Automated Colorectal Tumour Segmentation in DCE-MRI using Supervoxel Neighbourhood Contrast Characteristics

Benjamin Irving[1], Amalia Cifor[1], Bartłomiej W. Papież[1], Jamie Franklin[2], Ewan M. Anderson[2], Sir Michael Brady[3], and Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering,
Department of Engineering Science, University of Oxford, UK,
benjamin.irving@eng.ox.ac.uk
[2] Department of Radiology, Oxford University Hospitals NHS Trust, Oxford, UK,
[3] Department of Oncology, University of Oxford, UK,

Abstract. Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a powerful protocol for assessing tumour progression from changes in tissue contrast enhancement. Manual colorectal tumour delineation is a challenging and time consuming task due to the complex enhancement patterns in the 4D sequence. There is a need for a consistent approach to colorectal tumour segmentation in DCE-MRI and we propose a novel method based on detection of the tumour from signal enhancement characteristics of homogeneous tumour subregions and their neighbourhoods. Our method successfully detected 20 of 23 cases with a mean Dice score of 0.68 ± 0.15 compared to expert annotations, which is not significantly different from expert inter-rater variability of 0.73 ± 0.13 and 0.77 ± 0.10. In comparison, a standard DCE-MRI tumour segmentation technique, fuzzy c-means, obtained a Dice score of 0.28 ± 0.17.

1 Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is becoming a key modality for monitoring of tumour progression and response to therapy. DCE-MRI scans show contrast uptake in tissue, and, in principle, provide information about the perfusion and vascularity. A bolus of contrast agent is injected into a peripheral vein and travels through the circulatory system and into the extravascular-extracellular space (EES) – leading to characteristic tissue signal enhancement curves. These can be parameterised by pharmacokinetic (PK) models [9] and provide a measurement of tumour changes to monitor patient outcomes [2]. It is increasingly recognised that tumour heterogeneity is present in colorectal tumours due to the structure of the vasculature, hypoxia and necrosis, and that it is key to assessing the likely efficacy of therapy.

Accurate tumour segmentations are required for DCE-MRI in order to correctly characterise these changes. This is a challenging and time consuming task to perform manually, partly because of the complex signal enhancement patterns, limited soft tissue contrast and low resolution of the 4D DCE-MRI scan.

Alternatively, high resolution T2-weighted images can be acquired, manually annotated and non-rigidly registered to the DCE-MRI, but this is also time consuming. Instead, DCE-MRI contrast uptake curves potentially offer tumour specific features for automated segmentation.

In DCE-MRI, PK models [9] are popular methods of quantifying contrast enhancement but results are sensitive to the choice of compartment model, which in turn depends on the tissue. In practice, no single model may be adequate for the whole region, and Hamy et al. [6] use principal component analysis (PCA). To-date DCE-MRI colorectal tumour segmentation remains relatively unexplored. Fuzzy c-means has become established as a method to segment tumours in DCE-MRI images [3], which has the advantage that it is unsupervised. It iteratively assigns a fuzzy label to each voxel based on the distance of the enhancement curve from each cluster centre. However, the complex anatomy of the lower abdomen renders unsupervised clustering less effective. Fulkerson et al. [5] use a quick-shift superpixel representation of a 2D image for object recognition using a support vector machine classifier with features extracted from each superpixel and its neighbours. Mahapatra et al. [8] also develop a supervoxel approach for localisation of regions contraining Crohn's disease in conventional 3D MRI. Both [5], [8] use 2D and 3D supervoxel representations. However, unlike previous analysis, we develop a method for 4D contrast varying imaging using novel PCA and SLIC coupling using heterogeneous local neighbourhood characteristics.

We propose a novel method to automatically segment colorectal tumours from DCE-MRI scans as a key first step toward representing and quantifying heterogeneity. The method generates a supervoxel representation of the dynamic image and detects supervoxel regions that contain tumour based on features of the supervoxel and neighbourhood characteristics. Our contributions include: a novel method to automatically segment colorectal tumours from DCE-MRI (the authors are not aware of any existing method), robust learning from poorly defined masks, extension of simple iterative clustering (SLIC) [1] to 4D DCE-MRI, and supervoxel neighbourhood based learning for encoding tumour characteristics. Sec. 2 outlines the dataset, in Sec. 3 our segmentation method is introduced, and the method is evaluated on 23 colorectal cancer cases in Sec. 4.

2 Materials

T1-weighted DCE-MRI scans were acquired from 23 patients with rectal adenocarcinomas using a 1.5 T GE scanner with a gradient echo, fat-suppressed sequence (LAVA) (TR=4.5 ms, TE=2.2 ms and flip angle 12^o). The scans were acquired prior to downstaging chemo-radiotherapy. Multihance^{TM}was injected just after the scan start, and images were acquired every 12 s for between 20 and 25 successive periods at a resolution of $0.78 \times 0.78 \times 2.0$ mm. High resolution small FOV axial-oblique T2-weighted images were acquired prior to the DCE-MRI scan with a resolution of $0.39 \times 0.39 \times 3.30$mm (TR=14ms, TE=12ms, flip angle 40^o). Scans were acquired axial to the tumour at the point of invasion in order to minimise partial volume effects.

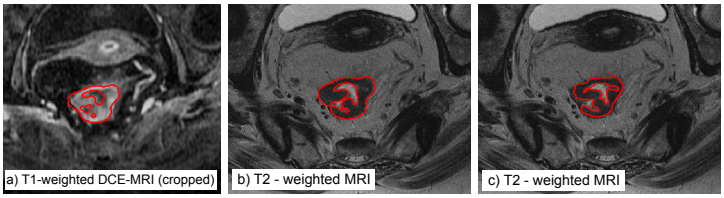

Fig. 1. Colorectal MRI images with annotations a) DCE-MRI axial slice (with tumour annotation) b,c) Corresponding T2 MRI showing inter-rater variability

Colorectal tumours were delineated on the high resolution T2-weighted scans by a radiologist and registered to the DCE-MRI scans. The T2 and DCE-MRI scans were acquired consecutively and the majority of the alignment is performed using the DICOM coordinates. The T2 scans were then resampled and registered to the baseline DCE-MRI (prior to contrast) to correct for minor abdominal motion in the time between scans [7]. Using this transformation, masks could be aligned to the DCE-MRI volume, and were checked by a radiologist. However, manual segmentation of colorectal tumours is a challenging task because of: partial volume effects in the axial plane, complex anatomy in the lower rectum making it difficult to delineate normal anatomical structures, wall thickening due to venous congestion, and mucinous tumours. Fig. 1 shows a single axial slice of the T1-weighted DCE-MRI (at a fixed time) and the T2-weighted MRI. Two annotations are shown to illustrate expert variability, and this sets a fundamental limit on the evaluation of any method against the radiologist "ground truth".

3 Method

Our colorectal tumour segmentation method uses a PCA representation of the contrast uptake curves as input to an n-feature Simple Linear Iterative Clustering (SLIC) algorithm of the heterogeneous tumour and surrounding tissue in DCE-MRI. This section describes the preprocessing, supervoxel extraction, derivation of features from the supervoxels, and the segmentation of unseen cases.

Preprocessing: A subregion is defined by detecting the MRI foreground based on Otsu thresholding to find the patient boundary. The patient boundary was used to automatically crop the DCE-MRI to an ROI surrounding tumour (approximately 25% of the original volume) based on the general location of the rectum. **Processed tumour volume-of-interest.** The registered DCE-MRI tumour annotation provides the ground truth to learn to characterise the tumour in the image. Small failures in the ground truth result in poor learning and, therefore, supervoxels entirely contained within the annotation were used with linear discriminant analysis (LDA) to assign a posterior probability to each voxel of being part of the tumour. Supervoxels containing at least 20% of the annotation with a high tumour probability were classified as tumour. This provided

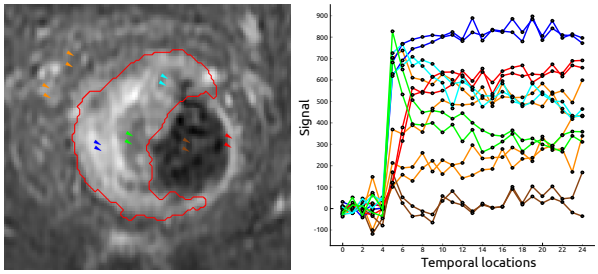

Fig. 2. Signal enhancement curves for the tumour and surrounding tissue where shape relates to tissue perfusion and vascularity. *Green*, *blue* and *cyan* are selected from regions of the tumour, *red* shows the rectal wall and mucosa, *orange* shows surrounding tissue, and *brown* the lumen. This figure illustrates the heterogeneity in the tumour, as well as similarities between subregions of the tumour.

a more robust ground truth for training (Fig. 3). **Injection time and image normalisation.** Injection time may vary between scans and was detected from the steepest gradient in contrast enhancement in the image (contrast in the arteries). Signal enhancement curves were also normalised by the 80th percentile of the maximum contrast.

Supervoxel representation: Fig. 2 shows signal enhancement curves for regions in the tumour (red mask) and surrounding tissue – illustrating the heterogeneity in the tumour, as well as similarities within subregions. Examples such as this motivate use of a supervoxel representation to cluster consistent subregions and extract the connectivity between subregions of the tumour, lumen and wall.

PCA applies a linear transform to project corresponding points into uncorrelated space and for dimensionality reduction. The eigenvectors from the covariance matrix of the features are used to apply this projection: $\mathbf{b} = \Phi^T(\mathbf{x} - \bar{\mathbf{x}})$ where $\mathbf{x}$ is the signal enhancement curve, $\mathbf{b}$ is the representation in uncorrelated space and Φ^T is the transposed eigenvectors of the covariance. A single enhancement curve is represented as the mean shape and a linear combination of each principal component by $\mathbf{x} \approx \bar{\mathbf{x}} + \Phi\mathbf{b}$. Standard deviation of each mode is given by the eigenvalue $\sigma_i = \sqrt{\lambda_i}$. The enhancement curves are represented by the first 5 principal modes (99% of the variation). Fig. 3a shows variation of the curves with two standard deviations from the curve mean ($\mu \pm 2\sigma_i$).

SLIC [1] typically generates superpixels from an image using an adaptation of k-means clustering and penalising distance from the cluster centre. SLIC has been shown to be a fast method with good performance [1]. SLIC initialises k cluster centres by sampling the grid regularly with distance $S = \sqrt{(N/k)}$, where N denotes the number of voxels. A distance function is defined that usually combines a color/intensity distance and spatial distance measure, and voxels are assigned to the closest cluster by searching a 2S x 2S x 2S neighbourhood (3D volume). We have extended SLIC to an n-feature image to enable extraction of

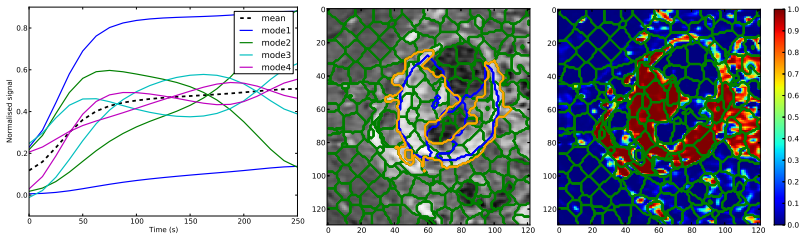

Fig. 3. SLIC supervoxel generation a) Two standard deviations of the first 4 modes of variation from the mean enhancement curve b) Supervoxels shown on a single slice and at a single time point of the 4D DCE-MRI volume (registered tumour mask in blue and preprocessed in orange) c) Posterior probabilities used for mask preprocessing.

supervoxels from a DCE-MRI image based on the features of the enhancement. The feature distance (d_f) and spatial distance(d_s) are defined as:

$$d_f^2 = \frac{1}{n}\sum_{k}^{n}(b_{jk} - b_{ik})^2, \qquad d_s^2 = \sum_{k}^{3}(x_{jk} - x_{ik})^2 \tag{1}$$

where b_{jk} is the kth principal component of the jth voxel, and (x_{j1}, x_{j2}, x_{j3}) is the 3D coordinate of the jth voxel in mm. As discussed, we use n=5 components. The range of each principal component is normalised between [0, 1] for the volume. The distances can be combined as a relative distance measure by:

$$D = \sqrt{d_f + (d_s/r)} \tag{2}$$

where $r = \frac{S}{\text{compactness}}$ is a weight. *Compactness* was chosen to be 0.05 and the average cluster size was chosen be 400 voxels by qualitatively assessing the ability of the supervoxel algorithm to correctly separate thin structures such as mucosal walls for the first 4 cases. Fig. 3 shows the modes of variation used to characterise the curves and a 2D axial slice of the 3D supervoxel representation.

Features from supervoxel neighbourhoods: Principal components are found for the voxel contrast enhancement curves, and n-feature SLIC was used to generate supervoxel clusters (v_i). The mean and standard deviation of the PCA modes (b_i) in each supervoxel (v_i) were used as features ($\mathbf{f}_i$) to characterise the enhancement. Features were all normalised between [0, 1]. Superpixel connectivity is used by Fulkerson et al. [5] to capture neighbourhood variation. We use rotationally invariant supervoxel feature magnitude of the gradient to capture changes related to the tumour, because the tumour can be characterised as a contiguous region, with structure in the tumour heterogeneity, that may abut lumen that contains non-enhancing air or stool, and a thin rectal wall/mucosa (Fig. 2). The supervoxel adjacency graph $G(V, E)$ was found where each supervoxel $v_i \in V$ and edges (E) connect adjacent supervoxels (v_i, v_j). The vector

(d_{ij}) between centroids of adjacent supervoxels (v_i, v_j) and the unit vector $\hat{d}_{ij}$ was then calculated. Using G, rotationally invariant descriptors related to the gradient magnitude the neighbouring supervoxel features were used to encode neighbourhood variation and approximated for each feature by:

$$f_{\bigtriangledown i} = \sqrt{(f_{x+1} - f_{x-1})^2 + (f_{y+1} - f_{y-1})^2 + (f_{z+1} - f_{z-1})^2} \tag{3}$$

where $f_i \in \mathbf{f}_i$ is a single feature from six neighbouring supervoxels with $\hat{d}_{ij}$ most representative of the six-connected neighbour orientations. Therefore, each supervoxel is represented by 10 features from the mean and standard deviation of the modes of each supervoxel, and 10 features from the neighbourhood.

Classification and Evaluation: To date we have evaluated two classifiers: LDA and random forests, which are increasingly used in medical imaging applications [8]. Patients 1-4 were used for parameter tuning and leave-one-out cross validation (LOOCV) was used to evaluate the algorithm by testing on each patient and training on the remainder. The classifier was trained on all training set supervoxels and used to classify supervoxels in a test case. Finally, connected region size and the central moment of inertia are used to exclude additional smaller regions (shown in Fig. 4). These results were compared to the standard fuzzy c-means algorithm for DCE-MRI tumour segmentation [3].

4 Results

The registered expert annotations from the T2 images provide the ground truth in order to train and evaluate our algorithm on DCE-MRI. The tumour centre is accurate but edges may vary making annotation preprocessing an important step (Sec. 3). Variability in the T2 manual annotations and registration error were assessed to quantify the uncertainty. Ten cases were relabelled by two radiologists and inter-rater voxelwise Dice scores of 0.73 ± 0.13 and 0.77 ± 0.10 were found when comparing to the original annotations. Registration error was quantified by an expert correcting the registered mask on the DCE-MRI using the T2 as reference for four cases (case 1-4). The Dice score between the registered mask and corrected mask was 0.94 ± 0.01. Fig. 4 shows the original annotations, tumour probabilities assigned to each supervoxel during classification, and the final segmentation for two cases. Segmentations are consistent but with variability on the border which agrees with inter-observer variability. Our method successfully detected 20 cases but failed to detect 3 cases (16, 18, 20), which showed slower tumour contrast uptake. The detected cases were segmented with a voxelwise mean Dice score of 0.68 ± 0.15 (compared to the processed ground truth), and 0.56 ± 0.13 for the original ground truth. The mean Dice scores for the automated algorithm were not significantly different from either of the experts (p=0.33 and p=0.06 using a Wilcoxon rank sum test). Fuzzy c-means [3] was considerably poorer with 0.28 ± 0.17 and 0.19 ± 0.14 for the processed and original masks,

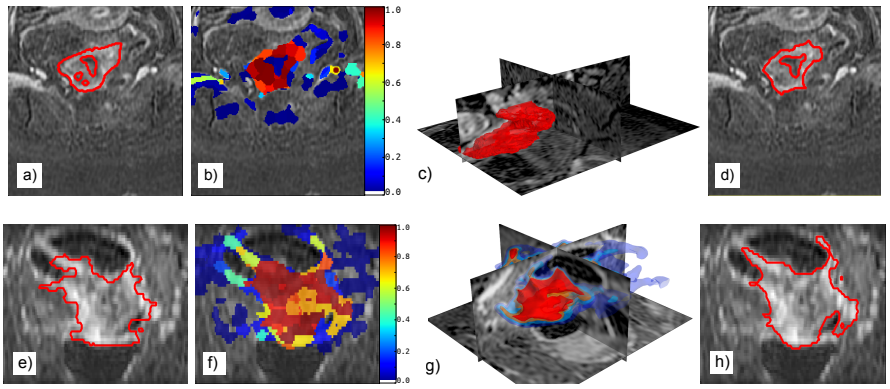

Fig. 4. A single slice through tumour example segmentations, a-d) case 6 axial slice, e-h) case 2 sagittal slice: Expert segmentations (a, e), Tumour probabilities assigned to the supervoxels during classification (transparent has $prob < 0.02$) (b, f), and (d, h) segmentations from our method. 3D representations are also shown of d) and f).

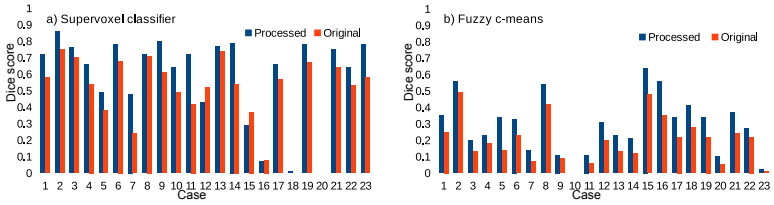

Fig. 5. Dice score for the supervoxel classification method and fuzzy c-means method compared to the processed ground truth and the original ground truth.

respectively. Fig. 5 shows the Dice score for each case using our supervoxel classifier and fuzzy c-means, compared to the preprocessed mask and original mask. RF achieved poorer results than the LDA results implemented here, which was probably due to the limited size of the training set leading to more sophisticated classifiers overfitting to the training set. This method was developed in Python and C++, and took an average of 3.7 ± 2.4 minutes to segment an unseen case (4 times faster than an expert of approximately 15 minutes).

5 Discussion

We introduced, a novel method to segment colorectal tumours directly from 4D DCE-MRI scans using the signal enhancement patterns of the tumour and surrounding regions. This has not been previously addressed, achieves results equivalent to the inter-observer variability in segmentation for most cases, and

is considerably better than the standard fuzzy c-means technique. It also demonstrates the potential of contrast enhancement curves to quantify tissue differences and tumour heterogeneity. Our method is automatic, except for two cases, where the presence of uterine fibroids resulted in similar appearance in DCE-MRI, and required manual exclusion (cases 8, 11). Future research will apply this method to a larger dataset, including post-therapy cases, and include a more explicit model-based representation of the tumour to distinguish fibroids. As an initial assessment of motion correction, we registered the DCE-MRI temporal sequences that showed considerable motion with an adapted feature-based diffeomorphic registration method for motion compensation [4]. We obtained up to an 11% improvement and will further explore this potential increment in future work.

Acknowledgements: This research is supported by the CRUK/EPSRC Oxford Cancer Imaging Centre and the Oxford EPSRC IAA. We also thank Prof. Fergus Gleeson as lead of the DCE-MRI imaging trial.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34, 2274–2281 (2012)
2. Bhushan, M., Schnabel, J., Chappell, M., Gleeson, F., Anderson, M., Franklin, J., Brady, M., Jenkinson, M.: The Impact of Heterogeneity and Uncertainty on Prediction of Response to Therapy Using Dynamic MRI Data. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I., LNCS, vol. 8149, pp. 316–323. Springer, Heidelberg (2013)
3. Chen, W., Giger, M.L., Bick, U.: A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. Acad. Radiol. 13, 63–72 (2006)
4. Cifor, A., Risser, L., Chung, D., Anderson, E., Schnabel, J.: Hybrid feature-based diffeomorphic registration for tumor tracking in 2-d liver ultrasound images. IEEE Trans. Med. Imag. 32, 1647–1656 (2013)
5. Fulkerson, B., Vedaldi, A., Soatto, S.: Class segmentation and object localization with superpixel neighborhoods. IEEE Int. Conf. Comput. Vis. pp. 670–677 (2009)
6. Hamy, V., Dikaios, N., Punwani, S., Melbourne, A., Latifoltojar, A., Makanyanga, J., Chouhan, M., Helbren, E., Menys, A., Taylor, S., Atkinson, D.: Respiratory motion correction in dynamic MRI using robust data decomposition registration - Application to DCE-MRI. Med. Image Anal. 18, 301–13 (2014)
7. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, M., Schnabel, J.A.: MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. Med. Image Anal. 16, 1423–35 (2012)
8. Mahapatra, D., Schuffler, P., Tielbeek, J., Makanyanga, J., Stoker, J., Taylor, S., Vos, F., Buhmann, J.: Automatic Detection and Segmentation of Crohn's Disease Tissues From Abdominal MRI. IEEE Trans. Med. Imag. 32, 2332–2347 (2013)
9. Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B., Lee, T., Mayr, N.A., Parker, G.e.a.: Estimating kinetic parameters from dynamic contrast-enhanced T 1-weighted MRI of a diffusable tracer: standardized quantities and symbols. J. Magn. Reson. Imaging 10, 223–232 (1999)