Objective Tumour Heterogeneity Determination in Gliomas

Dirk Simon[a], Jan Klein[b], Jan Rexilius[b], Bram Stieltjes[a]

[a]German Cancer Research Center, Department of Radiology, im Neuenheimer Feld, Heidelberg, Germany;
[b]MeVis Research, Universitaetsallee, Bremen, Germany

ABSTRACT

Diffusion weighted imaging (DWI) derived apparent diffusion coefficient (ADC) values are known to correlate inversely to tumour cellularity in brain tumours. The average ADC value increases after successful chemotherapy, radiotherapy or a combination of both and can be therewith used as a surrogate marker for treatment response. Moreover, high and low malignant areas can be distinguished. The main purpose of our project was to develop a software platform that enables the automated delineation and ADC quantification of different tumour sections in a fast, objective, user independent manner. Moreover, the software platform allows for an analysis of the probability density of the ADC in high and low malignant areas in ROIs drawn on conventional imaging to create a ground truth. We tested an Expectation Maximization algorithm with a Gaussian mixture model to objectively determine tumour heterogeneity in gliomas because of yielding Gaussian distributions in the different areas. Furthermore, the algorithm was initialized by seed points in the areas of the gross tumour volume and the data indicated that an automatic initialization should be possible. Thus automated clustering of high and low malignant areas and subsequent ADC determination within these areas is possible yielding reproducible ADC measurements within heterogeneous gliomas.

Keywords: Diffusion Weighted Imaging, Pattern Recognition, EM-Clustering, Gaussian Mixture Model

1. INTRODUCTION

Diffusion weighted magnetic resonance imaging[1,2] is a method of microanatomical imaging and was introduced into clinical practice in the 1990th. It measures the Brownian molecular motion of water. Two principal kinds of diffusion are being distinguished: isotropic diffusion and anisotropic diffusion. The former describes equal diffusion in all directions in free diffusing water. Biologic tissue that consists of fibers with coherent directions, show anisotropic diffusion, here diffusion is most prominent in the fiber direction.

Diffusion Weighted Imaging (DWI) describes the qualitative aspect of an examination[2] but a number of quantification steps can be based on diffusion measurement. Two basic measurement methods are being distinguished. In the first case, scalar values are saved into a so called ADC map. This map entails the averaged diffusion coefficents within a specific voxel. In the second case, at least 6 values are being determined which are combined in a diffusion tensor as an orthonormal basis of eigenvectors. This work is only considering the first case. Thus, only scalar grey values are being used.

It has been depicted that the mean ADC is correlated inversely with the cellularity in brain tumours.[3] High cellularity of the tissue correlates with tumour growth. It has been shown that the mean ADC increases after successful chemotherapy, radiotherapy or a combination of both. Thus, the mean ADC can be seen as a surrogate treatment marker. A low mean ADC signalizes that the tumour growth is high, whereas a high mean value indicates that the tumour growth is stagnating or low. However, the mean ADC is also increased by water retention, caused by oedema or necrosis. Also, the ADC can be used to separate high and low malignant tumour areas. Such a separation is important for resections, biopsies and radiation therapy.[4] It has been indicated that a separation between heterogeneous tumour regions through the evaluation of ADC maps is possible.[5,6] The first source deals with a partition between cystic, necrotic and solid tissue, whereas a differentiation between progressive and non progressive tumour could not be determined.[6] Usually, ADC values are extracted by means of

manual region of interest (ROI) drawing, causing a strong user dependency. Furthermore, the ROI-quantification is impeded by several factors such as poor data resolution and little contrast between high and low malignant areas. Thus, for the clinical use of DWI in brain tumour diagnosis and treatment planning, robust determination of the ADC in the different tumour areas is mandatory.

An automatic segmentation of brain tumours based on an Expectation Maximization algorithm is described.[7] Therein, the difficulty of segmenting brain tumours is listed. Six aspects are being mentioned as general characteristics of brain tumours:

1. vary greatly in size and position
2. vary greatly in the way they appear in MRI
3. may have overlapping intensities with normal tissue
4. may be space occupying or infiltrating
5. may enhance fully, partially, or not at all, with contrast agent
6. may be accompanied by surrounding edema

Additionally,[7] points out the following issues in the special case of gliomas:

1. ragged boundaries
2. initially only in white matter, possibly later spreading outside white matter
3. margins enhance with contrast agent, inside does not
4. accompanied by edema
5. infiltrating at first, possibly becoming space occupying when larger

In[8] another approach of segmenting brain tumours is explained. In this work, supervised Support Vector Machines (SVM) are being compared with one class unsupervised SVMs. It can be clearly seen that the one-class unsupervised SVM outperforms the supervised SVM in terms of segmentation accuracy, speed and reduced human intervention.

Our approach is also using an unsupervised machine-learning technique. Instead of unsupervised SVMs,[8] we are using the probabilistic soft clustering algorithm 'Expectation Maximization'[7] on ADC maps and show the applicability on these maps to determine heterogenous areals within gliomas. First, we show that the distribution of the ADC within the different tumour areas and the surrounding tissue principally allows for a clustering based on the ADC. Moreover, we could prove through the variation of the seed points that are needed for the initializing of the algorithm, that seed point positioning does not interfere with the clustering result.

Furthermore, we are coping with the mentioned problems of 'overlapping intensities with normal tissue' and 'ragged boundaries'.[7] In the process, we are also considering the fact of 'overlapping intensities of low malignant tissue with cerebrospinal fluid'. Our approach enables objective tumour heterogeneity determination with a minimum of human interaction.

2. METHODS

We included 6 patients with high grade gliomas. We used a 1.5 T and 3.0 T scanner, routine imaging and DTI was performed. DTI parameters: SSEPI, TR/TE 4700/78, FOV 240 mm 2.5 mm^3, 6 gradient directions, two b-values (0, 1000), 10 averages. The platform was developed using MeVisLab. Firstly, the software platform automatically determines the input data type and calculates the ADC map (figure 2), appropriately. Secondly, we registered the T1-image and the T2-image, respectively (moving image) onto the ADC map (fixed image) using normalized mutual information and an automatic linear registration algorithm. We evaluated the ADC distribution in different tissue types. For ADC distribution analysis, a physician drew conservative ROIs indicating low malignant (on T2) and high malignant areas (on T1) (Section 2.1). Next, we used an Expectation Maximization (EM) algorithm that applies a Gaussian mixture model for classification[9] (Section 2.2). For initialization, a seed ROI was set in both areas. Subsequently, the two areas were automatically clustered (figure 2). The clustering result was compared to the manually drawn ROIs.

2.1 Determining ADC

For automatic determination of the ADC map, we used linear least squares calculation (equation 1).

$$ln(S) = ln(S_0) - D \cdot b \Leftrightarrow \tag{1}$$

$$\hat{\theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (X^T \cdot X)^{-1} X^T \cdot (ln(S) - ln(S_0)) \Leftrightarrow \tag{2}$$

$$\hat{\theta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = X^+ \cdot (ln(S) - ln(S_0)). \tag{3}$$

The matrix X comprises the b-value and gradient directions. In the case of other input data, the ADC map can be determined, accordingly. The most important calculation schemes are *orthogonal* in the directions *read, phase, slice* and *trace*, respectively. The orthogonal ADC map $< ADC >$ is being calculated by equation 4.

$$< ADC >= \frac{ADC_x + ADC_y + ADC_z}{3} \tag{4}$$

$$where, ADC = \frac{-1}{b} \cdot ln\left(\frac{S}{S_0}\right). \Leftrightarrow \tag{5}$$

$$< ADC >= \frac{1}{3} \cdot \frac{-1}{b} \cdot \left(ln\left(\frac{S_x}{S_0}\right) + ln\left(\frac{S_y}{S_0}\right) + ln\left(\frac{S_z}{S_0}\right)\right). \tag{6}$$

The trace-image can be derived from equation 4. The trace calculation itself is shown in equation 7.

$$< ADC >= \frac{-1}{3} \cdot \frac{ln\left(\frac{S_x}{S_0}\right) + ln\left(\frac{S_y}{S_0}\right) + ln\left(\frac{S_z}{S_0}\right)}{b} \Leftrightarrow \tag{7}$$

$$< ADC >= \frac{-1}{3} \cdot \frac{ln\left(S_x \cdot S_y \cdot S_z\right) - 3 \cdot ln\left(S_0\right)}{b} \Leftrightarrow \tag{8}$$

$$< ADC >= -\frac{ln\left(S_x \cdot S_y \cdot S_z\right)^{\frac{1}{3}} - ln(S_0)}{b} \tag{9}$$

$$where, \quad S = (S_x \cdot S_y \cdot S_z)^{\frac{1}{3}} \quad denotes \quad the \quad trace \tag{10}$$

$$hence, < ADC >= -\frac{ln\left(\frac{S}{S_0}\right)}{b}. \tag{11}$$

If we have several trace weighted images, we can determine the ADC by linear regression (equation 12).

$$\overrightarrow{y} = \begin{pmatrix} ln(s_1) \\ ln(s_2) \\ \vdots \\ ln(s_n) \end{pmatrix} = \begin{pmatrix} 1 & b_1 \\ 1 & b_2 \\ \vdots & \vdots \\ 1 & b_n \end{pmatrix} \cdot \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \Leftrightarrow \tag{12}$$

$$\hat{\theta} = ((X^T \cdot X)^{-1}) \cdot X^T \cdot \overrightarrow{y} \Leftrightarrow \tag{13}$$

$$\hat{\theta} = X^+ \cdot \overrightarrow{y}. \tag{14}$$

2.2 The Expectation Maximization Algorithm initialized by seedpoints

K-Means and Expectation Maximization using a probabilistic model An Expectation Maximization Algorithm[9] is used as an unsupervised learning algorithm to discover similar examples from given input vectors, to be precise that structures are found in a set of unlabeled data. It is related to the hard clustering technique 'K-Means' und comprises the following steps: 'Expectation' and 'Maximization'. As a matter of principle, in the 'Expectation'-Step, the members are being assigned to the cluster. In 'K-Means', we speak of a hard assignment. Specifically, the so called 'indicator-function' is being minimized to achieve new cluster sets. Here, real partitions are being made so no overlap will occur (partitioned clustering). In contradiction to that, the EM-Algorithm uses overlapping and a soft assignment is performed. Instead of an indicator function, we use probabilities to determine the membership. These are the responsibilities or the posterior probabilities. In the Maximization-Step, these parameters are being used to recalculate the parameters of the model. In the simple case of 'K-Means', we only have the model parameter μ, whereas in the Expectation Maximization case, the model consists of a mixing coefficient π_k which comprises the probability of the cluster belonging to the whole model, μ_k which determines the mean value of the cluster model and Σ_k specifies the covariance-matrix if X, the sample, denotes a vector.

Exact steps of the Expectation Maximization Algorithm with probabilistic model

- Expectation-Step: In the Expectation-Step the cluster probabilities are being determined, i.e. γ_k is being calculated by equation 15.

$$\gamma_k(x) \equiv p(k|x) \Leftrightarrow \frac{p(k)p(x|k)}{\sum_{k=1}^{K} p(k)p(x|k)} \Leftrightarrow \frac{p(k)p(x|k)}{\sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)}. \tag{15}$$

Here, the Bayes Theorem is being used to calculate the cluster probabilities. $\sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)$ comprises the total probability and represents the whole Gaussian mixture model.

- Maximization-Step: The Maximization-Step uses the posterior probability to recalculate the model parameters (π_k, Σ_k, μ_k).

$$\mu_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma_k(x) x_n \tag{16}$$

$$\Sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma_k(x)(x_n - \mu_k^{new})(x_n - \mu_k^{new})^T \tag{17}$$

$$\pi_k^{new} = \frac{N_k}{N} \tag{18}$$

,where $N_k = \sum_{n=1}^{N} \gamma_k(x)$.

Now, we do have the following Log likelihood equation to maximize:

$$ln(p(X|\mu, \Sigma, \pi)) = \sum_{n=1}^{N} ln \left\{ \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k) \right\}. \tag{19}$$
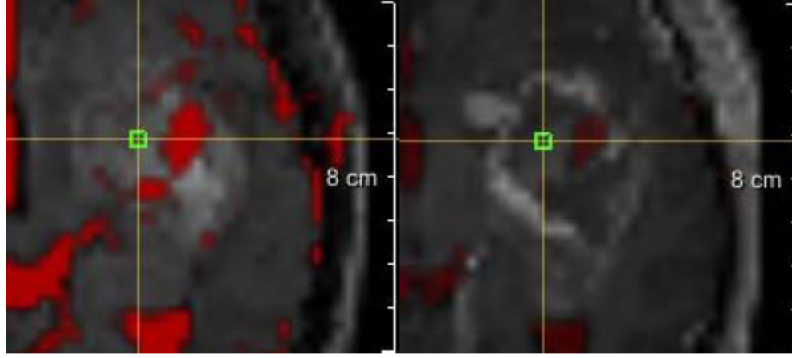
Figure 1. (i): T2-FLAIR overlayed with ADC map in red (crosshair in high malignant area), (ii): T1 overlayed with ADC map
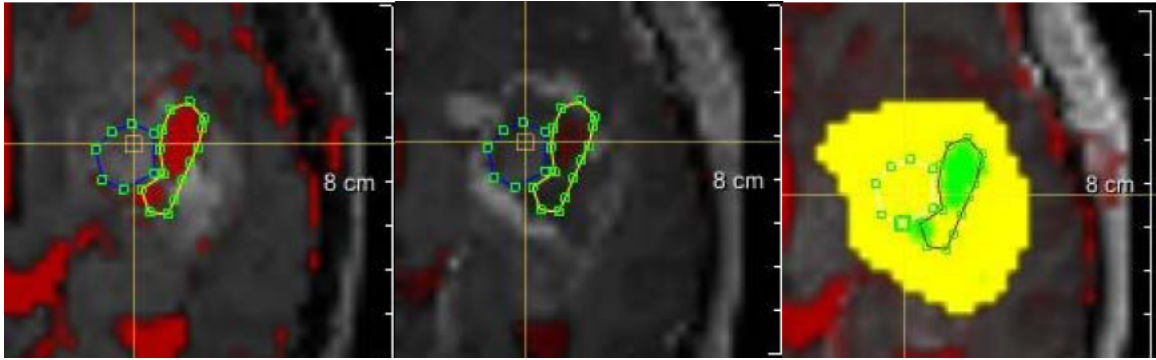

Figure 2. (i): Conservative ROI drawn by the physician: T2-FLAIR overlaid with ADC map in red, (ii) Conservative ROI drawn by the physician: T1 overlayed with ADC map in red. The blue circle indicates the conservative high malignant area, the yellow circle indicating the low malignant area, (iii) Clustered area: yellow indicates the posterior probability (γ_k) for the high malignant area, green the low malignant area. Higher values correspond to a higher probability, whereas lower values indicate lower probability.

3. RESULTS

3.1 Gaussian distributions in clustered area

The manual ROI-placement yielded Gaussian ADC distributions in the different tumour areas and healthy tissue. Moreover, peaks are distinguishable by their means and a clustering algorithm with Gaussian distributions could be applied (figure 3). Eventhough, a full range χ^2-test was not significant, a limited Gaussian fitting excluding normal tissue and necrosis (high malignant) (24074, 38786) $\chi^2(0.95; 50 - 3 = 47) = 64.0 > 7.39$ reaches significance. For low malignant tumour the limited χ^2-test (44890, 58350) results in $\chi^2(0.95; 100 - 3 = 97) = 119.87 > 51.15$. For 6 patients the limited test for high malignant tumour (28720, 41300) (figure 3) shows $\chi^2(0.95; 50 - 3 = 47) = 64.0 > 3.79$. Low malignant tumour (46670, 60150) leads to $\chi^2(0.95; 100 - 3 = 97) = 119.87 > 51.48$. Thus, a Gaussian mixture model with two Gaussians can be applied.

The results from the automated clustering compare well to the hand-drawn ROIs proving the validity of the clustering (Fig. 2). Furthermore, the objectiveness of the clustering results could be proven by an automatic test and a variation of seed points (see Section 3.3).

3.2 Testing the objectivity of cluster seedpoints

To prove the objectivity of the clusters in dependency of the drawn seed points, we implemented a method to automatically draw the seed points and then compare the mean ADC of the clustered area. Firstly, we determined the center of gravity of the high malignant and low malignant area, respectively, drawn in advance by the physician. Furthermore, new seed points were being generated with a specific radius in concentric circles
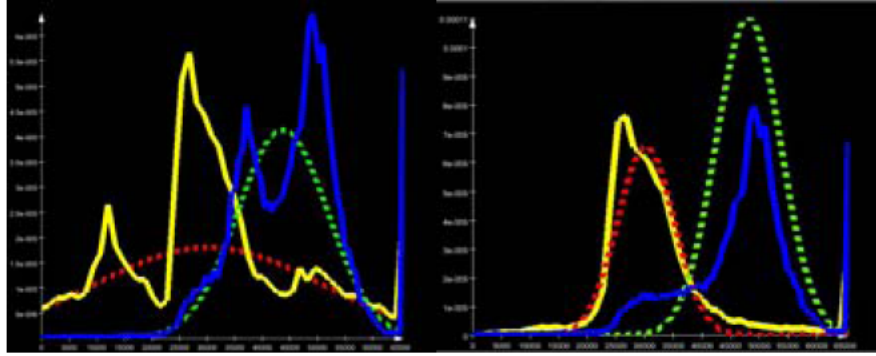
Figure 3. (i) Averaged over 6 patients (high malignant) (μ, σ) : (30275.9, 20180.8) and 5 patients (low malignant) (μ, σ) (43675.3, 8162.65) with Gaussian fits (dotted).(ii) Averaged over 4 patients with Gaussian fits (dotted) high malignant: (μ, σ) : (30181.8, 4898.42), low malignant: (μ, σ): (48059.5, 5623.49) 3D histograms containing 10 slices for each patient. Yellow indicating high malignant and blue indicating low malignant tumour. Two outliers show the variabilty that leads to difficulties in clustering border zones (see discussion)
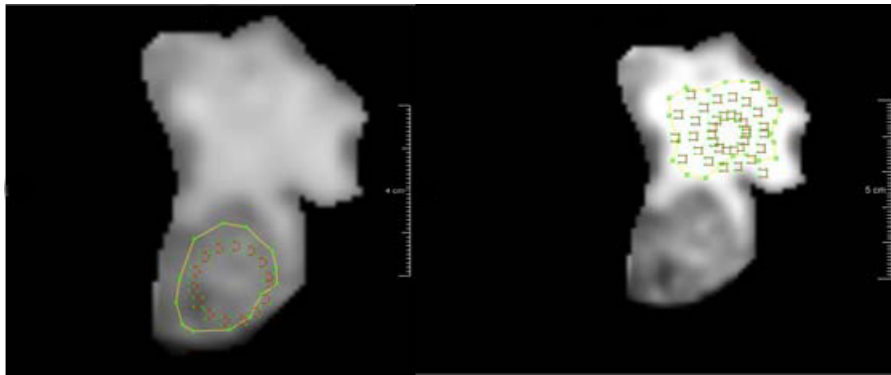

Figure 4. (i) conservative high malignant area with all generated sample seed points, (ii) conservative low malignant area with all generated sample seed points

with a specific degree value until the border of the bounding box was reached (figure 4). For the evaluation of seed point dependency, the standard deviation and mean ADC of the masked, clustered areas were determined.

Steps to be performed (figure 5)

- Load GTV * to determine the maximum area of the clustering (manually done)

- Choose degree interval of different concentric circle CSO † (manually done)

- Choose radius step of the concentric circles (manually done)

- Choose start slice where the physician indicated conservative high malignant and low malignant area (manually done)

- For all slices do (automatically done)

  – Generate ROI High Malignant and cluster region

  – Generate ROI Low Malignant and cluster region

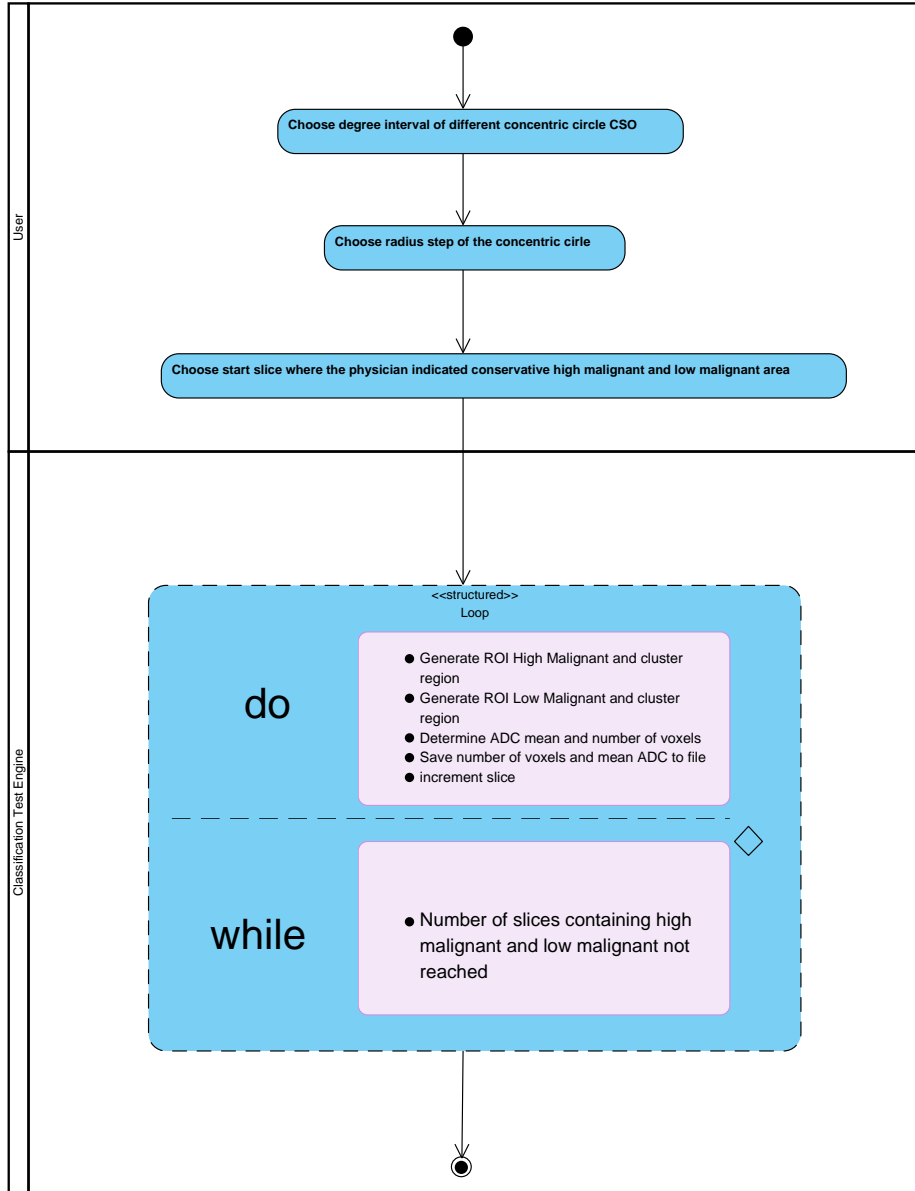---

*Gross Tumour Volume

†Contour Segmentation Object

Figure 5. Algorithm to show that initialization seeds do not interfere clustering result

- Determine ADC mean and number of voxels
- Save number of voxels and mean ADC to file

We found no changes in mean ADC in dependency of the chosen seedpoint. Also, no changes in the obtained cluster size were found. This shows that the definition of the initialization seeds does not interfere the clustering result.

3.3 Refining clustering result

In Section 1 some general problems concerning gliomas where mentioned.

- overlapping intensities with normal tissue

Figure 6. (i) GTV after interval thresholding, (ii) Clustering after interval thresholding, (iii) Clustering before interval thresholding Clustered area: green indicates the posterior probability (γ_k) for the high malignant area, yellow the low malignant area. Higher values correspond to a higher probability, whereas lower values indicate lower probability
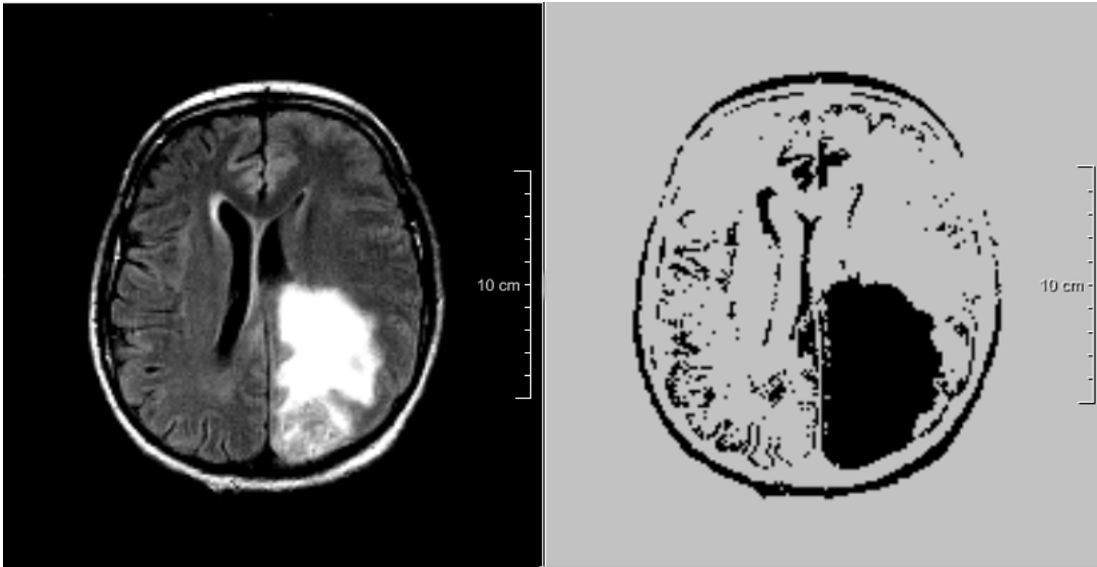

Figure 7. (i) T2-FLAIR-Image, (ii) Binary T2-FLAIR mask

- ragged boundaries

- overlapping intensities of low malignant tissue with cerebrospinal fluid

For the first two of the above mentioned isues, we applied an interval threshold on the GTV (figure 6). However, the clustering result (figure 6) is still not satisfying and we are considering using more sophisticated techniques.

For the last issue we used the T2-FLAIR image (figure 7) where cerebrospinal fluid appears hypointense. Firstly, we generated a binary mask (figure 7) so that only CSF was visible and secondly, we applied the mask onto the ADC map (figure 8). As we can clearly see, the cerebrospinal fluid is completely suppressed and the overlap between CSF and low malignant tumour areas is excluded.

4. CONCLUSIONS

In this work, we showed that tumour heterogeneity can be delineated in an objective way with minimal interference of the physician. Thus, the separation of different tumour areas, which is important for resections, biopsies
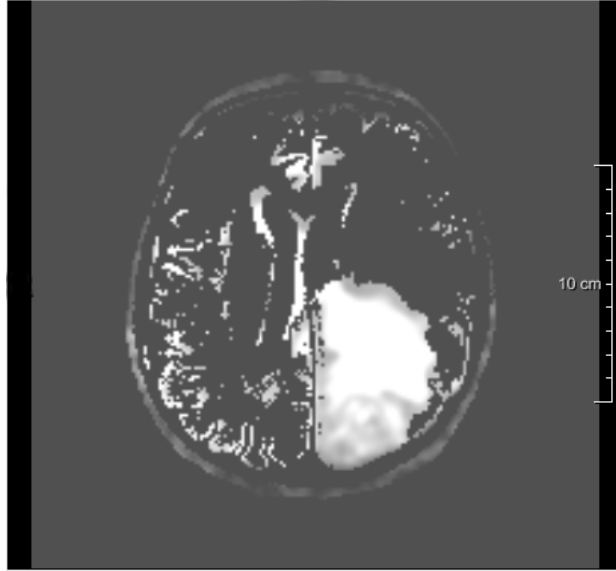
Figure 8. Applied binary mask onto ADC map

and radiation therapy,[4] can be performed in a fast, nearly user independent manner. Also, our data indicate that an automatic initialization should be possible. However, there are challenges considering the boundaries of normal tissue near low malignant tissue. Due to comparable ADC values and partial volume effects, normal tissue may be marked as high malignant tumour, whereas CSF and necrosis (spot in high malignant area) may be marked as low malignant tumour. The last issue has been successfully solved by using a binary FLAIR mask (Section 3.3). However, the high malignant vs. normal tissue problem requires a more complex approach using pre-existing knowledge such as Markov Random Fields.[10,11,12] Furthermore, we could also implement penalties so that impossible combinations are prevented.[13] In the tumour the following combinations are unlikely.

- high malignant tumour surrounded by low malignant tumour or normal tissue

- normal tissue around high malignant tumour

- high malignant tumour around low malignant tumour

An additional aspect could be the usage of other derived quantification parameters such as the Fractional Anisotropy (FA) to achieve a higher accuracy considering these issues.

REFERENCES

[1] Mori, S., [*Introduction to Diffusion Tensor Imaging*], Elsevier (2007).

[2] Tofts, P., [*Quantitive MRI of the Brain, Measuring changes caused by disease*], Wiley (2003).

[3] Chenevert, T. L., Stegman, L. D., Taylor, J. M., Robertson, P. L., Greenberg, H. S., Rehemtulla, A., and Ross, B. D., "Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors.," *J Natl Cancer Inst* **92**, 2029–2036 (December 2000).

[4] Stadnik, e. a., "Diffusion-weighted mr imaging of intracerebral masses: Comparison with conventional mr imaging and histologic findings.," *AJNR* **22**, 969–976 (May 2001).

[5] Brunberg, e. a., "In Vivo MR Determination of Water Diffusion Coefficients and Diffusion Anisotropy: Correlation with Structural Alteration in Gliomas of the Cerebral Hemnispheres," *American Society of Neuroradiology* **16**, 361–71 (February 1995).

[6] Robert D. Tien, e. a., "MR Imaging of High-Grade Cerebral Gliomas: Value of Diffusion-Weighted Echo-planar Pulse Sequences," *American Roentgen Ray Society* **162**, 671–677 (March 1994).

[7] Marcel Prastawa, e. a., "Automatic brain tumour segmentation by subject specific modification of atlas priors," **10**(12), 1341–1348 (2003).

[8] Jianguo Zhang, e. a., "Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine," (2004).

[9] Dempster, e. a., "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society* **11**(4), 1–38 (1977).

[10] Perez, P., "Markov Random Fields and Images," **11**(4), 413–437 (1998).

[11] Ross Kindermann, L. S., [*Markov Random Fields and Their Applications*], American Mathematical Society (2000).

[12] Li, S. Z., [*Markov Random Fields, Modeling in Computer Vision*], Springer (1995).

[13] Koen van Leemput, e. a., "Automated model-based tissue classification of mr images of the brain," *IEEE Transactions On Medical Imaging* **18**, 897–908 (October 1999).