentation is performed. Speier et al. (2011) recently presented a skull-stripping method dedicated to glioma images, which was able to handle images containing resection cavities.

2.2.2. Image Features for Segmentation Algorithms   The features used for segmentation of brain tumors largely depend on the type of tumor and its grade because different tumor types and grades can vary a lot in appearance (e.g. contrast uptake, shape, regularity, location, etc.). Additionally, feature selection will also depend on the sub-compartment of the tumor, which is to be segmented.

The most common feature used for brain tumor segmentation are the image intensities. This is based on the assumption that different tissues have different graylevels. Another type of features frequently used are local image textures because it has been observed that different tumor areas exhibit different textural patterns (Kassner and Thornhill (2010)). Textures can be computed according to different strategies (Tuceryan and Jain (1998)). Alignment-based features make use of spatial prior knowledge, which is often encoded by registration of a standard atlas to the patient image or by making use of symmetries between left and right brain hemisphere. Intensity gradients or edge-based features can be used for evolving a contour towards the tumor border. Recently, context-aware features modeling mid- or long-range spatial contexts similar to Lepetit et al. (2005) and Shotton et al. (2011) are becoming more popular.

Depending on the data available, all these features can either be computed from one single modality or from multi-modal images. Researchers have also combined different features from different modalities in order to improve their segmentation results.

For the task of brain tumor image analysis, Schmidt et al. (2005) explored different features for voxel-based classification and concluded that combining textural and alignment-based features allowed for substantial performance increases. Ahmed et al. (2011) investigated the efficacy of different image features and feature fusion strategies for pediatric brain tumor images. They argued that in multi-modal images, texture features had advantages over intensity or shape features.

2.2.3. Segmentation Algorithms Based on the previous section, we categorize segmentation algorithms according to the features they use. Therefore, we distinguish region- or edge-based methods, which mostly rely on deformable models, and

Table 1. Segmentation methods for brain tumor images using deformable methods or region-based methods. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision. SA means semi-automatic, FA fully-automatic. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Modalities	Method	Accuracy	Speed	Dim	S	Type
Edge-based methods							
Wang et al. (2009)	T_1	Fluid vector flow	0.6 (Tanimoto)	5s	2D	SA	
Sachdeva et al. (2012)	T_1, T_{1c}, T_2	Content-based active contour model	0.72-0.98		2.5D	SA	G, M
Region-based methods							
Ho et al. (2002)	T_1, T_{1c}	Level-set evolution with region competition	0.85-0.93 (Jaccard)		3D	FA	G, M
Rexilius et al. (2007)	T_{1c}, T_2, T_{2flair}	Multispectral histogram model adaptation for region-growing	0.73 (Jaccard)	10min	3D	SA	G
Harati et al. (2011)	T_{1c}	Fuzzy connectedness	>0.9 (similarity index)	2.5min		FA	G, M

classification or clustering methods, which make use of voxel-wise intensity and texture features. Many methods employ additional constraints for regularization, this is discussed in a separate section. Another group of methods is based on atlas-based segmentation. Atlas-based methods will be described later in sections 2.3 and 2.4 because the segmentation relies on registration methods.

Region- or Edge-based Methods

Deformable models make use of regional characteristics or edge detection in the images (McInerney and Terzopoulos (1996)). In most cases, a level-set is evolved towards the tumor boundary by searching for the largest gradient in the image or by employing region properties. Wang et al. (2009) employed a fluid vector flow model to evolve a contour towards the tumor boundary edge in T_1-weighted images. Sachdeva et al. (2012) made use of content-based intensity and texture patterns to evolve an active contour towards the tumor boundary in different MRI modalities.

Ho et al. (2002) derived a tumor probability map from the difference image between T_1 and T_{1c}, which formed the basis for the evolution of a region-competition level-set algorithm applied. Rexilius et al. (2007) initialized a region-growing algorithm with a tumor map, which was obtained from a multi-spectral histogram model adaptation. Harati et al. (2011) suggested a fuzzy-connectedness algorithm, which made use of region properties but had no spatial constraints. Seeding was automatically performed on T_1-weighted images with a so-called tumor detector matrix. Characteristics of each method are summarized in table 1.

Classification and Clustering

In fact, most of the segmentation algorithms proposed so far, are based on classification or clustering approaches. This is mostly owed to the fact that with these methods, multimodal datasets can be handled easily because they can operate on any chosen feature vector. In most cases, the features on which these algorithms

operate include voxel-wise intensities and frequently also local textures. The general idea is to decide for every single voxel individually, to which class it belongs based on its feature vector. Classification requires training data to learn a classification model, based on which new instances can be labeled. Clustering, on the other hand, works in an unsupervised way and groups data based on certain similarity criteria (Wang and Summers (2012)).

Clustering was introduced into the brain tumor segmentation community by Schad et al. (1993) who analyzed texture patterns of different tissues. Phillips et al. (1995) employed fuzzy c-means clustering and Vaidyanathan (1995) compared this to kNN-clustering for tumor volume determination during therapy on multispectral 2D image slices. Clark et al. (1998), from the same group, further developed this approach to incorporate knowledge-based techniques. Later, Fletcher-Heath et al. (2001) also combined fuzzy clustering with knowledge-based techniques for brain tumor segmentation from multi-sequence MRI.

Lately, there have been tremendous advances in developing more powerful discriminative classification methods and these new methods have also found their way into the field of medical image analysis. Cai et al. (2007) and later Verma et al. (2008) from the same group, used a high number of MRI modalities (DTI channels in addition to the conventional ones) to create voxel-wise intensity-based feature vectors, which they classified with support vector machines (SVM) (Schoelkopf and Smola (2002)). They were able to not only segment the healthy tissues, but also segment sub-compartments of healthy and tumor regions. Ruan et al. (2007) used a very similar approach based on SVMs, but with a lower number of modalities and they only segmented one tumor region. Later, they claimed that feature selection using kernel class separability, could slightly improve the results compared to their previous approach (Ruan et al. (2011)). Jensen and Schmainda (2009) explored different neural networks to detect brain tumor invasion from multi-parametric MRI (structural, diffusion and perfusion images). Kanaly et al. (2011) chose a simpler approach by thresholding the voxels of the difference image of pre- and post-contrast T_1-weighted MRI after intensity normalization. Zikic et al. (2012) lately proposed a computationally efficient approach based on decision forest classification (Criminisi et al. (2011)) with context-aware features and an additional generative model as an input, which is able to identify tumor sub-compartments from multi-modal images (see figure 3). They claimed that context-aware features eliminate the need for a post-processing step imposing smoothness constraints by spatial regularization. Geremia et al. (2012) had the idea to generate synthetic tumor images, which can be used to train a discriminative regression forest algorithm using different groups of features. They argued that this approach allowed them not only to segment patient images, but also to estimate latent tumor cell densities. Characteristics of each method are summarized in table 2.

Classification and Clustering with Additional Constraints
It has been remarked that simple voxel-wise classification or clustering methods do
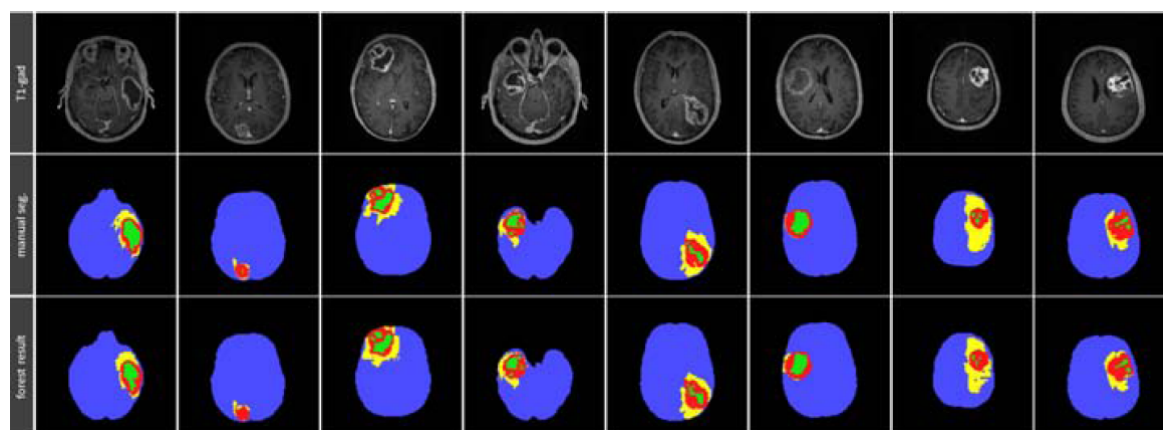
Figure 3. Segmentation results generated from multi-sequence 3D MR images, shown on T_{1c}-weighted axial slices of different patients. The second row shows the manually defined ground truth, the last row the results of a fully automatic algorithm, which used T_1, T_{1c}, T_2, T_{2flair} and DTI MRI sequences as an input. The pathologic tissues are separated into active, necrotic and edema compartments. Example from Zikic et al. (2012).

Table 2. Segmentation methods for brain tumor images using classification or clustering, grouped by different approaches. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision. SA means semi-automatic, FA fully-automatic. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Modalities	Method	Accuracy	Speed	Dim	S	Type
Classification or clustering methods							
Schad et al. (1993)	T_1, T_2	Texture analysis and clustering			2D		HGG, Met
Phillips et al. (1995)	T_1, T_2, PD	Clustering					HGG
Vaidyanathan (1995)	T_1, T_2, PD	Clustering		2-5min	2D		HGG, M
Clark et al. (1998)	T_1, T_2, PD	Knowledge-based techniques	0.69-0.99 (% match)		2D	FA	HGG
Fletcher-Heath et al. (2001)	T_1, T_2, PD	Fuzzy clustering + knowledge-based techniques	0.53-0.91 (% match)		2.5D	FA	
Cai et al. (2007)	T_1, T_{1c}, T_2, T_{2flair}, DTI	Probabilistic segmentation based on multi-modality MRI	0.73-0.98 (classification rate)		3D	FA	HGG
Ruan et al. (2007)	T_1, T_2, T_{2flair}, PD	SVM Classification	0.99 (TP)		3D	FA	
Verma et al. (2008)	T_1, T_{1c}, T_2, T_{2flair}, DTI	Multiparametric tissue classification	0.34-0.93 (classification rate)		3D	FA	HGG
Jensen and Schmainda (2009)	T_1, T_{1c}, T_2, T_{2flair}, DWI, DTI, DSC	Different classifiers on multiparametric MRI	0.78-0.86 (overlap)		3D	SA	HGG, M
Ruan et al. (2011)	T_2, T_{2flair}, PD	Image fusion for brain tumor follow-up		5min	3D	FA	
Kanaly et al. (2011)	T_1, T_{1c}	Difference image for volumetric tumor assessment			3D	SA	
Zikic et al. (2012)	T_1, T_{1c}, T_2, T_{2flair}, DTI	Decision forests for tissue-specific segmentation	0.7-0.9 (Dice)	2-3min	3D	FA	HGG
Geremia et al. (2012)	T_1, T_{1c}, T_2, T_{2flair}	Tumor cell density estimation for discriminative segmentation	0.55-0.83 (Dice)		3D	FA	G

not make use of the complete information contained in an image. Therefore, there have been numerous attempts to use additional information to further improve the segmentation result by adding additional constraints, which can be based on some form of neighborhood regularization or on imposing shape or localization constraints for the tumor. Neighborhood constraints are often imposed using a random field regularization method, while shape constraints are mostly handled by deformable models. Atlases can be used to restrict the tumor location and also for generative classification models.

Atlases have been used by a number of methods to incorporate spatial prior knowledge into the classification task. Kaus et al. (2001) suggested to use a brain atlas for guiding a kNN-classifier in their adaptive template-moderated classification algorithm. Other researchers used atlases not only to impose spatial constraints, but also to provide probabilistic information about the tissue model. Moon et al. (2002) and also Prastawa et al. (2003) employed a probabilistic tissue model and used an expectation maximization (EM) method (Dempster (1977)), which segmented tumor images by modifying an atlas with patient-specific information about tumor location from different MRI modalities. Later, Prastawa et al. (2004) extended this to work on mono-modal images. Menze et al. (2010) combined a healthy atlas with a latent tumor atlas to segment brain tumors from multi-sequence images using a generative probabilistic model and spatial regularization. Weizman et al. (2012) used localization based on an atlas for their multimodal segmentation of optic pathway gliomas, which performed classification with a probabilistic tissue model.

Deformable models (McInerney and Terzopoulos (1996)) can be used to impose shape constraints after previous tissue classification. Cobzas et al. (2007) and Cobzas and Schmidt (2009) combined tissue classification using a high-dimensional feature set with level-set evolution. Popuri et al. (2012) added a Dirichlet prior to the previous approach to better disambiguate tumor from healthy tissues. Khotanlou et al. (2009) combined fuzzy clustering and brain symmetry features with level-set evolution. Hamamci et al. (2011) introduced the tumor-cut algorithm, which combines tumor segmentation using cellular automata with a level-set evolving on the tumor probability map to impose spatial smoothness.

Neighborhood relationships are often used for spatial regularization after initial voxel-wise classification. In most cases this is performed with a random field method, either Markov Random Fields (MRF) (Kindermann and Snell (1980)) or Conditional Random Fields (CRF) (Lafferty et al. (2001); Kumar and Hebert (2006)). In the last years, algorithms based on some form of graph-cuts for segmentation (Boykov and Funka-Lea (2006)) have become increasingly popular. These algorithms make use of region-based properties by formulating neighborhood relationships as an energy minimization problem, which is solved by a graph-cut optimization method and the result is interpreted as a segmentation. In many cases, these methods require user-provided seeds, this is why the approaches are often semi-automatic. Lee et al. (2008) used conditional random fields for spatial regularization after previous voxel-wise SVM classification from multiple modalities. Wels et al. (2008) integrated a

probabilistic boosting tree classifier with a graph-cuts approach for considering the spatial relationships in pediatric brain tumor images. Birkbeck et al. (2009) proposed an interactive semi-automatic approach, which was able to include online user corrections and they argued that this method could not only reduce operator time, but also increased repeatability compared to manual segmentation. Hamamci et al. (2010) proposed a variation of the original graph-cut method using cellular automata for solving the shortest-path problem iteratively. Nie et al. (2009) used an EM algorithm, which was coupled with a spatial accuracy-weighted hidden Markov Random Field. This allowed them to consider anisotropic resolutions of different modalities. Zhu et al. (2012) also combined EM segmentation with MRF regularization, but added a post-processing pipeline including thresholding and morphological operations. Bauer et al. (2011a) claimed to be able to segment tumor and healthy tissues including sub-compartments based on SVM classification with integrated hierarchical CRF regularization. They also made use of prior knowledge about tissue adjacency probabilities. Hsieh et al. (2011) chose a simple region-growing and knowledge-based post-processing after fuzzy tissue classification to impose spatial coherence. Corso et al. (2007) took a different approach and suggested an extended graph-shifts algorithm, which performed energy minimization on a dynamic hierarchical representation of the image. In a later publication, they employed a multi-level approach for fusing Bayesian tissue classification with affinity assignments to perform segmentation by weighted aggregation (Corso et al. (2008)). Characteristics of each method are summarized in table 3.

2.2.4. Tumor Detection Algorithms In contrast to segmentation algorithms, detection algorithms only try to decide if tumor is present and output the approximate tumor location instead of providing a complete segmentation. Saha et al. (2011) located the tumor and drew a bounding box, instead of segmenting it. This could help in quickly analyzing large amounts of data. It was done with the help of a fast unsupervised change detection method searching for dissimilar regions across the symmetry line of the brain using Bhattacharya coefficient score. Ambrosini et al. (2010) employed a template-matching method to detect metastases on conventional MRI for screening purposes. Farjam et al. (2012) chose a similar approach, but improved on the spherical template generation process by considering variations in tumor size, lesion shape and intensities to achieve more accurate detection rates. Although these approaches do not perform a segmentation, the tumor detection could be used for initializing a segmentation method.

2.2.5. Summary and Conclusion - Segmentation Table 1, 2 and 3 present an overview of all the segmentation algorithms discussed and compare their most important characteristics and results. Most algorithms that have been suggested so far, rely on multi-modal MR images for tumor segmentation. The most promising approaches perform fully automatic segmentation based on voxel-wise classification in combination with spatial regularization to take the neighborhood information into account. They have the advantage of being very flexible and provide robust results within a

Table 3. Segmentation methods for brain tumor images using classification or clustering with additional constraints, grouped by different approaches. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision. SA means semi-automatic, FA fully-automatic. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Modalities	Method	Accuracy	Speed	Dim	S	Type
Classification or clustering with global constraints based on atlases							
Kaus et al. (2001)	T_1	Adaptive template-moderated classification	>95% (accuracy)	5-10min	3D	FA	LGG, M
Moon et al. (2002)	T_1, T_{1c}, T_2	Statistical classification with geometric prior	>0.9 (overlap ratio)		3D	FA	G, M
Prastawa et al. (2003)	T_1, T_{1c}, T_2	Subject-specific modification of atlas priors	0.49-0.71 (Jaccard)		3D	FA	
Prastawa et al. (2004)	T_2	Outlier detection	0.59-0.89 (Jaccard)	1.5h	3D	FA	G, M
Menze et al. (2010)	T_1, T_{1c}, T_2, T_{2flair}	Generative model for multi-modal segmentation	0.4-0.7 (Dice)		3D	FA	G
Weizman et al. (2012)	T_1, T_2, T_{2flair}	Classification and follow-up for optic pathway gliomas	0.69 (Dice)	20min	3D	FA	G
Classification or clustering with constraints based on deformable models							
Cobzas et al. (2007)	T_1, T_{1c}, T_2	Variational segmentation with high-dimensional feature set	0.16-0.76 (Jaccard)		3D	FA	G
Khotanlou et al. (2009)	T_1	Fuzzy classification and spatially constrained deformable models	0.92 (similarity index)	4.5min	3D	FA	HGG, M
Cobzas and Schmidt (2009)	T_{1c}, T_{2flair}	Level-set with embedded CRF	0.5-0.75 (Jaccard)		3D	FA	G
Popuri et al. (2012)	T_1, T_{1c}, T_2	Variational segmentation on a clustered feature set	0.58 (Jaccard)		3D	FA	G
Hamamci et al. (2011)	T_{1c}	Tumor-Cut	0.8-0.89 (Dice)	1s-16min	3D	SA	G,M, Met
Classification or clustering with local constraints based on neighborhood regularization							
Corso et al. (2007)	T_1, T_{1c}, T_2, T_{2flair}	Extended graph-shifts algorithm	0.87 (Jaccard)	1min	3D	FA	HGG
Corso et al. (2008)	T_1, T_{1c}, T_2, T_{2flair}	Segmentation by weighted aggregation	0.62-0.69 (Jaccard)	<1min	3D	FA	HGG
Lee et al. (2008)	T_1, T_{1c}, T_2	Pseudo conditional random fields	0.84-0.9 (Jaccard)		2D	FA	G
Wels et al. (2008)	T_1, T_{1c}, T_2	Discriminative model-constrained graph-cuts	0.78 (Jaccard)	1-2min	3D	FA	G
Birkbeck et al. (2009)	T_{1c}, T_{2flair}	Interactive graph-cuts		2min	2.5D	SA	
Nie et al. (2009)	T_1, T_2, T_{2flair}	Spatial accuracy-weighted hierarchical MRF	0.72-0.76 (Jaccard)	20-25min	3D	FA	G
Hamamci et al. (2010)	$T_{1contrast}$	Cellular automata segmentation	0.74-0.87 (Dice)		3D	SA	G,M, Met
Hsieh et al. (2011)	T_1, T_2	Fuzzy clustering plus region-growing	0.73 (% match)		3D	FA	M
Bauer et al. (2011a)	T_1, T_{1c}, T_2, T_{2flair}	Hierarchical SVM+CRF	0.77-0.84 (Dice)	<2min	3D	FA	G
Zhu et al. (2012)	T_{1c}, T_2	EM + MRF + Post-processing	0.25-0.81 (Jaccard)	4min	3D	SA	HGG

reasonable computation time. Imposing shape or localization constraints after previous classification is an effective approach too, but less flexible in handling different data and different kinds of tumors. However, fully automatic methods are subject to differences in MRI acquisitions, which can cause difficulties. Fewer examples rely on semi-automatic methods that incorporate user interaction. The interactive methods often rely on graph-cuts or deformable models with a user-defined initialization, which makes them very flexible, but eliminates the advantage of being completely objective. It is not possible to draw general conclusions about the accuracy of different methods, because so far they have all been evaluated on different datasets.

While preparing this review, it became apparent that many papers focus on segmentation algorithms and not on the features extracted from the image. Features might be more important especially when considering the variance in appearance of different tumor types and grades. In the future, it might be worthwile to take a closer
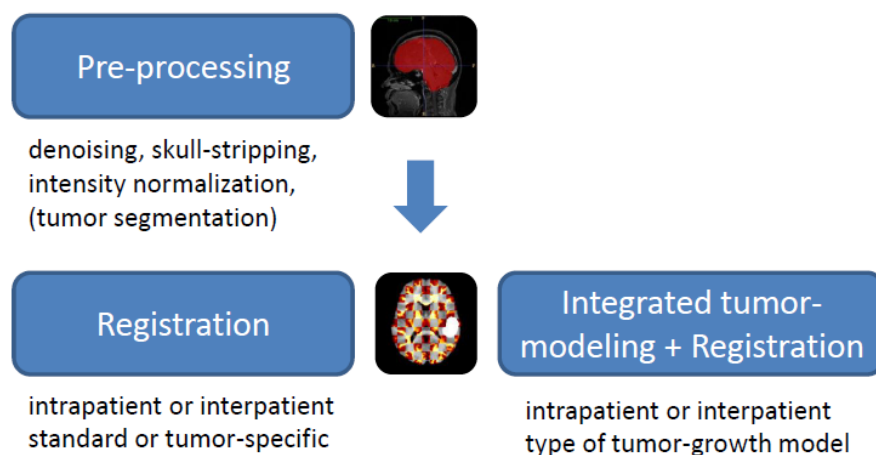
Figure 4. Illustration of the main blocks, which are used for building up the registration pipeline of most algorithms discussed in this article. We distinguish between algorithms, which are only based on registration and integrated algorithms combining tumor-growth modeling with registration. Most algorithms perform a linear registration for initial alignment before a non-rigid transformation is applied. The non-rigid transformation can be standard or tumor-specific.

look at relevant and meaningful features, and it would be interesting to explore how new features can be designed to obtain better results. This could be added to improvements in pre-processing for image standardization and in constraints for increased robustness.

2.3. Registration of Brain Tumor Images

Image registration of tumor-bearing brain images mainly focuses on two different objectives: intrapatient multi-modal or longitudinal registration (alignment) of images on the one hand, and interpatient spatial normalization of brain tumor images to a brain atlas on the other hand. The latter is commonly used for atlas-based segmentation purposes (Cabezas et al. (2011)) or for statistical analysis and functional/structural mapping of brain tumor images. It can also be applied for constructing statistical tumor atlases from multiple patients (Davatzikos et al. (2011)). A diagram illustrating the major blocks of most registration pipelines is shown in figure 4. The pre-processing steps are very similar to those discussed in section 2.2.1 and are therefore not repeated here. We decided to separate registration methods and methods integrating tumor-modeling with registration. Again, both can be sub-grouped into intra- or interpatient methods, which can use either standard or tumor-specific registration algorithms. Usually, an initial rigid alignment is followed by a non-rigid transformation step (Sotiras et al. (2012)). The feature metric, which is used by most of the discussed approaches is intensity-based registration thanks to its general applicability.

2.3.1. Standard Intrapatient Registration In the case of intrapatient multi-modal or longitudinal registration of images, the aim is to geometrically transform the images into a common reference space for simultaneous analysis of the different modalities, or for tumor growth monitoring over time. The problem of simply aligning the images without specifically considering any growth or deformation effects by the tumor was extensively studied by Mang et al. (2008), who concluded that in general, affine registration with a mutual information metric performed best. Another problem, which frequently occurs in registration of tumor-bearing brain images is the anisotropic voxel spacing. The individual modalities are usually acquired with different anisotropic resolutions, sometimes even in different orientations. The resulting interpolation problem has been addressed for the general case in Thévenaz et al. (2000). Nevertheless, in practice the inherent partial-volume effects still pose significant challenges for resampling all images in a common reference space. Characteristics of the method are summarized in table 4.

2.3.2. Advanced longitudinal Intrapatient Registration to handle Resections The problem of intrapatient registration is much more difficult when effects of surgery are considered because then, non-linear effects have to be included and in most cases it is not sufficient to apply standard registration algorithms. One application area for this is registration of pre-operative images to intra-operative images or post-resection tumor images. Clatz et al. (2005a) tackled the problem of non-rigidly registering intra-operative MR images to pre-operative scans. They were able to handle the mechanical brain deformation during surgery with a patient-specific finite element method (FEM) and a non-rigid block-matching method. Strict constraints on the computation time could be handled by a parallelized implementation and pre-computation of a large part of the processing pipeline. Archip et al. (2007a) extended this approach, but they registered additional pre-operatively acquired MRI modalities, so that this information could be exploited during surgery. They achieved near real-time computational speed and were able to provide enhanced visualization during neurosurgery. Wittek et al. (2007) presented a patient-specific bio-mechanical model of brain deformation and its application to image registration. In their method, they were able to make use of patient-specific non-linear bio-mechanical tissue properties for integrating FEM analysis into non-rigid registration of pre- and intra-operative images. Chitphakdithai and Duncan (2010) relied on non-rigid registration for considering changes after brain tumor treatment. They formulated registration as a maximum a posteriori (MAP) problem and solved it in an EM framework. This way, the probability term of the transformation could be seen as a similarity metric, for which an indicator map could model different correspondence assumptions for different tissue classes. Characteristics of each method are summarized in table 4.

2.3.3. Advanced longitudinal Intrapatient Registration for Tumor Growth Monitoring Another area where advanced intrapatient registration strategies are necessary is tumor growth monitoring from longitudinal image acquisitions. Konukoglu et al.

Table 4. Overview of registration methods for brain tumor images, grouped by different approaches for intrapatient registration. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision, Transf. for the transformation model. SA means semi-automatic, FA fully-automatic, L linear transformation model and NL non-linear transformation model. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Method	Accuracy	Speed	Dim	S	Transf.	Type
Standard intrapatient registration							
Mang et al. (2008)	Consistency of parametric registration	sub-voxel accuracy		3D	FA	L	G
Advanced longitudinal intrapatient registration to handle resections							
Clatz et al. (2005a)	Non-rigid registration to capture brain shift	<2.5mm	<35s (cluster)	3D	FA	NL	
Archip et al. (2007a)	Pre- and intra-operative alignment of MRI	<3.6mm	179s (cluster)	3D		NL	G
Wittek et al. (2007)	Model of brain deformation for image registration	1-30mm	15min	3D		NL	
Chitphakdithai and Duncan (2010)	MAP registration estimation framework	1.3 mm (AE)		3D	FA	NL	
Advanced longitudinal intrapatient registration for tumor growth monitoring							
Konukoglu et al. (2008)	Tumor monitoring	<38% (error)		3D	SA	NL	M
Niethammer et al. (2011)	Geometric metamorphosis	0.975 (overlap)		3D	FA	NL	
Pohl et al. (2011)	Detecting change in slowly evolving brain tumors	<23% (vol. diff.)	<5min	3D	SA	L	M
Angelini et al. (2012)	Differential analysis for tumor growth quantification	9-23% (error)		2D/3D	FA	NL	LGG

(2008) suggested a registration method for monitoring slowly evolving meningiomas, which performed semi-automatic tumor segmentation, non-rigid registration and change detection consecutively. They argued that their volume-change measurements were less user-biased than manual measurements. Niethammer et al. (2011) suggested a metamorphosis model, which jointly estimated a deformation in space and a change in image appearance for smooth transformation. Changes in image appearance were modeled by a global geometric deformation, and local matching was based on an image composition model in a large displacement diffeomorphic metric mapping (LDDMM) framework. Pohl et al. (2011) proposed to detect changes in slowly evolving meningiomas by semi-automatic segmentation, subsequent registration and final analysis of changes in local intensity patterns. Differential MRI analysis for quantification of low-grade glioma growth was applied by Angelini et al. (2012). They used a standard affine registration, but implemented a non-linear midway-based histogram mapping, which allowed them to compare and compute intensity difference maps directly. Glioma growth was finally quantified with a statistical analysis framework of the mapped intensity distributions. Characteristics of each method are summarized in table 4.

2.3.4. Standard Interpatient Registration The case of interpatient registration of tumor-bearing brain images is even more challenging than the alignment of images from the same patient because intersubject variations have to be integrated with topological changes. Isambert et al. (2008) circumvented this problem by using a general method for registering an atlas of a standard brain to the patient image using a multi-affine block-

matching strategy. Organs at risk in a clinical radiotherapy context were delineated from the registered atlas. Later, Deeley et al. (2011) did an extensive comparison of manual and automatic segmentation methods for brain structures in presence of space-occupying lesions. As an automatic method, they used a combination of multi-affine and non-rigid registration with final level-set refinement. They claimed that the accuracy of automatic and manual segmentations were in a similar range when compared against the STAPLE ground-truth, but the manual segmentations were still slightly better. Characteristics of each method are summarized in table 5.

2.3.5. Interpatient Registration with Lesion Masking Approaches In a comparison study, Andersen et al. (2010) showed that non-rigid registration approaches which masked out the lesion area from the similarity metric, achieved superior results compared to conventional registration taking the whole image region into account. Variations of these masking methods, which gave a different weight to tumor regions have been suggested by a number of groups.

Brett et al. (2001) masked the cost function of a combined affine and non-rigid registration method. The lesion had to be manually pre-segmented, to be used as a mask for the similarity criterium during registration. They experimented with different thresholds for the cost function mask. Stefanescu et al. (2004) chose a similar approach. They created a confidence map for the similarity metric used during registration and assigned zero confidence to all voxels inside the pre-segmented tumor mask. Additionally, they allowed adaptive regularization in different tissue regions.

Dawant et al. (2002) introduced small tumor seeds in the atlas at the approximate patient tumor location. Subsequently, they performed non-rigid registration, which also deformed the seed in the atlas to match the shape of the pre-segmented tumor in the patient approximately. Commowick et al. (2005) used statistical measures of anatomical variability to guide the regularization during the registration process. Regions of low variability were regularized more strongly. Lamecker and Pennec (2010) took up the idea of using confidence weights. They added an inpainting step to a Demons registration algorithm, which modeled a zero confidence within the pathologic region. Li et al. (2011) chose a Riemannian embedding for their registration algorithm, which was able to handle topological changes. The method required an a priori estimation of topological changes as an input. Parisot et al. (2011) did not take the tumor voxels into account for deformable registration of a dataset of images from different patients. The registration output has been used for graph-based spatial position mapping of low-grade gliomas. Later, they proposed a graph-based joint segmentation and registration framework (Parisot et al. (2012)). In the tumor region, which was identified by an Adaboost classifier, the registration criterion was relaxed to produce a smooth solution of the segmentation. Characteristics of each method are summarized in table 5.

2.3.6. Interpatient Registration with Two Distinct Deformation Models Another idea for interpatient registration is to incorporate a lesion model directly into the registration
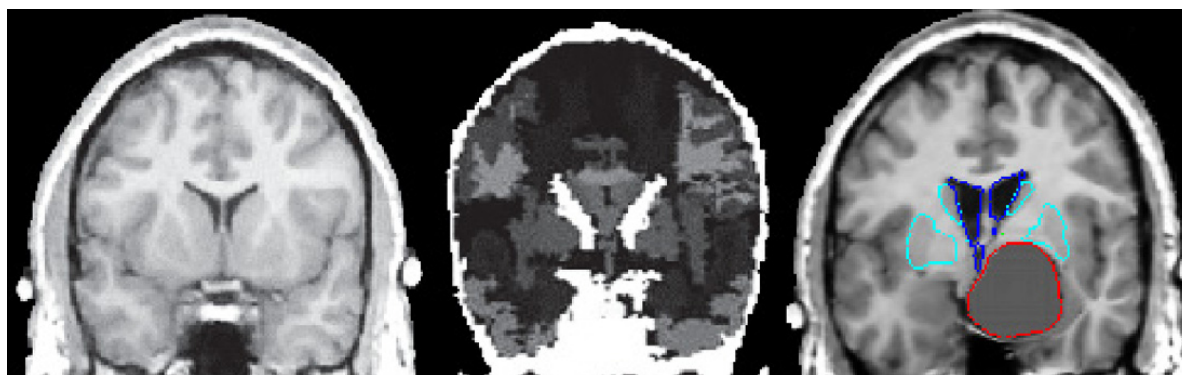
Figure 5. Registration based on two distinct deformation models. From left to right: atlas image, atlas label map and the segmented patient image (from Bach Cuadra et al. (2006)). The atlas is modified using a lesion-growth model to mimic the tumor-induced deformation and the segmented structures from the atlas are transferred to the patient image by non-rigid registration.

method. This allows for a decoupling of the deformations due to tumor growth and inter-subject variations, without having to model the complex tumor growth behavior explicitly. Bach Cuadra et al. (2004) suggested a model of lesion growth for atlas-based segmentation of tumor-bearing brain images. They incorporated a simplistic radial tumor growth model into a Demons-based non-rigid registration method. The lesion growth model modified a healthy atlas and adapted it to the tumor-bearing patient image, however the method was completely based on registration methods and did not include any kind of bio-mechanical tumor growth simulation. Later, Pollo et al. (2005) from the same group conducted more qualitative evaluation of the method. Finally, they further improved it by replacing the SSD-based Demons algorithm with an optical flow method based on mutual information (Bach Cuadra et al. (2006)). They reported better results for this method because the assumption of linear intensity correspondence between both images could be dropped (see figure 5 for results). Characteristics of each method are summarized in table 5.

2.3.7. Summary and Conclusion - Registration Most registration methods make use of a combination of rigid/affine and non-rigid registration with specific adaptations for tumor-bearing images. While non-rigid methods are obviously more accurate than standard rigid registration, they also carry additional risks of mis-registrations, especially when topological changes due to the tumor are involved. This risk can be reduced by tumor-specific adaptations, however these methods are in general more difficult to handle than standard registration approaches. It is not possible to draw a general conclusion about the accuracy and performance of different methods because different data has been used for validation, but we will briefly summarize the qualities of the different approaches.

The case of intra-subject alignment of longitudinal / multi-modal images can be easily solved by standard registration methods if tumor effects are neglected. The

Table 5. Overview of pure registration methods for brain tumor images, grouped by different approaches for interpatient registration. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision, Transf. for the transformation model. SA means semi-automatic, FA fully-automatic, L linear transformation model and NL non-linear transformation model. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Method	Accuracy	Speed	Dim	S	Transf.	Type
Standard interpatient registration							
Isambert et al. (2008)	Atlas-based segmentation in radio-therapy clinical context	0.3-0.85 (Dice)	7-8min (cluster)	3D	FA	NL	G, M
Deeley et al. (2011)	Comparison of manual and automatic segmentation methods	0.4-0.9 (Dice)		3D	FA	NL	HGG
Interpatient registration with lesion masking approaches							
Brett et al. (2001)	Cost function masking	0.5-2.6mm (RMS)		3D	SA	NL	
Dawant et al. (2002)	Seeded atlas deformation			3D	SA	NL	G, M
Stefanescu et al. (2004)	Confidence map for non-rigid registration		5-10min (cluster)	3D	FA	NL	
Commowick et al. (2005)	Statistical measures of anatomical variability for registration	0.84-0.94 (spec)	1h	3D	FA	NL	
Lamecker and Pennec (2010)	Demons and deformation inpainting		9min	3D	SA	NL	G
Li et al. (2011)	Riemannian embedding			3D	SA	NL	M
Parisot et al. (2011)	Graph-based spatial position mapping			3D	SA	NL	LGG
Parisot et al. (2012)	Joint tumor segmentation & dense deformable registration	0.8 (Dice)	6min	3D	FA	NL	LGG
Interpatient registration with two distinct deformation models							
Bach Cuadra et al. (2004)	Model of lesion growth for atlas-based segmentation	37.2mm^3 (MSE)		3D	SA	NL	M
Pollo et al. (2005)	Model of lesion growth for atlas-based segmentation			3D	SA	NL	M
Bach Cuadra et al. (2006)	Deformation field estimation for atlas-based segmentation using MLG and MI	0.89-1.0 (TP)		3D	SA	NL	M

problem is much more difficult if registration is used for longitudinal tumor growth monitoring or when registering pre- and post-resection images. Methods for registering images during surgery often make use of FEM models for considering the effect of brain-shift, while in longitudinal registration methods usually a special way of handling the tumor-induced appearance changes is employed. For intra-operative images, computation time is crucial and therefore emphasized by all authors. An overview and comparison of the presented techniques can be found in table 4.

The inter-subject normalization task is also very challenging due to the effects of the tumor mass in combination with inter-subject variations. It appears that masking out the tumor area from the calculation of the similarity metric is too simplistic and more sophisticated methods, which assign different confidence weights to different areas during registration, or which allow for a composition of two different kinds of deformations are a more promising way to solve the problem, however this also affects computation time. An overview and comparison of the presented techniques can be found in table 5.

2.4. Integrated Approaches for Tumor Growth Modeling and Registration / Segmentation

The major problem when registering a brain tumor image with a healthy atlas is the missing correspondence between pathologic patient image and healthy atlas. One idea to circumvent this problem is to seed the atlas with a tumor prior and simulate tumor growth in the atlas according to the patient image, followed by a final non-rigid registration step.

2.4.1. Types of Tumor Growth Models

Tumor growth simulation approaches used in image-based modeling can fall into three different categories: on the microscopic level, cellular growth models are used, while on the macroscopic scale diffusion-reaction models or bio-mechanical models are employed.

Cellular growth models focus on cell-to-cell interactions and are able to model proliferation or apoptosis of individual cells using rule-based cellular automata (Kansal et al. (2000), Stamatakos et al. (2007), Stamatakos et al. (2010)).

Reaction-diffusion models simulate the proliferation and spread of the tumor based on a reaction-diffusion partial differential equation model. The general assumption is that the cancer cells spread along the fiber direction with different diffusion coefficients in gray matter and white matter (Swanson et al. (2000), Swanson et al. (2002)).

Bio-mechanical models take into account the mass-effect of large tumors and model the deformation of the surrounding tissues based on their material properties, usually employing a finite element method (Davatzikos et al. (2001), Zacharaki et al. (2008a), Chen et al. (2010)). For the challenge of handling the large-scale deformations occurring in brain tumor growth, it has been suggested to use an Eulerian instead of a Lagrangian formulation (Hogea et al. (2007a)).

Multi-scale models integrate the methods mentioned previously in order to gain a more precise formulation. Most models simulate proliferation based on a reaction-diffusion equation and couple this with a bio-mechanical model for the mass-effect (Clatz et al. (2005b), Hogea et al. (2007b), Hogea et al. (2008b), Hogea et al. (2008a), Prastawa et al. (2009)). An exception is May et al. (2011), where a cellular proliferation model is coupled to an Eulerian FEM model to simulate the mass-effect. Marias et al. (2011) recently made a first attempt to integrate all three levels of complexity: cellular, diffusion and bio-mechanical into one single multiscale model.

Personalization of these models is difficult, especially for registration purposes when only limited data is available for this task. Most authors employ image-based modeling, making ad-hoc assumptions about the connection between image intensities and cell densities, which again rely on segmentations of tumor and healthy tissues. Other model parameters are either initialized with standard values from literature, or optimization methods are employed for personalizing the models by choosing individual parameters.

2.4.2. Purely Bio-mechanical Growth Models combined with Registration

Kyriacou

et al. (1999) were the first ones to introduce the concept of combining bio-mechanical tumor growth modeling with non-rigid registration in order to adapt a patient image to a healthy atlas. They assumed a uniform strain of the tumor and a non-linear elastic behavior of the surrounding tissues. First, shrinkage of the tumor in the patient image was performed to obtain an artificial healthy patient image. A healthy atlas was then registered to the tumor-free patient image with an elastic registration method. Subsequently, tumor growth was performed on the registered atlas using a regression method to invert the tumor shrinkage process from the previous step. Finally, they obtained a patient-adapted atlas including pathology. The concept was demonstrated on 2D image slices.

Later, Mohamed et al. (2006) from the same group chose a different approach. Instead of shrinking the patient tumor, they grew the tumor in the atlas image according to the pathologic patient image and applied a non-rigid registration to adapt the modified atlas to the patient image. In order to handle the significant computational cost of the tumor growth modeling in 3D, they employed an approach based on a statistical model of tumor-induced deformation using principal component analysis (PCA). For each case, the most likely parameters were estimated and the deformation could be applied based on the pre-built statistical model. Zacharaki et al. (2008b) further developed this approach as a multi-resolution framework for registration of brain tumor images. In their approach, they also used a statistical model of tumor induced deformation, but in a hierarchical framework for parameter optimization. Local information was incorporated into the tumor growth model and the registration methodology was improved. In a next step, Zacharaki et al. (2009a) improved the tumor growth model compared to their previous method. They employed a piecewise Eulerian simulator with a uniform outward pushing pressure model for the bulk tumor, which did not require the simplified PCA model for tumor growth while still being computationally efficient. Additionally, parameter optimization was performed in a parallelized fashion for further speed improvements.

Bauer et al. (2011b) proposed a more simplistic model for tumor growth, which they combined with a non-rigid Demons registration. A healthy atlas was seeded automatically and bio-mechanical tumor growth was modeled assuming a radial expansion force. The expansion was propagated over the image by means of a Markov Random Field method on the deformation field. Characteristics of each method are summarized in table 6.

2.4.3. Diffusion-based Bio-mechanical Growth Models combined with Registration
Gooya et al. (2011a) employed a more sophisticated diffusion-reaction model for tumor growth, which was coupled with an Eulerian method to simulate the mass-effect. They operated on multi-sequence images to obtain a probability map of the different tissue classes using SVM classification techniques. The patient specific tumor was grown in the atlas and a Demons like algorithm was employed for the localized atlas-to-patient transformation for which the tissue probability maps from the multi-sequence
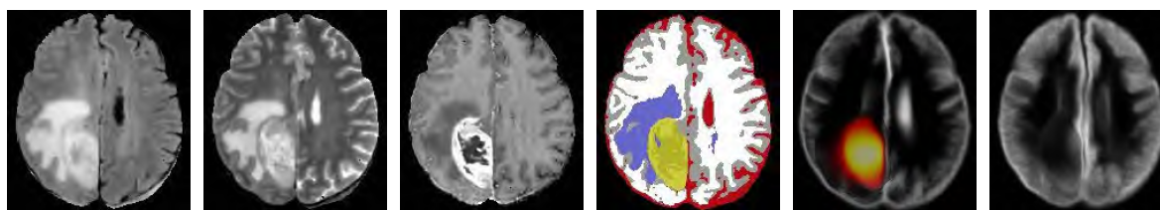
Figure 6. Integrated tumor-growth modeling and non-rigid registration. Axial slices of the three modalities (FLAIR, T_2, $T_{1contrast}$), which were used, are shown on the left. The segmentation result of the GLISTR method is shown in the center. The images on the right depict the obtained probability maps for tumor (color), CSF and gray matter (grayscales) respectively after the combined tumor growth simulation and registration. From Gooya et al. (2012).

classifier were considered for the cost function. The whole process was formulated in an expectation maximization framework to jointly estimate tumor growth and the spatial transformations for adapting the atlas to the patient image. In an extension, Gooya et al. (2011b) dropped the requirement for a pre-classification and proposed a joint segmentation and registration model including tumor growth. To this end, they used again an expectation maximization algorithm to jointly estimate tumor growth parameters and registration deformation field to obtain the tissue posterior probabilities in the patient image derived from the atlas. The approach was recently further improved and extended (Gooya et al. (2012), see figure 6), where the authors also reported on employing the method for the construction of statistical atlases for gliomas. Characteristics of each method are summarized in table 6.

2.4.4. Bio-mechanical Growth Method Based on a Cellular Model combined with Registration Bauer et al. (2012) suggested to use a bio-physio-mechanical tumor growth model for atlas-based segmentation of tumor-bearing brain images. They seeded the atlas with a patient specific prior and simulated tumor growth with a cellular automaton modeling tumor cell proliferation, which was coupled to an Eulerian FEM method to handle the tumor mass-effect. The modified atlas was registered to the patient image with a Demons algorithm and the label map was propagated for segmentation. Recently, they extended this to be applicable to radiotherapy and neurosurgery (Bauer et al. (2013)). The tumor outline was defined using classification methods and important structures were segmented by warping a subcortical label map with an enhanced registration metric. Characteristics of the methods are summarized in table 6.

2.4.5. Summary and Conclusion - Integrated Approaches Several authors have proposed integrated approaches that combine tumor growth modeling with segmentation methods based on registration of an atlas (see table 6 for a summary). To this end, initial correspondence between healthy atlas and pathologic patient image is established with the help of tumor growth simulation. It has been observed that the initially proposed purely bio-mechanical methods, modeling the tumor-mass effect only, were

Table 6. Overview of combined tumor growth modeling and registration/segmentation methods, grouped by different approaches. Evaluation was performed on different datasets, empty cells indicate no reported information. Dim stands for dimensionality, S for supervision, Transf. for the transformation model. SA means semi-automatic, FA fully-automatic, L linear transformation model and NL non-linear transformation model. Type describes the tumor type, which can be: G - glioma, HGG - high grade glioma, LGG - low grade glioma, M - meningioma, Met - metastasis.

Authors	Method	Accuracy	Speed	Dim	S	Transf.	Type
Purely bio-mechanical growth model							
Kyriacou et al. (1999)	Bio-mechanical tumor shrinkage and growth for registration	<3.5mm	<100min	2D	SA	NL	Met
Mohamed et al. (2006)	Statistical model for tumor-induced deformation	<19mm	35min	3D	SA	NL	G
Zacharaki et al. (2008b)	ORBIT	<42mm	14h	3D	SA	NL	
Zacharaki et al. (2009a)	Non-diffeomorphic registration with tumor growth	0.67-0.74 (Dice)	<6h (cluster)	3D	SA	NL	G, Met
Bauer et al. (2011b)	MRF lesion growth model	0.7-0.95 (Dice)	<30min (GPU)	3D	SA	NL	G
Diffusion-based bio-mechanical growth methods							
Gooya et al. (2011a)	Registration with EM algorithm and diffusion modeling	0.29-0.63 (Jaccard)	6-14h (cluster)	3D	FA	NL	G
Gooya et al. (2011b)	Joint Segmentation and Registration	0.76-0.85 (Dice)		3D	FA	NL	G
Gooya et al. (2012)	GLISTR: Glioma image segmentation & registration	0.75-0.84 (Dice)	3-6h	3D	FA	NL	G, Met
Cellular-based bio-mechanical growth methods							
Bauer et al. (2012)	Multi-scale modeling for atlas-based segmentation	0.56-0.95 (Dice)	10-36h	3D	SA	NL	G
Bauer et al. (2013)	Integrated segmentation for radiotherapy & neurosurgery	0.41-0.72 (Dice)	6-48h	3D	FA	NL	G

not satisfactory. Therefore, the more realistic models simulate the mass-effect coupled to either a microscopic cellular proliferation model or a macroscopic diffusion model. A logical next step would be to integrate all three levels: cellular, diffusion and bio-mechanical layer to obtain a more meaningful model. The non-rigid registration methods applied after tumor growth simulation, are mostly standard algorithms which are well-known in the medical image analysis community. It would be interesting to see if additional knowledge about tissue distributions and characteristics in brain tumor images could be useful. However, before the approaches presented in this section can become clinically relevant, computation time has to be drastically reduced.

3. Discussion & Outlook

3.1. Critical Review of the Current State of the Art

In this review article, we gave an overview of the state of the art in MRI-based medical image analysis for brain tumor studies. The focus was on segmentation and registration, whereas some of the registration methods also included tumor growth modeling. The first attempts in this field were made almost two decades ago, but it can be observed that in recent years the methods are becoming mature and an increase of their use in clinical practice is expected.

The majority of segmentation approaches operate on multi-sequence MRI data, employing classification methods using different features and taking spatial information

in a local neighborhood into account. The trend is not to segment the tumor only, but also to delineate tumor sub-compartments and different healthy regions on images from standard clinical acquisition protocols. This provides the physician with a more comprehensive information, on which diagnosis, tumor monitoring and therapy planning can be based. Apart from the evaluation of accuracy and robustness, an important criterium is computation time. Processing times of a few minutes are the current standard. Performing a better problem-oriented selection of the segmentation technique instead of choosing the technique first and then try to make it work on the current problem, could be accomplished in the future by paying more attention to feature selection than the segmentation algorithm.

The registration methods can be separated into intra- and inter-patient registration. In intra-patient registration, the main challenge is to handle effects of tumor growth or resection, which is mostly done by tumor image-specific extensions to standard registration algorithms. The majority of inter-patient registration approaches focus on registration of tumor-bearing brain images with a normal atlas. This can be used for atlas-based segmentation or for constructing statistical brain tumor atlases. One attempt to handle the missing correspondence between healthy atlas and pathologic patient image is solely based on registration methods, a different idea is to combine tumor growth modeling with registration methods. While the registration approaches are more general, the integrated approaches tend to be more accurate, however the tumor-growth model adds additional complexity, which also introduces additional risks. A major problem, especially for the integrated approaches, is their tremendous computation time. Therefore, these methods should be considered as pure research methods at the moment, and will not reach clinical use until improvements in computational speed are attained.

Comparing the traditional segmentation techniques and the atlas-based segmentation methods, which rely on registration, it is clear that the traditional segmentation techniques are more flexible and can be more easily adapted to handle multiple modalities simultaneously. Additionally, it is possible to treat individual tumor sub-compartments more easily. Atlas-based methods have advantages when segmenting tissues and structures surrounding the tumor, especially subcortical structures or functional areas. These structures can be segmented more easily if there is atlas-encoded information about spatial localization available. This can have important implications for surgery or radiation therapy. However, many of the approaches incorporating tumor-growth modeling have difficulties in handling multifocal lesions.

While preparing this article, it became apparent that some papers in the literature lack in precision, describing what exactly is being done, what type and grade of tumor is being considered, what image data is being used or how the algorithm performs in terms of robustness, accuracy and speed. This can be easily seen from the many empty fields in the presented tables. Particularly in the case of medical image analysis for brain tumor studies, better attention should be paid to differences according to the tumor type and grade, but also to robustness of the algorithm. Eventually, it would be useful

to test any new method on a standard database of brain tumor images to allow for a fair comparison against the state of the art. The MICCAI BraTS dataset ‖ would be one candidate for such a database (see also section 2.1).

3.2. Clinical Requirements and Applicability

Although a lot of research has been done in this field in the last years, application in the clinics is still limited. Despite the huge amount of work involved and the limited objectiveness, in many cases clinicians still rely on manual tumor delineations. This is probably mostly due to a lack of communication between researchers and clinicians. Many of the tools developed so far are pure research tools, which are not easy to handle for clinicians. Therefore, a major effort should be spent on embedding the developed tools into more user-friendly environments in the future. So far, in most commercial workstations, only very simple methods are implemented (e.g. thresholding). One reason for this might be, that the technology transfer from the bench to the bedside is still a major challenge because the performance achieved in a well controlled research environment can often not be reproduced in clinical routine. Recently, more researchers have tried to consider standard clinical acquisition protocols when developing their methods, instead of focusing on feasibility studies that employ pure research data as image material. This will hopefully aid in spreading the application more quickly. Another important aspect is computation time: real-time segmentation will be hard to achieve, but computation times which are beyond a few minutes, are unacceptable in clinical routine. For everyday use, robustness is crucial. This includes robustness against slight changes in the acquisition protocol, but also the flexibility to handle new acquisition protocols. If a method tends to fail completely on some cases, doctors will lose their trust in the automatic approach and will not be willing to change their habits. Therefore, thorough demonstration of the robustness should be one of the major steps for each new method.

Currently, image acquisition protocols are not completely standardized. In order to be able to make the best possible use of automatic methods for medical image analysis, it is essential to have image data, which has been acquired according to a well-defined protocol across different clinical sites. We expect such a standardization would aid significantly in improving the applicability and spread the use of automatic methods because the need for fine-tuning according to specific imaging characteristics will be eliminated.

Standard therapeutic procedures for brain tumors include resection, followed by fractionated radiation therapy combined with temozolomide (TMZ) chemotherapy (Stupp et al. (2009)). Due to recent changes in glioma treatment, the traditional Macdonald criteria (Macdonald et al. (1990)) for tumor response evaluation are increasingly recognized to be not adequate any more, as effects of pseudo-progression and pseudo-response are not covered by these criteria. Since the formulation of the

‖ http://www2.imm.dtu.dk/projects/BRATS2012/

Macdonald criteria, many novel MR-techniques have become available, now widely used in clinical routine, that allow for the non-invasive retrieval of morphological, functional and metabolic properties of brain tumors (e.g. diffusion weighted imaging, dynamic susceptibility contrast enhanced perfusion imaging, MR-spectroscopy (MRS)). These techniques may be able to fill the gap in the proper evaluation of tumor progression in the future, together with improvements in tumor volume estimation from medical image analysis by forming an integrated platform for advanced tumor assessment.

Eventually, a dialogue should be initiated to convince practitioners that although a software tool might lack in accuracy compared to the manual method in certain cases, it offers more objectiveness, which is very important in longitudinal studies. However, while being objective, the automatic segmentation obviously has to be sufficiently close to what a radiologist perceives as the correct tumor segmentation. If this cannot be achieved by fully automatic methods, it might be worthwhile to consider an interactive method that is initialized or corrected by the user. In practice, the variability introduced by interactive user input is usually much smaller than the variability of completely manual segmentation results. In this case, ease of use for the clinician is crucial for any method to be accepted. It should also not be neglected that for automatic tumor image analysis methods to be used as a routine diagnosis tool, certification is mandatory. This probably means that the application will not reach a breakthrough in the clinics before commercial software packages are being offered. Current clinical viewing systems are often not even well equipped yet for visualizing changes in multiple image classes over time.

3.3. Publicly Available Toolboxes

Only very few implementations of the presented methods are publicly available. This impedes comparison of new methods to existing approaches and also hinders large-scale evaluation with clinical data from different sites. To the best of our knowledge, at the moment there are only three methods publicly available for download, which are dedicated to the analysis of brain tumor images:

Prastawa et al. (2009) provide their "TumorSim" software for simulation of synthetic brain tumor images ¶. Bauer et al. (2011a) offer the "DoctorNo" suite of tools (Bauer and Reyes (2011)) as a plugin for the GUI-based "DoctorEye" software platform (Skounakis et al. (2010)) for download [+]. Gooya et al. (2011a) provide the "GLISTR" toolkit on their website [*].

3.4. Future Directions and Trends

It can be expected that the current efforts for registration and segmentation of tumor-bearing brain images on standard MRI protocols will enter routine clinical procedures in

---

¶ www.nitrc.org/projects/tumorsim
[+] www.istb.unibe.ch/content/research/medical_image_analysis/software
[*] www.rad.upenn.edu/sbia/software/glistr

the next years (Levy and Rubin (2011)). In order for this to happen, proof of robustness and sufficient speed of the methods is mandatory, as outlined before. This development is likely to be pushed with the anticipated shift from the diameter-based MacDonald (Macdonald et al. (1990)) and RECIST (Therasse et al. (2000)) criteria to volume-based criteria for assessment of tumor response to treatment as suggested by the RANO group (Wen et al. (2010)). However, it remains to be investigated how well automatic methods can handle the impact of treatment effects and related pseudo-progression or pseudo-response on the segmentation outcome, which will also be interesting for clinical trials (Henson et al. (2008)). Nevertheless, the objectiveness of automatic methods in tumor segmentation is undisputed and Angelini et al. (2012) have made use of this in a first comprehensive study on longitudinal brain tumor data. It is likely that this line of research will be exploited more in the future.

Apart from the standard imaging methods, recently more advanced MRI modalities like perfusion imaging, diffusion tensor imaging (DTI) and magnetic resonance spectroscopy (MRS) are getting more attention (Cha (2006), Law (2009)) in brain tumor analysis. These modalities are mostly used for tumor grading and for localization of the most active areas of the tumor. Researchers have started to apply machine learning approaches for tumor grading from perfusion images (Slotboom et al. (2008), Zacharaki et al. (2009b), Emblem et al. (2009), Zöllner et al. (2010), Zacharaki et al. (2011)) and also for analysis of combined MRI/MRS data (Luts et al. (2007), Fayed et al. (2008), Georgiadis et al. (2011)). PET imaging has the potential to provide better prognostic information and distinguish recurrent tumor from radiation necrosis (Chen (2007); Dhermain et al. (2010)).

Another area, which is likely to get increased attention is patient-specific image-based modeling of brain tumor growth (Konukoglu et al. (2010), Menze et al. (2011)). This research could help to predict tumor progression and optimize treatment options on an individual basis.

Acknowledgments

References

Ahmed, S., Iftekharuddin, K. M., and Vossough, A. (2011). Efficacy of texture, shape, and intensity feature fusion for posterior-fossa tumor segmentation in MRI. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 15(2):206–13.