RESEARCH

Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization

Nicolas Sauwen[1,2*†], Marjan Acou[3], Diana M Sima[1,2], Jelle Veraart[4], Frederik Maes[5], Uwe Himmelreich[6], Eric Achten[3] and Sabine Van Huffel[1,2]

Correspondence:
icolas.sauwen@kuleuven.be
Department of Electrical
ngineering (ESAT), STADIUS
entre for Dynamical Systems,
ignal Processing and Data
nalytics, KULeuven, Kasteelpark
renberg, Leuven, Belgium
ull list of author information is
vailable at the end of the article
Equal contributor

Abstract

Background: Segmentation of gliomas in multi-parametric (MP-)MR images is challenging due to their heterogeneous nature in terms of size, appearance and location. Manual tumor segmentation is a time-consuming task and clinical practice would benefit from (semi-)automated segmentation of the different tumor compartments.

Methods: We present a semi-automated framework for brain tumor segmentation based on non-negative matrix factorization (NMF) that does not require prior training of the method. L1-regularization is incorporated into the NMF objective function to promote spatial consistency and sparseness of the tissue abundance maps. The pathological sources are initialized through user-defined voxel selection. Knowledge about the spatial location of the selected voxels is combined with tissue adjacency constraints in a post-processing step to enhance segmentation quality. The method is applied to an MP-MRI dataset of 21 high-grade glioma patients, including conventional, perfusion-weighted and diffusion-weighted MRI. To verify robustness against user input variability, each patient's dataset is repeatedly analyzed with randomly selected seeding points from the pathological subcompartments.

Results: Using L1-regularized semi-automated NMF segmentation, mean Dice-scores of 65%, 74% and 80% are found for active tumor, the tumor core and the whole tumor region. The mean of the lower quartile Dice-scores across patients are 60%, 68% and 77%, respectively. The mean of the higher quartile Dice-scores across patients are 74%, 80% and 85%. Lower Dice-scores are found without L1-regularization and when only considering conventional MRI data.

Conclusions: The reported segmentation scores are competitive with state-of-the-art in literature, even when looking at the lower quartile values across patients. Robust results were found for most patients, although careful voxel selection is mandatory to avoid sub-optimal segmentation.

Keywords: MRI; segmentation; brain tumors; non-negative matrix factorization; unsupervised classification

Background

High-grade gliomas (HGGs) account for 80% of all malignant primary brain tumors [1]. Standard treatment of HGG consists of complete or partial resection of the

gross tumor volume, followed by radiotherapy and/or chemotherapy to attack the remaining tumor cells and to counteract tumor recurrence. So far, recent advancements in glioma research have only had little effect on patient outcome. The 5 year survival rate of the most common histopathological subtypes, anaplastic astrocytoma and glioblastoma multiforme (GBM) are 26% and 5%, respectively.

MRI has become the imaging modality of choice for evaluating tumor progression and the efficacy of a chosen treatment strategy. GBM, the most malignant and most common type of glioma, tends to invade the healthy tissue, such that tumor margins extend beyond the imageable component of the tumor based on conventional MRI (cMRI). Numerous recent studies have recommended incorporating additional imaging biomarkers from advanced MRI modalities such as perfusion-weighted imaging (PWI) and diffusion-weighted imaging (DWI) [2–4]. PWI is used for studying tumor angiogenesis, as HGGs are known to stimulate vascular ingrowth [5]. PWI measurements in the peritumoral region have been reported to differentiate GBM from metastases, suggesting the detection of tumor infiltration [6]. PWI was also found useful in differentiating tumor recurrence from radiation necrosis [7]. DWI assesses the mobility of water molecules within tissue due to Brownian motion. Because glioma infiltration disrupts the organization of the white matter tracts, DWI is potentially useful to characterize the extent of infiltration. Diffusion tensor imaging has been reported to better delineate tumor margins in gliomas than cMRI [8].

Assessment of tumor extent plays a key role at all stages of the treatment process. An outline of the gross tumor volume is made upon planning surgical resection and for defining the radiotherapy target volume. During follow-up, the RANO criteria are applied for monitoring tumor growth, assessing the maximal diameter of the lesion in 2 orthogonal directions on axial slices [9]. Clinical practice would benefit from volumetric measurements of the tumor and its subcompartments. With the recent emergence of the dose-painting concept in radiotherapy, segmentation of the active tumor region would allow focusing the radiation energy to the most active part of the tumor in a non-uniform way [10]. Volume measures of the active tumor and necrotic region have been reported to be significant predictors of patient

outcome to treatment [11, 12]. Furthermore, the RANO criteria acknowledge that volumetric measurements might be favorable compared to cross product diameters in cases of irregularly shaped tumors, multi-focal tumors and tumors with cystic or necrotic components. However, segmentation of HGGs in MRI is a challenging task, due to their heterogeneous nature: several stages of the disease can occur throughout the same lesion and diffuse boundaries exist between active tumor, necrosis, edema and the surrounding healthy brain. Manual segmentation of 3D images is also time-consuming, which is the main reason why volumetric tumor delineation has not become widespread in clinical practice.

In recent years, significant advancements have been made in the field of automated brain tumor segmentation. Both semi-automatic and fully automatic methods have been proposed [13]. A popular approach is to combine imaging biomarkers from different MRI sequences on a voxel-wise basis, thereby increasing specificity and reducing overlap between tissue classes. Nowadays, supervised classification methods are receiving most attention [14]. These methods rely on an extensive set of training images with manually annotated ground truth to learn decision boundaries between the tissue classes in feature space. The most popular methods include random forests [15, 16], support vector machines [17, 18] and neural networks [19]. Additional constraints may be imposed to further enhance performance. For instance, one can assume that a voxel is likely to belong to the same tissue class as its neighbors, instead of classifying all voxels independently. Markov Random Fields [20] and Conditional Random Fields [21] are common approaches to encourage such spatial consistency, either as a regularization term in the optimization function [15, 17] or as a post-processing step [22].

Unsupervised classification methods are also being considered for tumor segmentation. These methods are very flexible, as they don't require an extensive training dataset with a uniform acquisition protocol. They learn classification rules directly from the imaging data at hand, based on some similarity criterion. Popular approaches include fuzzy C-means clustering (FCM) [23, 24], Gaussian mixture modeling [25, 26], hidden Markov Random Fields [27, 28] and non-negative matrix factorization (NMF) [29, 30]. Due to the absence of training data, unsupervised

methods rely more strongly on the incorporation of prior knowledge to obtain competitive results. As gliomas exhibit a wide variability in terms of size, shape, location and appearance, imposing proper prior knowledge is difficult. Prastawa *et al.* used a normal brain atlas to obtain healthy tissue priors, and added a tumor prior based on the assumption of tumor enhancement on T1+contrast (T1C) [31]. These priors along with the MRI features were then fed to a Gaussian mixture model for tissue classification. The authors acknowledged that their algorithm might not adequately detect tumors which are non-enhancing or tumors with a large mass effect. Fletcher-Heath *et al.* applied knowledge-based FCM to segment non-enhancing gliomas [23]. Clusters containing tumor as well as CSF were separated based on PD-enhancement. The tumor region was then separated from the ventricular CSF by assuming symmetry of the ventricles across the hemispheres. Finally, extraventricular CSF and other false positives are removed from the tumor mask using morphological post-processing steps, assuming only one tumor volume. Kazerooni *et al.* applied FCM combined with manually initiated region growing to segment GBMs on a MP-MRI dataset combining T2, DWI and PWI [24]. The active tumor region was found by applying FCM to the PWI data, assuming hyperperfusion throughout the region. Saha *et al.* presented a computationally efficient tumor detection algorithm, by assuming symmetry across the hemispheres in healthy brain [32]. A bounding box is obtained within one hemisphere which maximizes the histogram-based dissimilarity with the corresponding contralateral region based on the Bhattacharya coefficient. Görlitz *et al.* applied a hidden Markov Random Field model to MRSI data of brain tumors [33]. They used a graph-cut algorithm to efficiently solve the spatio-spectral model, resulting in a binary tumor mask without differentiating the tumor subcompartments. Similarly, Ortega-Martorell *et al.* applied Convex NMF to MRSI data of brain tumor-bearing mice to delineate the entire pathological brain area [30]. Evidently, many unsupervised segmentation studies incorporate prior knowledge in the form of simplifying rules, limiting their applicability to a subset of gliomas with specific characteristics: the tumor is assumed to be enhancing or non-enhancing, deformation of the surrounding healthy tissue is assumed to be small, multi-focal lesions can not be detected, the lesion does not extend to the contralateral hemisphere, etc. Some methods are limited to a particular set of MRI parameters, as they rely on one or more specific imaging biomarkers. Other methods simplify the

segmentation problem by only providing an outer contour of the pathological region, without differentiating active tumor from the other tumor subcompartments.

Semi-automated segmentation algorithms use a different approach. They incorporate prior knowledge in the form of user-specific input, either as an initialization or as a post-processing step. Competitive performance has been reported for semi-automated algorithms [14]. Kwon *et al.* combined a normal brain atlas with a tumor growth model, based on user-defined tumor seeds [34]. This allowed them to generate a patient-specific atlas which also models the tumor mass effect. Through the user-defined seeding, multi-focal gliomas could be handled as well. Hamamci *et al.* initialized a cellular automata (CA) algorithm based on the maximal tumor diameter as drawn by the user on T1C images [35]. An active level-set surface is then initialized from the user-defined maximal diameter, to impose spatial smoothness on the segmented pathological region. Finally, CA is combined with Otsu's method to differentiate active tumor from necrosis. The authors acknowledged the limited applicability of their method to enhancing tumors. Havaei *et al.* initialized a k-nearest neighbors (kNN) classifier with user-selected voxels in the tumor subcompartments as well as in the healthy brain regions [36]. In addition to the MRI imaging features, spatial coordinates of the seeding points were also exploited in the feature set. Spatial consistency was encouraged by adding regularization in the form of Markov Random Fields and Conditional Random Fields, which further improved segmentation results. The main drawback of semi-automated methods is that they don't provide reproducible results, as the segmentation depends on subjective user input. Robustness against user input variability is therefore a key aspect of these methods. In the current study, we propose a semi-automated brain tumor segmentation method based on regularized non-negative matrix factorization. User-defined seeding points in the pathological regions are combined with a sophisticated seeding method for the normal brain tissues to initialize the NMF algorithm. Piece-wise spatial smoothness as well as sparseness of the NMF tissue abundance maps are encouraged through L1-regularization. Morphological post-processing based on the spatial location of the user-defined seeds is exploited to further remove false positives. The proposed method is applicable to all types of gliomas with any MP-MRI dataset, as the user-defined prior knowledge is patient-specific. We illustrate seg-

mentation performance on an MP-MRI dataset of 21 HGGs combining cMRI, PWI and DWI. To verify robustness against user input variability, each patient's dataset is repeatedly analyzed with randomly selected seeding points from the pathological subcompartments.

Methods

Multi-parametric MRI dataset

Twenty-one patients who were diagnosed with a HGG were enrolled in the study: 1 grade II astrocytoma with focal progression to grade III anaplasia, 2 grade II oligodendrogliomas with focal progression to grade III, 1 oligo-astrocytoma with focal progression to grade III, 6 anaplastic oligo-astrocytomas, 1 anaplastic astrocytoma with focal progression to GBM and 10 GBMs. The Ghent University Hospital local ethics committee allowed a retrospective analysis of the data. The MR examinations were performed on a 3T Siemens Trio Tim scanner (Erlangen, Germany), using a standard 12-channel phased array head coil. All patients underwent an MP-MRI acquisition protocol, consisting of cMRI, PWI and DWI. cMRI consisted of a 3-dimensional T1-weighted gradient-echo sequence (MPRAGE) before and after contrast administration, with isotropic voxels (176 sagittal slices, field of view (FOV) read = 220mm, voxel size $0.9 \times 0.9 \times 0.9\text{mm}^3$, Repetition Time (TR) = 1550ms, Echo Time (TE) = 2.39ms, Inversion Time (TI) = 900ms, matrix size = 256×256, flip angle = $9°$) and a 3-dimensional T2-weighted inversion recovery sequence (FLAIR) with isotropic voxels (176 sagittal slices, FOV read = 250mm, voxel size $1 \times 1 \times 1\text{mm}^3$, TR = 6000ms, TE = 421ms, TI = 2100ms, matrix size = 256×238). PWI was performed by using a lipid-suppressed, T2*-weighted echo-planar imaging sequence with the following parameters: TR = 1000ms, TE = 29ms, FOV = $230 \times 230\text{mm}^2$, slice thickness = 5.0mm, matrix size = 128×128, in-plane voxel size = $1.8 \times 1.8\text{mm}^2$, 15 slices, distance factor = 30%, flip angle = $90°$. A series of 90 multi-section acquisitions was acquired at 1 second intervals. The first 10 acquisitions were performed before contrast agent injection to establish a pre-contrast baseline. At the tenth acquisition, a 0.1mmol/kg body weight bolus of gadobutrol (Gadovist, Bayer) was injected with a power injector (Spectris, Medrad Inc., Indianola, PA) at a rate of 4ml/s through a 18-gauge intravenous catheter, immediately followed by a 20 ml bolus of sodium chloride solution at 4ml/s. Relative

cerebral blood volume (rCBV) maps were derived from the dynamic signal intensity curves in DSCoMAN (Duke University, Durham, NC) using the method proposed by Boxerman *et al.* which incorporates leakage correction [37]. Axial diffusion-weighted images were acquired using a fast single-shot gradient-echo echo-planar imaging sequence with diffusion gradient b-values of 0, 500 and 1000s/mm^2 (voxel size $2.0 \times 2.0 \times 3.0$mm^3, TR = 5400ms, TE = 80ms, FOV read = 264mm, 3 orthogonal directions). An affine coregistration was applied to account for eddy currents. Apparent Diffusion Coefficient (ADC) maps were derived from the 3 b-values using weighted linear least squares fitting [38]. The b0 images were also added to the dataset, serving as a T2-weighted reference.

Six MRI features were obtained from the raw acquired data after pre-processing: T1, T1C, FLAIR, rCBV, ADC and b0. All MP-MRI features were coregistered and resampled to the same spatial resolution of $1 \times 1 \times 3$mm^3. cMRI data were skull-stripped and T1C served as a reference for rigid coregistration in SPM8 (Wellcome Trust Centre for Neuroimaging, University College London), using the normalized mutual information criterion [39] and cubic B-spline interpolation for reslicing. Analyses were performed on 10 axial slices located around the tumor centre. Additional intensity-based features were added to the feature set to include localized spatial information. An in-plane local neighborhood of 3×3 and 5×5 voxels was used to calculate average intensity values that were assigned to the central voxel. These spatially averaged intensity values were added for all MP-MRI features. Each feature's full range was rescaled linearly to [0-1]. A total of 18 MRI features was finally obtained, making up the rows of the input matrix X.

Non-negative matrix factorization

NMF provides a rank-r approximation of a non-negative input matrix X by the product of 2 non-negative factor matrices, W and H:

$$X \approx WH \quad \text{with} \quad X \in R_+^{m \times n}, W \in R_+^{m \times r} \quad \text{and} \quad H \in R_+^{r \times n} \tag{1}$$

with m being the number of input features and n the number of data points in X. NMF reveals an additive parts-based structure of the input data. It represents

each column of X by a weighted sum of the r columns of W. As we are dealing with image intensities, the non-negativity constraint applies naturally. Each column of X corresponds to one voxel's MP-MRI feature set. Each column of W represents a tissue-specific signature, i.e. an MP-MRI feature vector corresponding to one pure tissue type. As such, each voxel's MP-MRI feature vector is approximated as a weighted sum of (tissue-specific) source vectors. Each column of H represents the weights of the tissue types for one voxel. One row of H contains the abundances of one particular tissue type over all the voxels, which can be transformed back into the image space to obtain a tissue abundance map. The most commonly used objective function to measure the quality of the factor approximation is the Frobenius norm, which is based on the squared Euclidean distance:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}\|X - WH\|_F^2 \quad ,\forall i,j : W_{i,j} \geq 0, \quad H_{i,j} \geq 0 \tag{2}$$

As NMF is an NP-hard problem, a locally optimal solution is usually found rather than a global optimum. A wide range of algorithms has been developed to solve the above NMF formulation [40]. We are using the structured data fusion framework [41] as implemented in Tensorlab [42]. A transformation of variables is used to convert the constrained optimization problem in Equation 2 into an unconstrained problem, by squaring the entries of the factor matrices. The Gauss-Newton algorithm with dogleg trust region (GNDL) is used to solve the resulting non-linear least-squares problem [43]. At each iteration of the GNDL algorithm, a step is calculated by iteratively solving a linearized version of Equation 2 using conjugate gradients. The structured data fusion framework allows straightforward incorporation of regularization terms into the objective function. An L1-regularization term was included to promote both spatial consistency as well as sparseness of the tissue abundance maps:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}(\|X - WH\|_F^2 + \lambda\|(L+I)H\|_1) \tag{3}$$

The L1-regularization term in Equation 3 in fact consists of 2 terms. The first term, LH, promotes piece-wise smoothness on the tissue abundance maps. L is a

sparse $n \times n$ matrix, with each row containing a vectorized Laplacian kernel. As such, each row of L applies a two-dimensional second order spatial derivative to the corresponding voxel. An in-plane neighborhood of 4 voxels is considered for the Laplacian kernel. Due to the relatively low out-of-plane resolution of the MRI data, neighboring voxels in adjacent slices are not considered. The second term applies L1-regularization to H directly, imposing sparseness to the abundance maps. A maximum number of 500 iterations and a convergence tolerance of 10^{-6} for the relative difference of two subsequent values of the objective function were used as stopping criteria for the NMF analyses. The regularization coefficient λ was empirically set to 0.1.

NMF initialization

As NMF poses a non-convex optimization problem, initialization of the factor matrices (W_0 and H_0) is required. Careful initialization is important, as it might speed up the convergence process and influence the final segmentation result. The pathological tissue sources are initialized based on voxel selection. The user is required to select one or more voxels in each pathological region, i.e. active tumor and, if present, necrosis and edema. For each selected voxel, a candidate source vector is added to W_0 as the averaged feature vector of the selected voxel and its 4 in-plane neighboring voxels. For each pathological tissue class, we then calculate the correlation coefficient between the candidate source vectors as their inner vector product after normalization. The initial pathological sources are obtained after merging candidate source vectors with a correlation coefficient higher than 0.95 in W_0 and replacing them by their average, thereby reducing complexity of the NMF model.

As the normal brain tissue types (i.e. white matter, grey matter, CSF and blood vessels) are still highly abundant in the affected brain, an automated seeding procedure is used to obtain their initial source vectors. To cope with the variance within the tissue classes, 2 sources are assigned to each normal tissue type, leading to a total of 8 normal tissue sources. We use the Successive projection algorithm (SPA) [44] to obtain an initial estimate of the normal sources. SPA returns a subset of voxels from the input matrix X with minimal collinearity in feature space. All columns of X are first projected onto the orthogonal subspace of the already initialized patho-

logical sources. A first normal seeding point is selected as the voxel with the highest L2-norm in the orthogonal subspace. All columns are then further projected onto the orthogonal subspace of the added voxel, and the next seeding point is selected as the point with the highest L2-norm among these projected columns. The above procedure is repeated until 8 voxels have been selected. Similarly to the pathological sources, the normal sources are obtained as the averaged feature vector of each selected voxel and its 4 in-plane neighboring voxels, assuming these to belong to the same tissue type as the selected voxel. As SPA is known to be sensitive to outliers [45], an additional FCM procedure is applied, using the initial pathological sources and the normal sources obtained from SPA to initialize the cluster centroids. FCM alternatingly updates the cluster centroids and the cluster membership values. As we already obtained proper initialization of the pathological tissue types through user input, the pathological cluster centroids are forced to remain the same throughout the FCM updating procedure. The final FCM centroids make up the columns of W_0. The columns of H_0 are found by applying non-negative least squares fitting to the corresponding column of X and W_0. To reduce computation time, a downsampling factor of 2 was applied to the x- and y-direction of each image slice. After NMF analysis, the obtained tissue abundance maps were upsampled again through linear interpolation. Performance was not affected by this downsampling procedure.

Morphological post-processing

After NMF analysis, each voxel is assigned to the source for which it has the highest abundance value. Voxels that were assigned to different sources belonging to the same pathological tissue class are merged to obtain a preliminary tissue segmentation mask. A morphological post-processing procedure, consisting of 2 steps is then applied to the pathological segmentation masks to further remove false positives, by exploiting the known location of the user-selected voxels. Spatial consistency of the tissue regions is assumed, therefore the segmentation masks are analyzed in terms of (3D-)connected components. For each selected pathological tissue type, only the connected component closest to each user-defined voxel is withheld in the corresponding tissue segmentation mask (Step 1). As the user-defined voxels are not necessarily selected close to the centre of the tissue region, the distance between a selected voxel and a connected component is defined as the minimal distance be-

tween the selected voxel and any voxel from the connected component. It is possible that the same connected component is found for several or all user-selected voxels of the same tissue class.

To avoid missing disjoint regions of a tumor component for which no voxels were selected by the user, an additional step is performed which assumes spatial connectivity of the various pathological components (Step 2). Necrotic areas are always adjacent to an active tumor region. Therefore, any connected component of the preliminary necrosis mask which has a common edge with any withheld active tumor component is also included. Similarly, components of the preliminary active tumor mask are also included if they have a common edge with any withheld necrotic component. The same procedure is also applied to edema, by verifying spatial adjacency to the active tumor region(s). Figure 1 illustrates the morphological post-processing procedure for the necrotic tissue mask. One small necrotic component is not withheld after Step 1, as the user didn't click in this region. However, it was still recovered in Step 2 by verifying spatial connectivity with the active tumor mask.

Validation

Segmentation results of the pathological tissue regions were compared against manual segmentation by an experienced radiologist. The Dice-score was used to quantify the spatial alignment between semi-automated and manual segmentation:

$$Dice_{\text{tissue}} = 2 \times \frac{A_{\text{tissue,NMF}} \cap A_{\text{tissue,manual}}}{A_{\text{tissue,NMF}} + A_{\text{tissue,manual}}} \tag{4}$$

where $A_{\text{tissue,NMF}}$ is the area segmented by NMF and $A_{\text{tissue,manual}}$ the area segmented by the radiologist for the same tissue type. In analogy to previous work [14], Dice-scores are reported for active tumor, the tumor core (active tumor + necrosis) and the whole tumor (tumor core + edema). To verify robustness of the semi-automated method to user input variability, the NMF analysis was repeated 20 times per patient with different selection of the seeding points. In each run, an automatic random selection of 3 points was performed in each pathological region, based on the manual segmentation masks.

Besides NMF with spatial regularization and sparseness ($\mathbf{NMF}_{\text{spatial_sparse}}$), additional NMF analyses were performed to assess the added value of the regularization term and the advanced MRI modalities: NMF without regularization ($\mathbf{NMF}_{\text{no_reg}}$), NMF with only spatial regularization ($\mathbf{NMF}_{\text{spatial}}$), and NMF with spatial regularization and sparseness when only considering cMRI data ($\mathbf{NMF}_{\text{cMRI}}$). For the $\mathbf{NMF}_{\text{cMRI}}$ analyses, rCBV and ADC were omitted from the MP-MRI feature set but b0 was withheld, acting as a surrogate for T_2-weighted MRI.

Results

Figure 2 gives an example of the regularized NMF results and tissue segmentation for a GBM patient with a typical ring-enhancing lesion. Some of the MP-MRI input maps are shown on the first row. The NMF abundance maps as well as the final segmentation masks are shown for the pathological tissue types on the second row.

Figure 3 shows the dispersion of the Dice-scores for active tumor per patient over 20 runs. 16 out of 21 patients have a median Dice-score of at least 60%. 15 patients have a lower quartile Dice-score of at least 50%. 17 patients have an interquartile range not higher than 15%. 7 patients had at least one run with a Dice-score lower than 30%. A large spread in the Dice-scores is seen for patient 6, who has a non-enhancing anaplastic astrocytoma with focal areas of enhancing GBM. Variability in the results was caused by the difficult differentiation of non-enhancing tumor from edema. Patient 8 suffers from a non-enhancing bifocal GBM, with only a small enhancing area in the smaller lesion. Due to the random selection of voxels in the active tumor region, in some runs voxels were only selected from one of both lesions, resulting in a failure to detect the other lesion. A large dispersion in the Dice-scores is also found for patient 18, exhibiting a heterogeneous and irregularly shaped GBM with varying degrees of enhancement in the active tumor region. Active tumor and necrosis were not well differentiated in some runs, mainly due to ambiguous voxel selection near the pathological tissue boundaries. The low outlier scores for patient 10 are due to the difficult differentiation of non-enhancing tumor from edema. The low outlier score for patient 13 is explained by a suboptimal and unrepresentative random voxel selection in the active tumor region.

Boxplots for the tumor core are shown in Figure 4. 18 out of 21 patients have a median Dice-score higher than 65%. For 17 patients, the lower quartile Dice-score was at least 60%. The interquartile range was not higher than 15% for 17 patients. 7 patients have at least one run with a Dice-score lower than 40%. For 5 out of the 7 patients, these low Dice-scores were found to be outliers (i.e. at a distance of more than 1.5 times the interquartile range from the lower quartile value).

Figure 5 shows the boxplots for the whole tumor region. A median Dice-score higher than 70% was found for 19 out of 21 patients. 18 patients have lower quartile value higher than 60%. 5 patients have at least one run with a Dice-score lower than 45%. For 4 out of these 5 patients, such low values were found to be outliers.

Table 1 gives a comparison of the mean Dice-scores over all patients for the different NMF analyses ($\mathbf{NMF}_{\text{no_reg}}$, $\mathbf{NMF}_{\text{spatial}}$, $\mathbf{NMF}_{\text{spatial_sparse}}$ and $\mathbf{NMF}_{\text{cMRI}}$). Overall, the best performance is obtained with $\mathbf{NMF}_{\text{spatial_sparse}}$. When considering only spatial regularization ($\mathbf{NMF}_{\text{spatial}}$), segmentation results are mainly worse for the active tumor region, where the Dice-scores are even lower than for $\mathbf{NMF}_{\text{no_reg}}$. The spatial smoothing was found to be too severe for several GBMs with a narrow ring-enhancing active tumor compartment. Compared to $\mathbf{NMF}_{\text{no_reg}}$, Dice-scores are higher with $\mathbf{NMF}_{\text{spatial_sparse}}$ for active tumor and the tumor core, with an increase of 2% in most cases. The Dice-score for the whole tumor region is 0 to 1% higher for $\mathbf{NMF}_{\text{spatial_sparse}}$ compared to $\mathbf{NMF}_{\text{no_reg}}$. $\mathbf{NMF}_{\text{spatial_sparse}}$ does not show any improvement compared to $\mathbf{NMF}_{\text{cMRI}}$ for the active tumor region. An increase in Dice-score of 1 to 2% is found for the tumor core and an increase of 3 to 4% for the whole tumor region.

Discussion

Most of the unsupervised brain tumor segmentation algorithms in literature incorporate prior knowledge in the form of basic assumptions or heuristics, which often limits their general applicability to a subset of gliomas or to a particular set of MP-MRI data. Prior knowledge in the form of user input is flexible, allowing to incorporate patient-specific information regarding appearance, location and/or shape of the tumor. Using our semi-automated framework, we were able to tackle

Table 1 Mean Dice-scores [%] for NMF without regularization, with spatial regularization (λ=0.1), with spatial regularization and sparseness (λ=0.1), and using only cMRI data with spatial regularization and sparseness. For each tissue class, the first column reports the mean of the 25^{th} percentile Dice-score across all patients, the second column the mean of the mean Dice-score across all patients and the third column the mean of the 75^{th} percentile Dice-score.

	Active tumor			Tumor core			Whole tumor		
	25_{th} prcntile	Mean	75_{th} prcntile	25_{th} prcntile	Mean	75_{th} prcntile	25_{th} prcntile	Mean	75_{th} prcntile
$\mathbf{NMF}_{\text{no_reg}}$	58	63	72	66	72	79	76	80	84
$\mathbf{NMF}_{\text{spatial}}$	55	60	70	65	71	79	76	80	85
$\mathbf{NMF}_{\text{spatial_sparse}}$	60	65	74	68	74	80	77	80	85
$\mathbf{NMF}_{\text{cMRI}}$	60	65	74	67	72	79	73	77	82

the main limitations of many unsupervised classification methods like NMF. Most unsupervised algorithms require proper initialization to come to a valid locally optimal solution. Several studies cope with this by running the segmentation algorithm numerous times with a randomized initialization, then selecting the final solution based on a predefined objective function [25, 29]. We are combining user-defined voxel selection for initializing the pathological sources with a sophisticated initialization of the normal brain tissue sources based on SPA and FCM. As SPA is sensitive to outliers, its output is fed into the FCM algorithm, thereby providing a deterministic initialization for FCM.

Due to the lack of labelled training data, automatic assignment of a tissue label to each segmented region is non-trivial for unsupervised methods. Several studies do not explicitly propose an automated labelling strategy [26, 28], while others assume specific characteristics of the tumor compartments in specific MRI images (such as e.g. contrast enhancement of active tumor) [23, 24]. In our proposed methodology, tissue labelling automatically results from the voxel selection in the pathological regions.

Unsupervised methods model the input with a limited number of degrees of freedom, based on a global optimization criterion. As such, they tend to obtain data clusters of similar size and they do not promote the detection of small pathological tissue regions [46]. Some studies cope with this by using an over-determined

model, i.e. by setting the number of clusters/sources higher than the actual number of tissue classes. This approach also allows the detection of regions with different characteristics belonging to the same tissue class (e.g. enhancing and non-enhancing tumor). However, it requires an intelligent post-processing for fusing the clusters from the same tissue class. Juan-Albarracin *et al.* applied hierarchical agglomerative clustering to merge similar classes using the Jensen-Shannon divergence criterion to measure distances between the probability density functions of different clusters [25]. Hsieh *et al.* obtained 32 segmented regions after FCM combined with region growing, then eliminated non-pathological areas and merged pathological areas based on assumptions regarding tumor size, symmetry across the hemispheres and spatial adjacency [47]. We used on average 12 to 13 sources to model 6 to 7 tissue classes, with the number of sources depending on the number of initial pathological sources that were merged. Fusing of the segmented regions was straightforward, as pathological regions had already been labelled through voxel selection.

False positive regions were excluded from the pathological tissue masks by exploiting the spatial location of the selected voxels in a morphological post-processing procedure. As the pathological tissue classes are assumed to form spatially consistent regions, only the connected component closest to each selected voxel is withheld in the segmentation masks. Loss of tumor components not selected by the user is avoided by assuming spatial adjacency of the pathological regions. Similar types of post-processing have been proposed to enhance segmentation results. Cordier *et al.* only withheld 1 or 2 connected components for each tissue region based on their size, assuming that the largest component(s) correspond to the true tissue region [48]. Menze *et al.* exploit knowledge about shape and location of the tumor based on an atlas tissue prior to remove false positives [26]. Bauer *et al.* apply spatial regularization in the form of a conditional random field, penalizing neighboring voxels with different labels depending on the likelihood of tissue adjacencies [15]. Havaei *et al.* initialize a k-nearest neighbors (kNN) classifier based on voxel selection by the user [36]. They do not exploit the spatial location of the selected voxels in a post-processing step, but instead they add voxel coordinates to the feature set of the kNN classifier.

Comparison to other segmentation studies is sometimes hampered by the fact that different segmentation metrics are being used, or because of a different definition of pathological subregions. We have quantified our results for active tumor, the tumor core and the whole tumor region using the Dice-score, in accordance with the Multimodal Brain Tumor Segmentation (BRATS) challenge held at the Medical Image Computing and Computer Assisted Intervention (MICCAI) conference. Looking at the results reported for HGGs on the BRATS 2012 and 2013 datasets [14], it can be seen that our results are competitive, even when looking at the mean of 25^{th} percentile values (see Table 1). For the whole tumor region, Dice-scores are close to or even higher than 80%, which is in the range of inter-observer variability [14]. Overfitting to the presented dataset seems unlikely as we are using an unsupervised method with no tuning parameters. Excellent results compared to literature have been reported for other semi-automated segmentation methods as well, either based on a supervised [36] or an unsupervised classifier [24, 34]. It has to be noted that the current BRATS dataset only contains cMRI data. Comparison might therefore be more appropriate to the $\mathbf{NMF}_{\text{cMRI}}$ results, for which only cMRI data were considered. The inclusion of additional MRI modalities for brain tumor segmentation has been commonly suggested [2, 49] and initial studies have found improved segmentation results with extended MP-MRI datasets [29, 50]. When including ADC and rCBV into the MRI feature set, we found no advantage for the active tumor region, but Dice-scores increased by 1 to 2% for the tumor core and by 3 to 4% for the whole tumor region. These findings are in accordance with [29], where inclusion of PWI and DWI were found to be mainly advantageous for the tumor core and the whole tumor region.

Nowadays, a common approach to further improve segmentation results is to model the spatial dependency between the tissue labels of adjacent voxels, typically using a Markov Random Field (MRF) [26,36] or a Conditional Random Field (CRF) [15,17]. MRF and CRF are usually combined with a probabilistic classifier which segments all voxels independently. With MRF, a fixed penalty term is added to the cost function when adjacent voxels have different labels. This penalty term is also depending on the (dis)similarity of the feature vectors when using CRF. To the authors' knowledge, this is the first study to consider spatially regularized NMF for

tumor segmentation. We have added a spatial regularization term with sparseness to the NMF objective function. L1-regularization was used to promote piece-wise smoothness of the abundance maps, allowing for discontinuities at the tissue boundaries. Whereas MRF and CRF are directly applied to the voxel labels, we applied L1-regularization to the tissue abundance maps which represent continuous variables. In some patients, suboptimal segmentation results were found when using spatial regularization only. This is reflected in the lower Dice-scores for $\mathbf{NMF}_{\text{spatial}}$ in the active tumor region and in the tumor core region (Table 1). The original abundance maps have a relatively low degree of sparseness, such that piece-wise spatial smoothness was not always restricted to the true tissue boundaries. Results were improved by combining spatial regularization with sparseness. This kind of regularization resembles MRF more closely, since MRF implies absolute sparseness of the tissue labels. Over all patients, we found an increase in mean Dice-score by 2% for active tumor and for the tumor core when using spatial regularization with sparseness (Table 1). For the whole tumor region, an increase of 1% was found for the 25^{th} and 75^{th} percentile Dice-scores and no increase was found for the mean value. Similar improvements in segmentation results have been reported in the literature [17, 36].

As imaging characteristics vary considerably across glioma subtypes and grades, supervised classification methods require extensive training datasets with a uniform acquisition protocol. To increase specificity, separate classifiers are usually built for low- and high-grade gliomas [14]. Using training data from different patients to segment a new test case requires careful calibration of image intensities across different patients, which is non-trivial in the presence of pathology. Our method uses labelled data only from the patient under study. Feature vectors averaged over the local neighborhood of the selected voxels are assumed to be approximative prototypes of the pathological tissue regions to be segmented.

So far, most semi-automated segmentation studies did not quantify the effect of user input variability on their results. We have assessed user input variability by repeatedly selecting random voxels in each pathological region. Fairly robust segmentation results were found for most patients (see Figures 3, 4 and 5), but a

large spread in the Dice-scores and low outlier values were found in several patients as well. In some cases suboptimal results could be explained by ambiguous voxel selection near the tissue boundaries. Figure 6 illustrates a case of suboptimal voxel selection for patient 1. One of the necrotic voxels was selected at the edge of the manually segmented necrotic region (red voxel in B, other necrotic voxels were selected on other slices). This voxel is located very near the active tumor region and does not show the hypo-intense T1C signal which is characteristic for necrosis. The resulting NMF segmentation is shown in C for active tumor and in D for necrosis. The necrotic voxel selected at the boundary results in an overestimation of the necrotic region and an underestimation of the tumor region. The low outlier Dice-score for patient 1 for active tumor in Figure 3 corresponds to this voxel selection. Another patient had a bi-focal tumor, and in some runs only one of both lesions was detected as no voxels were selected in the other lesion. To obtain more consistent segmentation results, we require the user to select voxels in a smart way. Spatially distributing the selected voxels is one way of covering intra-tissue heterogeneity. As our method exploits spatial information, at least one voxel should be selected in each lesion in case of multi-focal lesions. A clinician is more likely to select voxels which are clearly located in one tissue region than voxels near diffuse tissue boundaries. Careful voxel selection is expected to further improve segmentation results and reduce variability.

Conclusion

We have presented a semi-automated brain tumor segmentation method, based on NMF with L1-regularization to promote spatial consistency and sparseness of the tissue abundance maps. The semi-automated framework was applied to an MP-MRI dataset consisting of cMRI, PWI and DWI. User-defined voxel selection is applied to initialize pathological sources of the NMF analysis, and to exploit knowledge about the spatial location of the tumor to remove false positives in a post-processing step. In this way, we aimed at incorporating prior knowledge while maintaining general applicability of our method to any type of glioma and to any MP-MRI dataset. Sensitivity to user input variability was explored through repeated analyses with different voxel selection. Robust results were found for most patients, although careful voxel selection is mandatory to avoid sub-optimal segmentation. The reported

segmentation scores are competitive with state-of-the-art in literature, even when looking at the lower quartile values across patients.

List of abbreviations

ADC: Apparent diffusion coefficient

BRATS: Multimodal Brain Tumor Segmentation challenge

CA: Cellular automata

cMRI: Conventional MRI

CRF: Conditional random field

DWI: Diffusion-weighted imaging

FCM: Fuzzy C-means clustering

FLAIR: Fluid-attenuated inversion recovery

FOV: Field of view

GBM: Glioblastoma multiforme

GNDL: Gauss-Newton with dogleg trust region

HGG: High-grade glioma

kNN: k-nearest neighbors

MP-MRI: Multi-parametric MRI

MRF: Markov random field

NMF: Non-negative matrix factorization

PWI: Perfusion-weighted imaging

rCBV: Relative cerebral blood volume

SPA: Successive projection algorithm

T1C: T1-imaging with contrast enhancement

TE: Echo time

TI: Inversion time

TR: Repetition time

Ethics approval and consent to participate

The local ethics committee of the Ghent University Hospital allowed a retrospective analysis of the patient data.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analysed during the current study are not publicly available due to them containing information that could compromise research participant privacy/consent. The data are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

NS, DMS and MA received funding from the Research Foundation Flanders (FWO), grant number G.0869.12N. JV received funding from the Henri Benedictus Fellowship of the Belgian American Educational Foundation. SVH received funding from the government agency for Innovation by Science and Technology (IWT), grant number IM 135005. SVH received funding from the European Research Council under the European Union's Seventh Framework Programme, grant number 339804. SVH and DMS received funding from the European Research Council under the European Union's Seventh Framework Programme, grant number 316679. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author's contributions

EA and MA acquired and anonymized the patient imaging data. JV provided software code and advice to pre-process the DWI data. NS and DMS designed the semi-automated segmentation framework. NMF analyses were performed by NS. The study was supervised by SVH and EA. NS, DMS, SVH, FM and UH were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Author details

[1]Department of Electrical Engineering (ESAT), STADIUS Centre for Dynamical Systems, Signal Processing and Data Analytics, KULeuven, Kasteelpark Arenberg, Leuven, Belgium. [2]Department Medical Information Technologies, iMinds, Technologiepark 19, 9052 Zwijnaarde, Belgium. [3]Department of Radiology, Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium. [4]Department of Physics, iMinds Vision Lab, University of Antwerp, Edegemsesteenweg 200-240, 2610 Antwerp, Belgium. [5]Department of Electrical Engineering (ESAT), PSI Centre for Processing Speech and Images, KULeuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium. [6]Department of Imaging and Pathology, Biomedical MRI/MoSAIC, KULeuven, Herestraat 49, 3000 Leuven, Belgium.

References

1. Ostrom, Q.T., Gittleman, H., Liao, P., Rouse, C., Chen, Y., Dowling, J., Wolinsky, Y., Kruchko, C., Barnholtz-Sloan, J.: CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. Neuro-oncology **16**(suppl 4), 1–63 (2014)
2. Cha, S.: Update on brain tumor imaging: from anatomy to physiology. American Journal of Neuroradiology **27**(3), 475–487 (2006)
3. Young, G.S.: Advanced MRI of adult brain tumors. Neurologic Clinics **25**(4), 947–973 (2007)
4. Van Cauter, S., De Keyzer, F., Sima, D.M., Sava, A.C., D'Arco, F., Veraart, J., Peeters, R.R., Leemans, A., Van Gool, S., Wilms, G., *et al.*: Integrating diffusion kurtosis imaging, dynamic susceptibility-weighted contrast-enhanced mri, and short echo time chemical shift imaging for grading gliomas. Neuro-oncology **16**(7), 1010–1021 (2014)
5. Cha, S., Knopp, E.A., Johnson, G., Wetzel, S.G., Litt, A.W., Zagzag, D.: Intracranial mass lesions: Dynamic contrast-enhanced susceptibility-weighted echo-planar perfusion MR imaging 1. Radiology **223**(1), 11–29 (2002)
6. Bulakbasi, N., Kocaoglu, M., Farzaliyev, A., Tayfun, C., Ucoz, T., Somuncu, I.: Assessment of diagnostic accuracy of perfusion mr imaging in primary and metastatic solitary malignant brain tumors. American journal of neuroradiology **26**(9), 2187–2199 (2005)

7. Hu, L.S., Eschbacher, J.M., Heiserman, J.E., Dueck, A.C., Shapiro, W.R., Liu, S., Karis, J.P., Smith, K.A., Coons, S.W., Nakaji, P., *et al.*: Reevaluating the imaging definition of tumor progression: perfusion MRI quantifies recurrent glioblastoma tumor fraction, pseudoprogression, and radiation necrosis to predict survival. Neuro-Oncology **14**(7), 919–930 (2012)
8. Price, S., Jena, R., Burnet, N., Hutchinson, P., Dean, A., Pena, A., Pickard, J., Carpenter, T., Gillard, J.: Improved delineation of glioma margins and regions of infiltration with the use of diffusion tensor imaging: an image-guided biopsy study. American Journal of Neuroradiology **27**(9), 1969–1974 (2006)
9. Wen, P.Y., Macdonald, D.R., Reardon, D.A., Cloughesy, T.F., Sorensen, A.G., Galanis, E., DeGroot, J., Wick, W., Gilbert, M.R., Lassman, A.B., *et al.*: Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. Journal of Clinical Oncology **28**(11), 1963–1972 (2010)
10. Dhermain, F.: Radiotherapy of high-grade gliomas: current standards and new concepts, innovations in imaging and radiotherapy, and new therapeutic approaches. Chin J Cancer **33**(1), 16–24 (2014)
11. Lacroix, M., Abi-Said, D., Fourney, D.R., Gokaslan, Z.L., Shi, W., DeMonte, F., Lang, F.F., McCutcheon, I.E., Hassenbusch, S.J., Holland, E., Hess, K., Michael, C., Miller, D., Sawaya, R.: A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. Journal of Neurosurgery **95**(2), 190–198 (2001)
12. Ellingson, B.M., Cloughesy, T.F., Lai, A., Nghiemphu, P.L., Mischel, P.S., Pope, W.B.: Quantitative volumetric analysis of conventional mri response in recurrent glioblastoma treated with bevacizumab. Neuro-oncology, 206 (2011)
13. Gordillo, N., Montseny, E., Sobrevilla, P.: State of the art survey on mri brain tumor segmentation. Magnetic resonance imaging **31**(8), 1426–1438 (2013)
14. Menze, B., Reyes, M., Van Leemput, K., *et al.*: The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Transactions on Medical Imaging **34**(10), 1993–2024 (2015). doi:10.1109/TMI.2014.2377694
15. Bauer, S., Fejes, T., Slotboom, J., Wiest, R., Nolte, L.-P., Reyes, M.: Segmentation of brain tumor images based on integrated hierarchical classification and regularization. In: MICCAI BraTS Workshop. Nice: Miccai Society (2012)
16. Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O., Das, T., Jena, R., Price, S.: Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012 vol. LCNS 7512, pp. 369–376. Springer, ??? (2012)
17. Lee, C.-H., Wang, S., Murtha, A., Brown, M.R., Greiner, R.: Segmenting brain tumors using pseudo–conditional random fields. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 359–366 (2008). Springer
18. Verma, R., Zacharaki, E.I., Ou, Y., Cai, H., Chawla, S., Lee, S.-K., Melhem, E.R., Wolf, R., Davatzikos, C.: Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. Academic Radiology **15**(8), 966–977 (2008)
19. Jensen, T.R., Schmainda, K.M.: Computer-aided detection of brain tumor invasion using multiparametric mri. Journal of Magnetic Resonance Imaging **30**(3), 481–489 (2009)
20. Li, S.Z.: Markov Random Field Modeling in Image Analysis. Springer, ??? (2009)
21. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proceedings of the Eighteenth International Conference on Machine Learning, ICML, vol. 1, pp. 282–289 (2001)
22. Subbanna, N.K., Precup, D., Collins, D.L., Arbel, T.: Hierarchical probabilistic gabor and mrf segmentation of brain tumours in mri volumes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 751–758 (2013). Springer
23. Fletcher-Heath, L.M., Hall, L.O., Goldgof, D.B., Murtagh, F.R.: Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artificial intelligence in medicine **21**(1), 43–63 (2001)
24. Kazerooni, A.F., Mohseni, M., Rezaei, S., Bakhshandehpour, G., Rad, H.S.: Multi-parametric (ADC/PWI/T2-w) image fusion approach for accurate semi-automatic segmentation of tumorous regions in glioblastoma multiforme. Magnetic Resonance Materials in Physics, Biology and Medicine **28**(1), 13–22 (2015)
25. Juan-Albarracín, J., Fuster-Garcia, E., Manjón, J.V., Robles, M., Aparici, F., Martí-Bonmatí, L., García-Gómez,

J.M.: Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. PloS One **10**(5), 0125143 (2015)

26. Menze, B.H., Van Leemput, K., Lashkari, D., Weber, M.-A., Ayache, N., Golland, P.: A generative model for brain tumor segmentation in multi-modal images. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010 vol. 13, pp. 151–159. Springer, ??? (2010)
27. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging **20**(1), 45–57 (2001)
28. Nie, J., Xue, Z., Liu, T., Young, G.S., Setayesh, K., Guo, L., Wong, S.T.: Automated brain tumor segmentation using spatial accuracy-weighted hidden markov random field. Computerized Medical Imaging and Graphics **33**(6), 431–441 (2009)
29. Sauwen, N., Sima, D.M., Van Cauter, S., Veraart, J., Leemans, A., Maes, F., Himmelreich, U., Van Huffel, S.: Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. NMR in Biomedicine **28**(12), 1599–1624 (2015). doi:10.1002/nbm.3413
30. Ortega-Martorell, S., Lisboa, P.J., Vellido, A., Simões, R.V., Pumarola, M., Julià-Sapé, M., Arús, C.: Convex non-negative matrix factorization for brain tumor delimitation from mrsi data. PloS one **7**(10), 47824 (2012)
31. Prastawa, M., Bullitt, E., Moon, N., Van Leemput, K., Gerig, G.: Automatic brain tumor segmentation by subject specific modification of atlas priors 1. Academic radiology **10**(12), 1341–1348 (2003)
32. Saha, B.N., Ray, N., Greiner, R., Murtha, A., Zhang, H.: Quick detection of brain tumors and edemas: A bounding box method using symmetry. Computerized medical imaging and graphics **36**(2), 95–107 (2012)
33. Görlitz, L., Menze, B.H., Weber, M.-A., Kelm, B.M., Hamprecht, F.A.: Semi-supervised tumor detection in magnetic resonance spectroscopic images using discriminative random fields. In: Pattern Recognition, pp. 224–233. Springer, ??? (2007)
34. Kwon, D., Shinohara, R.T., Akbari, H., Davatzikos, C.: Combining generative models for multifocal glioma segmentation and registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 763–770 (2014). Springer
35. Hamamci, A., Kucuk, N., Karaman, K., Engin, K., Unal, G.: Tumor-cut: segmentation of brain tumors on contrast enhanced mr images for radiosurgery applications. IEEE transactions on medical imaging **31**(3), 790–804 (2012)
36. Havaei, M., Jodoin, P.-M., Larochelle, H.: Efficient interactive brain tumor segmentation as within-brain knn classification. In: ICPR, pp. 556–561 (2014)
37. Boxerman, J., Schmainda, K., Weisskoff, R.: Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. American Journal of Neuroradiology **27**(4), 859–867 (2006)
38. Veraart, J., Sijbers, J., Sunaert, S., Leemans, A., Jeurissen, B.: Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. NeuroImage **81**, 335–346 (2013)
39. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging **16**(2), 187–198 (1997)
40. Wang, Y.-X., Zhang, Y.-J.: Nonnegative matrix factorization: A comprehensive review. Knowledge and Data Engineering, IEEE Transactions on **25**(6), 1336–1353 (2013)
41. Sorber, L., Van Barel, M., De Lathauwer, L.: Structured data fusion. IEEE Journal of Selected Topics in Signal Processing **9**(4), 586–600 (2015)
42. Vervliet, N., Debals, O., Sorber, L., Van Barel, M., De Lathauwer, L.: Tensorlab v3.0 (March 2016). `http://www.tensorlab.net`. [Online; accessed 24-June-2016]
43. Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd Edition. Springer, ??? (2006)
44. Araújo, M.C.U., Saldanha, T.C.B., Galvão, R.K.H., Yoneyama, T., Chame, H.C., Visani, V.: The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. Chemometrics and Intelligent Laboratory Systems **57**(2), 65–73 (2001)
45. Gillis, N.: Successive nonnegative projection algorithm for robust nonnegative blind source separation. SIAM Journal on Imaging Sciences **7**(2), 1420–1450 (2014)
46. Krawczyk, B.: Learning from imbalanced data: open challenges and future directions. Progress in Artificial Intelligence, 1–12 (2016)

47. Hsieh, T.M., Liu, Y.-M., Liao, C.-C., Xiao, F., Chiang, I.-J., Wong, J.-M.: Automatic segmentation of meningioma from non-contrasted brain mri integrating fuzzy clustering and region growing. BMC medical informatics and decision making **11**(1), 1 (2011)
48. Cordier, N., Menze, B., Delingette, H., Ayache, N.: Patch-based segmentation of brain tissues. In: MICCAI Challenge on Multimodal Brain Tumor Segmentation, pp. 6–17 (2013). IEEE
49. Bauer, S., Wiest, R., Nolte, L.-P., Reyes, M.: A survey of MRI-based medical image analysis for brain tumor studies. Physics in Medicine and Biology **58**(13), 97 (2013)
50. Di Costanzo, A., Scarabino, T., Trojsi, F., Giannatempo, G.M., Popolizio, T., Catapano, D., Bonavita, S., Maggialetti, N., Tosetti, M., Salvolini, U., d'Angelo, V., Tedeschi, G.: Multiparametric 3T MR approach to the assessment of cerebral gliomas: tumor extent and malignancy. Neuroradiology **48**(9), 622–631 (2006)

Figures

Figure 1 Illustration of morphological post-processing after initial semi-automated NMF based segmentation of necrosis (A) and active tumor (C). Step 1: false positives are removed by withholding only the connected components closest to the user-defined seeding points (marked by an 'x' in A and C). Step 2: spatial adjacency of the connected components in the preliminary necrosis mask (green) to the withheld active tumor mask (red) is verified in E. The final necrosis mask is shown in yellow in F.

Figure 2 First row: coregistered MP-MRI maps of a GBM (patient 3), left to right: T1C, FLAIR, rCBV, ADC. Second row, left to right: NMF abundance maps for active tumor, necrosis and edema. The final segmentation masks are shown on the right for active tumor (red), necrosis (yellow) and edema (blue).

Figure 3 Boxplots showing the dispersion of the Dice-scores for active tumor. Boxplots show quartile ranges of the Dice-scores, '+' indicates outliers.

Figure 4 Boxplots showing the dispersion of the Dice-scores for the tumor core. Boxplots show quartile ranges of the Dice-scores, '+' indicates outliers.

Figure 5 Boxplots showing the dispersion of the Dice-scores for the whole tumor region. Boxplots show quartile ranges of the Dice-scores, '+' indicates outliers.

Figure 6 A close-up of a GBM lesion on an axial T1C slice (A). The manually segmented necrotic region is shown in yellow (B), the selected necrotic voxel is marked in red. Segmentation of the active tumor region (C) and necrotic region (D) on several slices, blue indicates NMF segmentation, green indicates manual segmentation and cyan indicates overlap.