Multimodal Multiclass Brain Tumour Image Segmentation Using Texture Features and Conditional Random Field

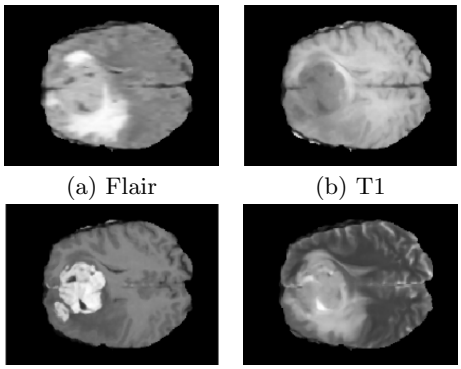
Haocheng Shen and Jianguo Zhang
Computer Vision and Image Processing Group
School of Computing, University of Dundee
Dundee, UK
hyshen@dundee.ac.uk, j.n.zhang@dundee.ac.uk

ABSTRACT

Brain tumour segmentation is an important and challenge task in monitoring the growth or shrinkage of the tumour in patient during therapy. We propose an automatic segmentation framework which applies Random Forest to classify texture features extracted from multimodal magnetic resonance (MR) images to different abnormal categories. Conditional Random Field (CRF) are then used to incorporate spatial constraints to achieve final results. The proposed method is evaluated on BRATS 2013 dataset.

1. INTRODUCTION

Brain tumour segmentation is crucial for monitoring the growth or shrinkage of the tumour in patient during therapy. However, it is a challenge problem due to the considerable intra-class variations in visual appearance of tumours from patient to patient: the structures are non-gird and they are varying in size and location. Multimodal magnetic resonance imaging (MRI) images are extensively used in brain disease diagnosis and radiotherapy due to their ability to provide complementary information for the diagnosis. The normally used modalities including T_1 -weighted MRI (T_1), T_1 -weighted MRI with contrast enhancement (T_{1c}), T_2 -weighted MRI (T_2) and T_2 -weighted MRI with fluid-attenuated inversion recovery (T_{FLAIR}) are used to enhancing different compartment of the tumour. An example of one slice in different modalities is shown in Figure 1. In current clinical practice, the delineation of tumour boundaries is still perform manually, which is time-consuming and tedious for radiologists and is also of limited use for an objective quantitative analysis. Therefore, we propose an automatic brain tumour segmentation framework using texture features from multimodal MRI and evaluate on BRATS 2013 dataset showing promising results.



(a) Flair (b) T1 (c) T1c (d) T2

Figure 1: Example of MRI multi-modal images in BRATS dataset: (a) Flair; (b) T1; (c) T1c; (d) T2.

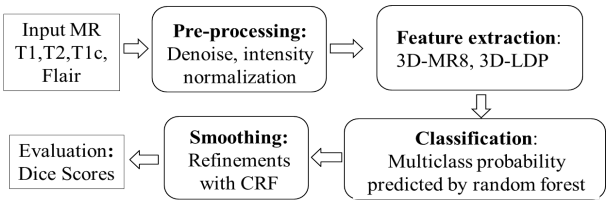


Figure 2: the proposed brain tumour segmentation framework

2. METHODS

We consider the brain tumour segmentation as a pixel-level classification problem. The framework is shown in Figure 2.

2.1 Feature Extraction

Feature extraction of voxels in the image dominates the performance of classification and segmentation. For multi-modality images, we extract feature of each voxel in each modality individually and then concatenate them together to form the final feature representation. Two types of local features are used to describe each voxel:

- **3D Maximum Response Filter (3D-MR8)**: The MR8 filter bank was proposed by [7] and has been proven robust in texture classification. It consists of 38 filters (a Gaussian and a Laplacian of Gaussian, an edge filter and a bar filter at 3 scales and 6 orientations). Measuring only the maximum response across orientations reduces the number of responses from 38 to 8 resulting in rotation invariant filters. We extend it into 3D version by calculating the original MR8 filter responses in saggital, axial and coronal views of the 3D volume separately, and then concatenate them together to form the feature representation of each voxel.

- **3D Local Difference Pattern (3D-LDP)**: LDP [6] avoid coarse quantizations of Local Binary Pattern (LBP) [5, 4] by keeping the original difference between the centre voxel and its neighbour in the patch rather than quantizing them into binary codes. We extend LDP into 3D version similarly to [8] and multi-resolution LDP patterns with radius from 1 to 3 based

on three orthogonal planes of a extracted 3D patch are calculated and then concatenated as a feature representation.

Each type of feature is encoded by Bag-of-words (BoW) and then concatenated to form the final BoW representation of the central voxel.

2.2 Classification

We choose Random Forest (RF) for classification of different abnormal tissue types in the brain. The RF is an ensemble learning algorithm that generates many weak classifiers and aggregates their results in order to make decision. Compared with SVM, RF is inherently designed for multiclass classification problem and is able to output posterior probability directly for post-processing. We rely on threshold-based entropy information gain as the split function of each node and the number of trees is set to 50.

2.3 Spatial Regularization

Classifier-based segmentation methods assume that the individual feature is independent and identically distributed (i.i.d.). Approaches based on random fields relax the i.i.d. assumption by incorporating spatial constraints and enforcing adjacent pixels belonging to the same class. Thus, we apply CRF framework similar to [2] on the probabilistic outputs of RF described above. Let $G(S, E)$ be the adjacency graph of voxels, with each voxel corresponding to a node $s \in S$, and every edge $(s_i, s_j) \in E$ indicating the neighbourhood relationship between two voxels s_i and s_j. Then CRF minimizes an energy of the form [2]:

$$-log(P(\mathbf{c}|G;w)) = \sum_{s_i \in S} \psi(c_i|s_i) + w \sum_{(s_i,s_j)\in E} \Phi(c_i, c_j|s_i, s_j) \tag{1}$$

We directly use the probability outputs $P(c_i = c|s_i)$ from RF to define the node (ψ) and edge potentials (Φ):

$$\psi(c_i|s_i) = -log(P(c_i|s_i)) \tag{2}$$

$$\Phi(c_i, c_j|s_i, s_j) = \frac{|c_i - c_j|}{1+|P(c_i|s_i) - P(c_j|s_j)|} \tag{3}$$

Where $\mathbf{s}_i$ is the feature representation of pixel i and the weight w in Eq. 1 represents the trade-off between the spatial regularization (edge-potential) and the confidence in the classification (node-potential) which is learned based on the cross-validation in the training set. We use the public library for graph-optimization [1] for the label inference.

3. RESULTS

The dataset we used is provide by MICCAI 2013 challenges on Multimodal Brain Tumour Segmentation [3]. It consists 20 high-grade glioma subjects in four modalities(T_1, T_{1c}, T_2, T_{FLAIR}) which were annotated by 3∼4 radiologists. The complete tumour region is subdivided into 4 structures: necrosis, enhancing core, non-enhancing core and edema. All 3D volumes are linearly co-registered to the T_{1c} image, skull stripped and interploated to 1mm isotropic resolution.

We used 5-fold cross validation on the dataset and the segmentation results were uploaded to the online evaluation platform and were evaluated automatically using the online evaluation tool[1]. And the evaluation was performed for three different tumour subregions: complete tumour(including all four tumour structures), tumour core(including all tumour structures except edema) and enhancing core(enhancing core only). The following Dice score is applied to measure the overlap between the predicted segmentation results(P) and the ground truth(T):

$$Dice(P, T) = \frac{|P_1 \wedge T_1|}{(|P_1| + |T_1|)/2} \tag{4}$$

The Dice score of our proposed methods for complete, core and enhancing are 0.78, 0.66, 0.61 respectively, which ranks top 10 among over 30 participated groups.

4. FUTURE WORK

In order to build automatic brain tumour segmentation system, complete pixel-level ground truth provided by radiologists are desirable. However, it is expensive and time-consuming to obtain. Our future work will focus on developing learning-based segmentation methods involving partial annotations (e.g., bounding box, clicks, curves) while obtaining competitive segmentation results compared to the ones using complete annotations.

5. REFERENCES

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222–1239, 2001.

[2] B. Fulkerson, A. Vedaldi, and S. Soatto. Class segmentation and object localization with superpixel neighborhoods. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 670–677. IEEE, 2009.

[3] B. Menze, M. Reyes, and K. Van Leemput. The multimodal brain tumorimage segmentation benchmark (brats). 2014.

[4] L. Nanni, A. Lumini, and S. Brahnam. Local binary patterns variants as texture descriptors for medical image analysis. *Artificial intelligence in medicine*, 49(2):117–125, 2010.

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *Computer Vision-ECCV 2000*, pages 404–420. Springer, 2000.

[6] G. Sharma, S. ul Hussain, and F. Jurie. Local higher-order statistics (LHS) for texture categorization and facial analysis. In *Computer Vision–ECCV 2012*, pages 1–12. Springer, 2012.

[7] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

[8] G. Zhao and M. Pietikainen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(6):915–928, 2007.

[1] http://virtualskeleton.ch/