

# Iterative Multilevel MRF Leveraging Context and Voxel Information for Brain Tumour Segmentation in MRI

Nagesh Subbanna  
McGill University, Canada  
nagesh@cim.mcgill.ca

Doina Precup  
McGill University, Canada  
dprecup@cs.mcgill.ca

Tal Arbel  
McGill University, Canada  
arbel@cim.mcgill.ca

## Abstract

*In this paper, we introduce a fully automated multi-stage graphical probabilistic framework to segment brain tumours from multimodal Magnetic Resonance Images (MRIs) acquired from real patients. An initial Bayesian tumour classification based on Gabor texture features permits subsequent computations to be focused on areas where the probability of tumour is deemed high. An iterative, multi-stage Markov Random Field (MRF) framework is then devised to classify the various tumour subclasses (e.g. edema, solid tumour, enhancing tumour and necrotic core). Specifically, an adapted, voxel-based MRF provides tumour candidates to a higher level, regional MRF, which then leverages both contextual texture information and relative spatial consistency of the tumour subclass positions to provide updated regional information down to the voxel-based MRF for further local refinement. The two stages iterate until convergence. Experiments are performed on publicly available, patient brain tumour images from the MICCAI 2012 [11] and 2013 [12] Brain Tumour Segmentation Challenges. The results demonstrate that the proposed method achieves the top performance in the segmentation of tumour cores and enhancing tumours, and performs comparably to the winners in other tumour categories.*

## 1. Introduction

The segmentation of pathology from healthy tissue in medical images is particularly challenging given the wide variability in shape, size, position and texture over a population of patient images. Figure 1 shows several MRI images of brain tumours, exhibiting large variation in appearance. Furthermore, brain tumours tend to have different boundaries in the different contrasts [1]. The problem is compounded by image acquisition biases, magnetization inhomogeneities, normalization errors, incorrect registration, etc. Consequently, significant challenges arise when adopting traditional machine learning techniques to model and

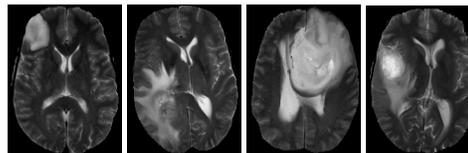


Figure 1. T2 MRI depicting tumours in different patient brain images. Note that the tumour size, position, texture and appearance are very different in different patient images.

learn these variabilities.

In the case of brain tumour segmentation, several techniques exist to separate the tumour from healthy tissue, such as locating outliers of the registration of healthy atlases to the tumour [2], or learning textural patterns that are common to tumours [3]. An important and even more challenging task is to segment and classify the various components of the pathology, such as the subclasses of a brain tumour (edema and tumour core, which includes solid tumour, enhancing tumour and necrotic core). Some techniques exist to distinguish the different tissues based on local intensity and texture features [4, 5, 6]. However, in the context of medical pathology, local features can be noisy and considering them in isolation can lead to multiple false positive hypotheses. Furthermore, local features are very vulnerable to the intensity normalization stage (especially given the variability of scanners and image acquisition parameters). As a result, too much dependence on local intensity values alone is risky. Higher level textures provide good coarse information, but techniques that are based on textures alone are less able to capture accurate boundary information.

In this paper, we develop a new, iterative, multi-stage graphical model, in order to leverage both local, voxel-level observations and class information, along with contextual regional information, based on textures and relative subclass spatial positioning. Initially, a Bayesian classification of the main tumour area is derived based on Gabor, multi-scale texture information. The classifier is designed so as to separate the tumour regions from the healthy tissues, without too much emphasis on the different types of tumour tis-

sue. Further computation is focused on the area of high tumour probability. An iterative, multi-stage Markov Random Field (MRF) framework is used to classify the tumour subclasses. At the voxel level, we utilize an adapted MRF which uses both local information, as well as neighbouring class and intensity features. This leads to over-segmentation and numerous false positive tumour subclass regions. A higher level MRF leverages both contextual texture information as well as the relative spatial consistency of the tumour subclass positions, in order to improve the classification. At this level, each node represents a subclass region hypothesis and the graphical model takes the form of an irregular lattice. Once computed, the higher level, regional information is then passed back down to the voxel-based MRF for further refinement and the two stages iterate until convergence. The main advantage of this iterative model is that the non-lattice region-level MRF corrects the local mischaracterizations of tissue types that typically occur due to noise, inhomogeneities and normalization problems. It also adds spatial and textural coherence to the segmentation. An overview of the technique is shown in Fig. 2.

We provide an empirical evaluation of our approach on publicly available patient brain tumour images from the MICCAI 2012 [11] and 2013 [12] Brain Tumour Segmentation Challenges, comparing it to the top performing techniques from these challenges. The results demonstrate that the proposed method performs best in the segmentation of tumour cores and enhancing tumours, and comparably to the past winners in the other tumour categories.

## 2. Methodology

In the case of pathology in medical images, there are often several contrast images of the patient available. For example, for patients with brain tumours, there are often MRI volumes with different contrasts (T1, T2, etc). Hence, every voxel has a corresponding  $N$ -dimensional vector, consisting of its intensity in the different contrasts. The goal is to segment various subclasses of the tumour, in our case tumour core, edema, enhancing tumour, and necrotic core. We now describe in detail the different stages of our model.

### 2.1. Customised Gabor features

Each contrast is processed using multi-window, 2D discrete Gabor transforms of the form suggested in [7]. For each slice of MRI, we have a set of  $R$  window functions whose template  $g_r[x, y; a, b, n_1, n_2, m_1, m_2, \sigma_{x_r}, \sigma_{y_r}]$  is given by:

$$e^{-\left(\frac{(x-n_1a)^2}{\sigma_{x_r}^2} + \frac{(y-n_2a)^2}{\sigma_{y_r}^2}\right)} e^{-j2\pi\frac{(m_1bx+m_2by)}{L}}, \quad (1)$$

where  $L$  is the total number of voxels in the slice in the  $X$  and  $Y$  directions,  $x$  and  $y$  are spatial coordinates within the

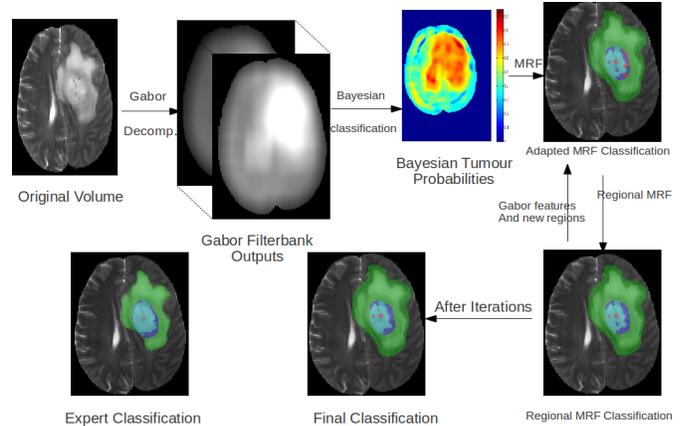


Figure 2. Flowchart depicting the various stages of the classification technique. In the MRF classification and expert labels, red represents necrotic core, green represents edema, dark blue represents solid tumour and light blue represents enhancing tumour. The Bayesian classifier produces probabilities for tumour vs. non-tumour at every point. The adapted voxel-based MRF is performed next, providing an initial segmentation of the different region classes of the tumour, based on grouping voxel labels. At the next higher level, contextual regional information is used to penalize unlikely region and spatial transitions between regional classes (e.g., some of the red area islands in the dark blue area disappear). Finally, after several iterations between the different MRF stages, the algorithm converges to a more likely set of labels at every voxel. Note the correction of the dark blue island that lies in the middle of the main light blue region.

slice,  $a$  and  $b$  are the magnitudes of the shifts in the spatial and frequency domains respectively,  $n_{1,2}$  and  $m_{1,2}$  are the indices of the shifts in the position and frequency domains respectively, and  $\sigma_{x_r}$  and  $\sigma_{y_r}$  are variance parameters of the  $r$ -th window. Let  $\mathbf{G}$  be the Gabor matrix whose columns are generated by picking all possible shift values for both  $a$  and  $b$  for all the  $R$  windows, with every  $x$  and  $y$  represented in each column. The filter bank coefficients  $\mathbf{c}$  are obtained by convolving each contrast volume slice by slice with  $\mathbf{G}$ . We use the same  $\mathbf{G}$  matrix for all contrasts.

**Training:** Each voxel in the training volumes is categorized for the purpose of this stage into one of two classes: tumour or healthy tissue. We fix the remapping window, and vary the analysis window  $g_r[\cdot]$  over the parameters  $(\sigma_{x_r}$  and  $\sigma_{y_r})$ , aiming to maximize the distance between healthy tissues and tumours. More formally, let  $\{\mathbf{f}_t\}$  and  $\{\mathbf{f}_h\}$  be the sets of voxels belonging to the tumour class and the healthy class respectively. The corresponding tumour coefficients  $\mathbf{c}_t$  in the combined space are obtained by a convolution of the Gabor filters centred at the tumour voxels. Similarly, the  $\mathbf{c}_h$  are obtained with Gabor filters centred at the healthy voxels. Ideally, the coefficients of the tumour and healthy class should be as different as possible. To achieve this goal,

the following optimization problem needs to be solved:

$$\arg \max_{\sigma_x, \sigma_y} \sum_{j,k} |c_j - c_k| \forall c_j \in \{\mathbf{c}_t\}, \forall c_k \in \{\mathbf{c}_h\} \quad (2)$$

where  $\sigma_x$  and  $\sigma_y$  are the vectors containing the values of  $\sigma_{x_r}$  and  $\sigma_{y_r}, \forall r = 1, \dots, R$ . We solve this optimization using graph cuts [9].

**Classification:** Each test volume is decomposed into its multi-window Gabor filter bank output,  $\mathbf{I}^G$ , using convolution at each voxel as described above. The class of each voxel  $i$ ,  $C_i$ , is then estimated using Bayesian classification:

$$P(C_i | \mathbf{I}_i^G) \propto P(\mathbf{I}_i^G | C_i) P(C_i), \quad (3)$$

where  $\mathbf{I}_i^G$  is the set of Gabor coefficients for voxel  $i$ . Through this approach, we obtain an initial estimate of the probability of the class of each voxel, given the texture features in the Gabor space.

## 2.2. Adapted Markov Random Field

The main purpose of this stage is to distinguish the different sub-types of tumour tissue and refine the boundaries of the tumour. The MRF is designed specifically to model probabilistically the differences in intensity between a voxel and its neighbours, in order to preserve the correct tumour boundaries. It uses larger clique sizes than in standard models, which typically use only pairs of voxels. This is due to the need to account for context information.

More precisely, let  $C_i$  be the class label of voxel  $i$ , which can have one of  $M$  possible values. We denote by  $\mathbf{I}_i$  the vector of intensities recorded at voxel  $i$  in the different contrasts. Let  $N_i$  be the set of all voxels neighbouring  $i$  in the different cliques in which  $i$  participates. Let  $\Delta \mathbf{I}_{N_i}$  denote the vector of differences in intensity between voxel  $i$  and its neighbouring voxels from  $N_i$ , and let  $\mathbf{C}_{N_i}$  denote an assignment of classes to the voxels in  $N_i$ . Then, the probability of class  $C_i$  at voxel  $i$  is modelled as:

$$\begin{aligned} P(C_i | \mathbf{I}_i, \Delta \mathbf{I}_{N_i}) &= \sum_{\mathbf{C}_{N_i}} P(C_i, \mathbf{C}_{N_i} | \mathbf{I}_i, \Delta \mathbf{I}_{N_i}) \quad (4) \\ &\propto \sum_{\mathbf{C}_{N_i}} P(\mathbf{I}_i, \Delta \mathbf{I}_{N_i} | C_i, \mathbf{C}_{N_i}) P(C_i, \mathbf{C}_{N_i}) \\ &= \sum_{\mathbf{C}_{N_i}} P(\Delta \mathbf{I}_{N_i} | \mathbf{I}_i, C_i, \mathbf{C}_{N_i}) P(\mathbf{I}_i | C_i, \mathbf{C}_{N_i}) P(\mathbf{C}_{N_i} | C_i) P(C_i) \\ &\propto \sum_{\mathbf{C}_{N_i}} P(\Delta \mathbf{I}_{N_i} | C_i, \mathbf{C}_{N_i}) P(\mathbf{I}_i | C_i) P(\mathbf{C}_{N_i} | C_i) P(C_i) \quad (5) \end{aligned}$$

In this equation,  $P(C_i)$  is the prior probability of class  $C_i$ ,  $P(\mathbf{I}_i | C_i)$  models the intensity at a voxel given the class,  $P(\Delta \mathbf{I}_{N_i} | \mathbf{C}_{N_i}, C_i)$  models the intensity difference between a voxel and its neighbours given the classes in the neighbourhood, and  $P(\mathbf{C}_{N_i} | C_i)$  models the likelihood of class

transitions. In the last line we made two conditional independence assumptions. The simplification  $P(\Delta \mathbf{I}_{N_i} | \mathbf{I}_i, C_i, \mathbf{C}_{N_i}) \simeq P(\Delta \mathbf{I}_{N_i} | C_i, \mathbf{C}_{N_i})$  is understandable considering that the difference between the voxel and its neighbours is dependent on the classes present in the neighbourhood but not on the absolute intensity value of the voxel. The assumption  $P(\mathbf{I}_i | \mathbf{C}_{N_i}, C_i) \simeq P(\mathbf{I}_i | C_i)$  means that the intensity of voxel is conditionally independent of its neighbours' intensities given the class label, which is intuitively clear as well.

To infer the class labels, the energies at all voxels  $i$  have to be minimized simultaneously. The energy at voxel  $i$  is given by:

$$\begin{aligned} U(C_i | \mathbf{I}_i, \Delta \mathbf{I}_{N_i}) &= -\ln P(C_i) - \ln P(\mathbf{I}_i | C_i) - \\ &- \sum_j (\ln P(\Delta \mathbf{I}_{i,j} | C_i, \mathbf{C}_j) - \alpha m(\mathbf{C}_j, C_i)), \quad (6) \end{aligned}$$

where the summation is over all cliques  $j$  in which voxel  $i$  participates,  $P(\Delta \mathbf{I}_{i,j} | C_i, \mathbf{C}_j)$  models the difference in intensity between  $i$  and the voxels in the  $j$ -th clique given classes  $C_i$  and  $\mathbf{C}_j$ ,  $m(\mathbf{C}_j, C_i)$  is the potential of transitioning from  $C_i$  to  $\mathbf{C}_j$  and  $\alpha$  is a weighting parameter<sup>1</sup> used to handle variations between inter-slice distance and intra-slice distance. Note that this equation corresponds to a decomposition of the summation in Eq. (5), in which we explicitly list all cliques in  $i$ 's neighbourhood.

**Training:** During training, the tumour region is masked out using the expert classification labels. The remaining tissues are non-linearly registered to a brain tissue atlas, allowing us to generate separate labels for grey matter, white matter, and cerebrospinal fluid. We consider an 8-neighbourhood around the voxel in the axial plane as well as the corresponding voxels in the slices above and below. The neighbourhood  $N_i$  consists of all size 2, 3, and 4 cliques that contain voxel  $i$ . Multivariate Gaussian distributions are used to model both class intensities and intensity differences between various classes. The class transition probabilities are extracted from the frequency of co-occurrence of different class combinations in the training volumes.

**Classification** The probabilities computed by the previous layer through Eq. (3) are used as priors for the tumour and the edema classes. The voxels which are more likely to be pathological according to these estimated probabilities are masked out, and healthy atlases are registered to the remaining regions to get the prior probabilities of the different types of healthy tissues, as described above. Iterated conditional modes (ICM) [10] are used to minimize the total energy. We note that this approach works well here because the initial class labels based on the Gabor features provide a good starting point for the optimization.

<sup>1</sup>In all experiments,  $\alpha = 1$

### 2.3. Regional MRF

This stage defines an MRF in which each node represents a “region” of the image. A region encompasses a set of voxels with the same current label (as provided by the second stage), which are contiguous in 3D. Each region node is connected to all other region nodes with whom it shares a border. More precisely, regions  $i$  and  $j$  are connected if at least one voxel of region  $i$  is adjacent to some voxel in region  $j$ . The number of possible neighbours of a region is not limited. The result of this construction is an irregular, non-lattice-based MRF.

The parameters of the MRF are designed to capture 3 types of information. The first component is based on the intensities and textures at every voxel in the region. The second component is the tissue prior at every voxel based on the Bayesian tumour probabilities. The third component captures the interaction between the node and its neighbours. The interaction energy is based on the number of voxels of the two regions that are adjacent to each other, modified by the likelihood of transition between the two regions. The total energy at the voxel is a sum of the intensity and texture similarities at the voxel with the class under consideration, the prior probability of the class at the voxel, and the interaction energy that depends on the number of voxels adjacent to each other, modulated by the number of adjacent voxels of the two classes.

More precisely, for a region  $j$ , we consider both intensity information,  $\mathbf{I}_j$ , as well as texture information, given by texture features  $\mathbf{G}_j$ . The probability of class  $C_j \in \{0, \dots, M-1\}$  at region node  $j$  can be modelled as:

$$\begin{aligned} P(C_j|\mathbf{I}_j, \mathbf{G}_j, \Delta\mathbf{I}_{N_j}) &= \sum_{\mathbf{C}_{N_j}} P(C_j, \mathbf{C}_{N_j}|\mathbf{I}_j, \mathbf{G}_j, \Delta\mathbf{I}_{N_j}) \\ &\propto \sum_{\mathbf{C}_{N_j}} P(\mathbf{I}_j, \mathbf{G}_j, \Delta\mathbf{I}_{N_j}|\mathbf{C}_j, \mathbf{C}_{N_j})P(\mathbf{C}_j, \mathbf{C}_{N_j}) \\ &= \sum_{\mathbf{C}_{N_j}} [P(\Delta\mathbf{I}_{N_j}|\mathbf{G}_j, \mathbf{I}_j, C_j, \mathbf{C}_{N_j})P(\mathbf{G}_j|\mathbf{I}_j, \mathbf{C}_{N_j}, C_j) \\ &\quad P(\mathbf{I}_j|C_j, \mathbf{C}_{N_j})P(\mathbf{C}_{N_j}|C_j)P(C_j)]. \end{aligned} \quad (7)$$

We use the same simplifying assumptions as in the previous stage, and furthermore assume that:  $P(\mathbf{G}_j|C_j, \mathbf{C}_{N_j}, \mathbf{I}_j) = P(\mathbf{G}_j|C_j)$  (i.e. the texture features only depend on the class at the node and they obey a naive Bayes assumption with respect to the intensity features at the node). Incorporation of all conditional independence assumptions in the formula, we get:

$$\begin{aligned} P(C_j|\mathbf{I}_j, \mathbf{G}_j, \mathbf{I}_{N_j}) &\propto P(C_j)P(\mathbf{G}_j|C_j)P(\mathbf{I}_j|C_j) \\ &\quad \sum_{\mathbf{C}_{N_j}} P(\Delta\mathbf{I}_{N_j}|C_j, \mathbf{C}_{N_j})P(\mathbf{C}_{N_j}|C_j) \end{aligned} \quad (8)$$

The equivalent MRF energy equation for each region is

given by:

$$\begin{aligned} U(C_j|\mathbf{I}_j, \mathbf{G}_j, \Delta\mathbf{I}_{N_j}) &= -[\ln P(C_j) + \ln P(\mathbf{I}_j|C_j) + \ln P(\mathbf{G}_j|C_j) \\ &\quad + \sum_k \ln P(\Delta\mathbf{I}_{j,k}|C_j, \mathbf{C}_k)] - \alpha\psi(\mathbf{C}_k, C_j), \end{aligned} \quad (9)$$

where the summation is over all cliques  $k$  in which region  $j$  participates, and  $\psi(\mathbf{C}_k, C_j)$  is the potential of transitioning from  $C_j$  to  $C_k$ . We now describe how the different required components of this formula are modelled for the regions.

#### Computation of $\mathbf{I}_j$

This is computed as the mean intensity over the voxels of region  $j$ .

#### Computation of $\Delta\mathbf{I}_{N_j}$

We consider all voxels of region  $j$  that have one or more voxels of region  $k$  adjacent to them. To compute  $\Delta\mathbf{I}_{N_j}$ , we consider neighbouring voxels at the border of the two regions. For every voxel in region  $j$  that has a neighbouring voxel in region  $k$ , we compute the intensity difference.  $\Delta\mathbf{I}_{N_j}$  is computed as the mean of these intensity differences of all pairs of adjacent voxels from the two regions.

#### Computation of $\mathbf{G}_j$

$\mathbf{G}_j$  are the texture features computed at every voxel. To compute them, we use all the Gabor coefficients computed at stage 1 from windows for every voxel  $s$  within region  $j$  such that the Gabor windows at voxel  $s$  lie completely within the region (i.e., if the decay of the window is to less than a thousandth compared to the peak at all the border voxels). In other words, all possible windows are checked at every voxel. If the window lies completely within the region, the Gabor coefficient produced by that window is chosen, otherwise, it is discarded.  $\mathbf{G}_j$  is mean of the regional texture computed for region  $j$ .

**Training and classification:** The Gabor features  $\mathbf{G}_j$ , intensities  $\mathbf{I}_j$ , and intensity differences  $\Delta\mathbf{I}_{N_j}$  of all the different tissue classes and class combinations are modelled as multivariate Gaussians. The region class transition potential is computed from the frequency of co-occurrence of the regions in the training set. Training proceeds otherwise as described in the previous section.

We only consider cliques of size 2. The neighbourhood of a region consists of every other region to which it is adjacent. We start with the labels obtained by the adapted intensity-based MRF and apply ICM for the inference, as these labels are often a good starting point.

### 2.4. Iterative inference

Once the regional model performs its computation, we have additional information which needs to be passed back to the voxel-based MRF, in order to improve the outcome. In addition to the class labels, we also pass down the texture features at every voxel, computed as described above. The voxel-level MRF used from this stage onwards is very

similar to the adapted MRF, except that it also incorporates the texture features. Hence, the energy is given by:

$$U(C_i | \mathbf{I}_i, \mathbf{G}_i, \Delta \mathbf{I}_{N_i}) = -[\ln P(C_i) + \ln P(\mathbf{I}_i | C_i) + \ln P(\mathbf{G}_i | C_i)] + \sum_j (\log P(\Delta \mathbf{I}_{i,j} | C_i, C_j) - \alpha m(C_j, C_i)), \quad (10)$$

where  $\mathbf{G}_i$  is the textural feature computed at the regional MRF level.

Once the voxel-level MRF re-computes the class labels, the regional model is re-computed, and the iterative inference proceeds as described above. The process stops when the changes observed in the class probabilities between two consecutive stages are smaller than a threshold.

### 3. Experiments and Results

We trained our algorithm on all 20 high grade and 10 low grade real clinical tumour volumes available from the MICCAI 2012 Brain Tumour Segmentation Training dataset [11]. The datasets include MRI with several different contrasts: *T1*, *T2*, *T1* after gadolinium enhancement (i.e. *T1c*) and *FLAIR*. All volumes were pre-processed with bias field correction and were normalized. The volumes include four tumour labels: necrotic core, edema, solid tumour and enhancing tumours, for high grade and low grade tumour cases. Training was performed separately for the high grade and low grade tumour cases due to their differences in appearance. The result of our segmentation framework is a label at every voxel in the 3D MRI volumes. The framework provides 8 possible labels: the four previously mentioned tumour tissues, and four healthy tissues (grey matter, white matter, cerebrospinal fluid and partial volume<sup>2</sup>). After training, our technique was evaluated on two different challenge data sets. The first was the MICCAI 2012 Brain Tumour Segmentation Challenge dataset available at [11]. This data set consists of 11 high grade and 4 low grade clinical tumour cases. The second set was the MICCAI 2013 Brain Tumour Segmentation Challenge data set available at [12], consisting of 10 high grade clinical tumour cases. The resulting labels were uploaded to the challenge websites, and all evaluations were performed by the online system. The metric used to evaluate segmentation results were based on Dice similarity coefficients. Although the labels for all the tumour subclasses were produced by the proposed framework, the online evaluation system provided only three evaluation metrics for each algorithm: the Dice coefficients for the entire tumour (i.e. all the tumour subclasses combined including edema), the tumour core (comprised of necrotic core, solid tumour and enhancing tumour) and the enhancing tumour alone. The class labels were computed on a Dell Optiplex 980 I7 machine. The entire

<sup>2</sup>Partial volume is a class corresponding to voxels that comprise a mixture of grey matter and cerebrospinal fluid

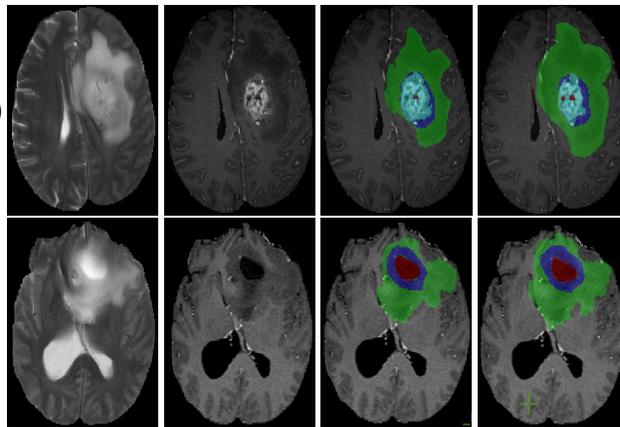


Figure 3. (a) Top row, high grade tumour case, bottom row, low grade tumour case. (a) The unlabelled T2 slice, (b) the unlabelled T1 MRI slice after injection of a contrast agent (c) expert labelling and (d) labels produced by our algorithm (red = necrotic core, green = edema, dark blue = solid tumour, light blue = enhancing tumour). Note that these correspond closely with the experts' labels.

process of segmentation takes roughly 75 minutes per volume.

Our choice of  $b$  for the Gabor filters was such as to provide 6 equally spaced orientations between 0 and  $\pi$  radians (sufficient in practice) and  $a = 1$ .

#### 3.1. Qualitative Results

Fig. 3 shows the results of our algorithm on a slice from a low grade tumour and one from a high grade tumour, against the experts' segmentation, along with the corresponding unlabelled T1c and T2 slices. Visually, in both cases, it can be seen that our results are comparable to the experts' labelling.

#### 3.2. Quantitative Results

Tables 1 and 2 show the results of our technique on real glioma cases (both high grade and low grade) and compare it against the two top performing algorithms in the two challenges (BRATS 2012 and BRATS 2013), based on Dice similarity coefficients. Our algorithm outperforms the winning algorithm by about 25% in the case of tumour cores and about 5% for enhanced tumours in the case BRATS 2012. In the case of the BRATS 2013, our algorithm outperforms the winner for tumour cores by about 10% and for enhanced tumours by about 9%. The tables also show that our performance for the entire tumour is comparable to the winners in the two challenges.

### 4. Conclusions and future work

In this paper, we presented a new iterative, multi-stage graphical model framework aimed at segmenting pathology

Method	Complete Tumour	Tumour Core	Enhancing Tumour
Zikic. et. al. [5]	<b>0.75</b>	0.54	<b>0.41</b>
Zhao. et. al. [13]	<b>0.82</b>	<b>0.55</b>	0.41
Iterative Multilevel-MRF	0.73	0.70	0.452

Table 1. Comparison of the segmentation results of the proposed method, "Iterative Multilevel-MRF", and the top two methods participating in the BRATS Challenge 2012, as presented on the live challenge website [11] for entire tumour, tumour core and enhancing tumour for real, clinical cases. Shown are average Dice similarity coefficient values. The winners of the challenge in each category are highlighted in bold. Our technique outperformed the winner by about 25% for tumour cores, and by about 5% for enhancing tumours.

Method	Complete Tumour	Tumour Core	Enhancing Tumour
Tustison. et. al. [14]	<b>0.87</b>	<b>0.78</b>	<b>0.74</b>
Meier. et. al. [15]	0.82	0.73	0.69
Iterative Multilevel-MRF	0.86	0.86	0.77

Table 2. Comparison of the segmentation results of the proposed method, "Iterative Multilevel-MRF", and the top two methods participating in the BRATS Challenge 2013 as presented on the website [12] against the experts' labels for the complete tumour, tumour core, and enhancing tumour for real, clinical cases. Shown are average Dice similarity coefficient values. The winners of the challenge in each category are highlighted in bold. Our technique performs comparably to the winners of the challenge for the entire tumour and outperforms the top method by about 10% better for tumour cores, and by about 5% for enhancing tumours.

from medical images. The model was designed to leverage the strengths of both a local, voxel-based MRF and a contextual, regional (non-lattice based) MRF, in order to penalize implausible regional labels and label combinations, while also attaining accurate boundaries. The framework was applied to the challenging problem of segmenting multiple brain tumour subclasses in real clinical patient MRI. The approach was trained and tested on public databases from the MICCAI BRATS challenge, and was shown to outperform the winning approaches for the case of tumour core and enhancing tumour segmentation, while performing comparably in the other segmentation tasks.

Here we focused on brain tumour segmentation, but the framework we presented is general and can be applied to other challenging segmentation tasks. We are currently adapting this approach to the challenging problem of segmenting lesions in brain images from patients with Multiple Sclerosis. In this context, pathologies tend to be smaller and quite similar to the surrounding tissue, so we are exploring alternative texture features. In general, the choices of inference method and of modelling both the features and their distributions deserve further investigation.

## References

- [1] Pohl, K.M., et. al.: A unifying approach to registration, segmentation, and Intensity Correction. In: LNCS 3750, Proc MICCAI. (2005) 310–318.
- [2] Prastawa, M., et. al.: A brain tumor segmentation framework based on outlier detection. In: Med. Image Ana. Volume 8(3) (2004) 275–283.
- [3] G. Farias et. al.: Brain Tumour Diagnosis with Wavelets and Support Vector Machines. In: Int. Conf. Intell. Systems and Knowledge Engg. (2008) 1453–1459.
- [4] Bauer, S., et. al.: Segmentation of Brain Tumour Images Based on Integrated Hierarchical Classification and Regularisation. In: BRATS MICCAI (2012).
- [5] Zikic, D., et. al.: Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In: MICCAI Volume 15(3) (2012) 369–376.
- [6] Subbanna, N.K., et. al.: Hierarchical Probabilistic Gabor and MRF Segmentation of Brain Tumours in MRI Volumes, in Proc. MICCAI (2013) 751–758.
- [7] Zibulski, M., et. al.: Discrete multiwindow Gabor-type transforms. In: IEEE Trans. Sig. Proc. Volume 45(6) (1997) 1428–1442.
- [8] Subbanna, N., et. al.: Existence Conditions for Non-Canonical Discrete Multiwindow Gabor Frames. In: IEEE Trans. Sig. Proc. Volume 55(10) (2007) 5113–5117.
- [9] Boykov, Y., et. al.: Fast Approximate Energy Minimization via Graph Cuts. In: IEEE Trans. Patt. Recn., Volume 23(11) (2001) 1222–1239.
- [10] Duda, R., et. al.: Pattern Classification. John Wiley and Sons (2000).
- [11] <http://www2.imm.dtu.dk/projects/BRATS2012/data.html>
- [12] <http://martinos.org/rtim/miccai2013/results.html>
- [13] Zhao, L., et. al.: Brain Tumor Segmentation based on GMM and Active Contour Method with a Model aware Edge Map. In: BRATS MICCAI 2012 (2012) 19–23.
- [14] Tustison, N., et. al.: ANTS and Arboles. In: BRATS MICCAI 2013 (2013) 47–50
- [15] Meier, R., et. al.: A Hybrid Model for Multimodal Brain Tumor Segmentation. In: BRATS MICCAI 2013 (2013) 31–36.