

# Unsupervised Segmentation of 5D Hyperpolarized Carbon-13 MRI Data Using a Fuzzy Markov Random Field Model

Charlie J. Daniels and Ferdia A. Gallagher\*

**Abstract**—Hyperpolarized MRI with  $^{13}\text{C}$ -labelled compounds is an emerging clinical technique allowing *in vivo* metabolic processes to be characterized non-invasively. Accurate quantification of  $^{13}\text{C}$  data, both for clinical and research purposes, typically relies on the use of region-of-interest analysis to detect and compare regions of altered metabolism. However, it is not clear how this should be determined from the five-dimensional data produced and most standard methodologies are unable to exploit the multidimensional nature of the data. Here we propose a solution to the novel problem of  $^{13}\text{C}$  image segmentation using a hybrid Markov random field model with continuous fuzzy logic. The algorithm fully utilizes the multi-dimensional data format in order to classify each voxel into one of six distinct classes based on its metabolic characteristics. Bayesian priors fully incorporate spatial, temporal and ratiometric contextual information whilst image contrast from multiple spectral dimensions are considered concurrently by using an analogy from color image segmentation. Performance of the algorithm is demonstrated on *in silico* data where the superiority of the approach over a reference thresholding method is consistently observed. Application to *in vivo* animal data from a pre-clinical subcutaneous tumor model illustrates the ability of the MRF algorithm to successfully detect tumor location whilst avoiding image artefacts. This work has the potential to assist the analysis of human hyperpolarized  $^{13}\text{C}$  data in the future.

**Index Terms**—Image segmentation, hyperpolarized MRI, carbon-13, Markov random field, fuzzy systems, image quantification.

## I. INTRODUCTION

WITH the increasing use of functional and molecular imaging techniques in medicine, there has been a significant increase in the size and the complexity of the data acquired. A major challenge in medical imaging is to determine the optimal use of these large data sets. For example, automated lesion identification and segmentation could greatly assist the interpretation of radiological images but is often hampered by the intrinsically noisy nature of most molecular imaging techniques when compared to conventional anatomical imaging. Hyperpolarized imaging with  $^{13}\text{C}$ -labelled compounds is an emerging clinical technique allowing *in vivo* metabolic processes to be characterized and quantified non-invasively. The method is based on dynamic nuclear polarisation (DNP); endogenously occurring metabolites labelled

with one or more  $^{13}\text{C}$  atoms are hyperpolarized at a cryogenic temperature and within a high magnetic field, before being rapidly dissolved using a heated fluid [1]. The molecules are able to retain liquid state polarization for a sufficient length of time to spectroscopically image both the substrate and its metabolic breakdown products with magnetic resonance spectroscopic imaging (MRSI), following their intravenous injection into an animal or human [2].

The conversion of hyperpolarized [1- $^{13}\text{C}$ ]pyruvate to [1- $^{13}\text{C}$ ]lactate is the most intensively studied exchange reaction to date and the first molecule to be translated into the clinic. Preclinical oncological studies have shown increased lactate exchange correlating with tumor grade [3], a change in this exchange rate to be a marker of treatment response [4] and a first-in-man study in prostate cancer has demonstrated the feasibility for human imaging [5]. [1- $^{13}\text{C}$ ]pyruvate may also be metabolized to produce [1- $^{13}\text{C}$ ]alanine and  $^{13}\text{C}$ -bicarbonate. The ability to rapidly detect small changes in metabolism *in vivo* can be applied to a number of areas within medicine including oncology, neurology [6] and cardiology [7], with a recent pilot study successfully imaging pyruvate metabolism in the healthy human heart [8].

By utilizing the chemical shift induced in the  $^{13}\text{C}$  frequency during chemical exchange, multiple metabolites can be simultaneously imaged in 3 spatial dimensions and at multiple time points with a resolution of a few seconds to generate intrinsically co-registered 5-dimensional data. Accurate and robust quantification of multidimensional  $^{13}\text{C}$  data is imperative in order to make reliable inter- and intra-patient comparisons and many methodologies have been suggested for this purpose. Most commonly, the forward exchange rate constant  $k_{PL}$  is derived by fitting variations on the following coupled differential equations to pyruvate and lactate timecourse data [4], [3], [9]:

$$\frac{dP}{dt} = I(t) - k_{PL}P + k_{LP}L - \rho_P P \quad (1)$$

$$\frac{dL}{dt} = -k_{LP}L + k_{PL}P - \rho_L L \quad (2)$$

where  $I(t)$  is a function describing the pyruvate inflow,  $k_{LP}$  is the backwards exchange rate constant and  $\rho$  describes the polarization loss for each metabolite due to both thermal  $T_1$  decay and through radiofrequency excitation during sampling. Simple model-free analyses have also been shown to accurately quantify lactate exchange [10], [11]. Both the lactate/pyruvate area under the curve (AUC) ratio and the lactate

Manuscript received 7th April 2017, revised 29th June 2017

Asterisk indicates corresponding author.

Charlie J. Daniels is with the Department of Radiology, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK (email: cjd65@cam.ac.uk)

\*Ferdia A. Gallagher is with the Department of Radiology, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK (email: fag1000@cam.ac.uk)

time-to-peak show excellent correlation with the exchange rate model approach, albeit by producing parameters less easily linked to the underlying biological reaction [11].

Clinical imaging research has conventionally made use of regions or volumes of interest (ROI or VOI) for the analysis and quantification of areas that are pathologically abnormal in comparison to normal tissue. A VOI must be accurate, robust and repeatable, however commonly used manual delineation and thresholding techniques both often suffer from high variability, as well as susceptibility to noise and image contrast [12], making statistical segmentation methods important for the reliable processing of large imaging datasets.

Unsupervised segmentation of  $^{13}\text{C}$  data offers a unique challenge due to its low spatial resolution, high noise, susceptibility to artefacts and its inherent 5-dimensionality, making analysis of this data using spatial contrast alone very difficult. Simple segmentation methods such as thresholding are made unsuitable by the rapidly changing temporal dynamics and spectral interdependency, whilst supervised learning methods are hampered by the lack of clinical data currently available to enable training. Here we propose the use of a fuzzy Markov random field (MRF) model which provides a framework for segmentation which fully incorporates spatial, temporal and spectral information whilst effectively handling the low image quality currently associated with hyperpolarized MRI.

Image segmentation describes the problem of sorting a set of image pixels or voxels into a number of distinct classes. For a set  $S = \{1, 2, \dots, s\}$  of image pixels, indexed by  $j$ , the MRF model considers two random fields; the observed, noise-corrupted field  $Y = (Y_j)_{j \in S}$  and the underlying ground-truth field  $X = (X_j)_{j \in S}$ . Here, the vector of real, positive image pixel intensity values is a realisation, defined here as a particular observed configuration,  $Y = \mathbf{y}$  of  $Y$  and the objective is to uncover the field  $X = \mathbf{x}$  representing the ‘‘true’’ image by means of iteratively optimizing an objective function. Here, the realisation  $\mathbf{x}$  represents the assignment vector of one of the set  $\Omega = \{1, 2, \dots, q\}$  of classes to each pixel  $j$ .

In addition to being an efficient approach for many problems, a main advantage of using an MRF model for medical image segmentation is the ability to formally incorporate multiple forms of contextual information. Modelling of local interactions between spatially and/or temporally adjacent pixels vastly reduces the influence of noise on the segmentation, as each pixel is classified based not only on its own contrast, but on the contrast of its neighbors. MRFs, and the closely related Markov random chains, have been extensively applied to segmentation problems in MRI [13], computed tomography (CT) [14] and positron emission tomography (PET) imaging [15].

A powerful way of accounting for poor image resolution is to incorporate fuzzy logic into the segmentation process. Rather than pixels being assigned to a ‘hard’ Boolean class, i.e. definitely tumor or definitely background with  $x_j = 0$  or 1, pixels are assigned a fuzzy membership level to each class. This is similar to, although distinct from, the probability of belonging to each class and instead provides a measure of uncertainty on the true content of the pixel. Similarities between PET and hyperpolarized  $^{13}\text{C}$  imaging

include low resolution and large partial volume effects [16]. Incorporation of fuzzy logic into PET segmentation algorithms proved successful in handling these challenges by providing a more realistic description of the blurred boundaries between regions of different classes when compared to a hard Boolean approach. This is particularly apt for both PET and  $^{13}\text{C}$ , as each large pixel is likely to contain elements of both tumor and background.

Two methodologies for applying fuzzy logic in medical image segmentation have dominated the literature; the fuzzy c-means (FCM) approach [17], [18] and the Markovian approach [19], [15]. Chatzis *et al* have proposed a hybrid FCM-MRF which combines the flexibility of the FCM model with the capacity to incorporate higher dimensional prior information of an MRF model [20]. Here, we build on their work and extend it to handle the 5D data produced by hyperpolarized  $^{13}\text{C}$  MRI, using an analogy with color image segmentation to fully incorporate the spectral dimension, in order to segment and quantify regions of interest. The primary focus is to detect the metabolism of  $[1-^{13}\text{C}]\text{pyruvate}$  to  $[1-^{13}\text{C}]\text{lactate}$ , however the capability of identifying  $[1-^{13}\text{C}]\text{alanine}$  and  $^{13}\text{C}$ -bicarbonate production is also included.

The performance of the algorithm is initially demonstrated here on an *in silico* dataset as well as *in vivo* animal data from rats with subcutaneously implanted tumors. This work demonstrates the feasibility of applying a fuzzy MRF approach to the novel and challenging problem of hyperpolarized  $^{13}\text{C}$  image segmentation in order to automatically detect regions of altered metabolism, thereby providing a useful tool to delineate tumor from normal tissue.

## II. THEORY

### A. The MRF Framework

Given a set of 2D or 3D image voxels  $S = \{1, 2, \dots, s\}$ , the objective of the MRF segmentation algorithm is to find the correct classification for each voxel which minimizes a Bayesian-derived objective function, separating the image into segments displaying similar features. The input information is the observed data; the measured voxel intensity values are a realisation  $Y = \mathbf{y}$ , where  $Y$  is the set of all possible measurements, and the desired output is to find the correct configuration  $X = \mathbf{x}$  of assigned class labels.  $Y$  and  $X$  are the ‘emitted’ and ‘hidden’ random fields respectively, where a random field is defined as having a strictly positive probability distribution:

$$p(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in X \quad (3)$$

An MRF is defined as any random field for which the joint probability distribution is a Gibbs (Boltzmann) distribution:

$$p(\mathbf{x}|\beta) = \frac{1}{Z} e^{-U(\mathbf{x}|\beta)} \quad (4)$$

Here  $\beta$  is the inverse temperature parameter, which is maximized at each iteration of the algorithm; in physical terms, this is equivalent to actively ‘cooling’ the system such that it converges to a state of minimum energy. In further analogy to statistical mechanics,  $Z$  is the partition function, defined as

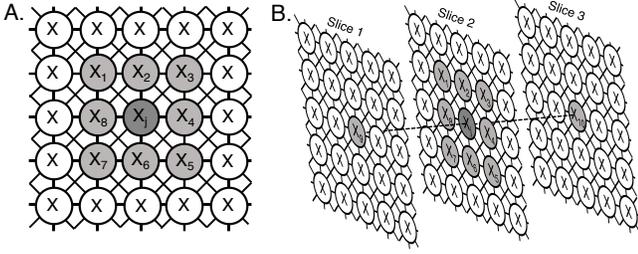


Fig. 1. Graphical depiction of image pixels  $x$  as nodes in a Markov random field and the connections between them. A. shows the 8-neighborhood in 2D (shown in gray) and B. the 10-neighborhood for 3D segmentation which form the cliques on which the state of each pixel  $x_j$  will depend.

the sum of the numerator over every possible realisation of  $\mathbf{x} \in X$ .

Exact computation of the partition function becomes computationally expensive to the point of being intractable for systems with a large number of sites; it is therefore simplified here by the commonly used mean-field approximation [21]. Under this approximation, the site of interest is decoupled from the field by averaging over its external interactions with each other site to produce a single mean effect and  $p(\mathbf{x}|\beta)$  is given by:

$$p(\mathbf{x}|\beta) = \prod_{j=1}^s p(x_j|\hat{\mathbf{x}}_{c_j}, \beta) \quad (5)$$

where  $\hat{\mathbf{x}}_{c_j}$  is an estimate of the class labels in the clique  $c_j$  obtained from the previous iteration of the algorithm.

The term  $U(\mathbf{x}|\beta)$  in (4) is the energy function containing the contextual prior information. For the set of voxels  $S = 1, 2, \dots, s$ , it is possible to define a set of cliques  $c_l$  such that for each site  $j \in S$ ,  $j \in c_l$  if and only if  $j \neq l$  and  $l \in c_j$ . This sets up the network of two-way interactions to be considered at each site; figure 1 depicts the 8- and 10-connectivity neighborhoods which form the cliques considered in the 2D and 3D contexts respectively. A main advantage of MRF models is the straightforward incorporation of additional prior information by linearly adding terms to form the energy function  $U(\mathbf{x}|\beta)$ . Three terms are considered here with equal weighting: a spatial term  $U_S(\mathbf{x}|\beta)$  utilizing the network discussed above; a temporal prior  $U_T(\mathbf{x}|\beta)$  testing for an expected shape profile; a radiometric prior  $U_R(\mathbf{x}|\beta)$  which tests for expected metabolite ratios. In accordance with Bayes theorem, the joint probability of the Gibbs prior distribution for the hidden field  $p(\mathbf{x})$ , and the emitted field distribution  $p(\mathbf{y})$  is given by:

$$p(\mathbf{x}, \mathbf{y}) = p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \quad (6)$$

It is reasonable to assume that the full conditional distribution can be written as the product of contributions from each individual site and that the distributions are of multivariate Gaussian form:

$$p(\mathbf{y}|\mathbf{x}) = \prod_{j=1}^s p(\mathbf{y}_j|x_j) = \frac{e^{(\mathbf{y}_j - \boldsymbol{\mu}_{x_j})^T \boldsymbol{\Sigma}_{x_j}^{-1} (\mathbf{y}_j - \boldsymbol{\mu}_{x_j})}}{\sqrt{(2\pi)^n |\boldsymbol{\Sigma}_{x_j}|}} \quad (7)$$

where  $\mathbf{y}_j$  is the vector of measured intensity values of dimensionality  $n$ ,  $\boldsymbol{\mu}_{x_j}$  is the vector of means over each dimension of

$\mathbf{y}_j$  defined for a given class  $x_j$  and  $\boldsymbol{\Sigma}_{x_j}$  is the corresponding covariance matrix.

Substituting (5) and (7) into (6), then taking the negative log, provides a convenient form of the posterior probability as the objective function to be minimized. Given the set of possible classes  $\Omega = \{1, 2, \dots, q\}$  where each  $x_j$  may take values 1 to  $q$ , the full negative log-likelihood function under the mean-field approximation takes the form:

$$L(\mathbf{x}|\mathbf{y}) = - \sum_{j=1}^s \log p(\mathbf{y}_j|x_j, \boldsymbol{\mu}_{x_j}, \boldsymbol{\Sigma}_{x_j}) - \sum_{j=1}^s \log p(x_j|\beta) \quad (8)$$

where the voxel-wise priors are given by:

$$p(x_j|\beta) = \frac{e^{-U(x_j|\beta)}}{\sum_{i=1}^q e^{-U(x_j=i|\beta)}} \quad (9)$$

### B. The Hybrid Fuzzy C-Means MRF Model

Fuzzy c-means clustering is a simple and commonly used algorithm for separating a set of  $s$  voxels into  $q$  partitions by minimizing the objective function:

$$L_\lambda = \sum_{j=1}^s \sum_{i=1}^q R_{i,j}^\lambda D_{i,j} \quad (10)$$

The ‘dissimilarity function’  $D_{i,j}$  describes the likelihood of voxel  $j$  belonging to a given class  $i$ , and  $R_{i,j}$  is the ‘fuzzy membership function’, which describes the certainty of  $j$  belonging to each class  $i$  and exhibits the following properties:

$$R_{i,j} = ]0, 1[ \quad \sum_{i=1}^q R_{i,j} = 1 \quad (11)$$

The parameter  $\lambda$  controls the level of fuzziness. Following the rationale set out by Chatzis *et al* [20], it is instructive to consider a variant of the above function which is regularized by relative entropy [22]:

$$L_\lambda = \sum_{j=1}^s \sum_{i=1}^q R_{i,j} D_{i,j} + \lambda \sum_{j=1}^s \sum_{i=1}^q R_{i,j} \log \left( \frac{R_{i,j}}{P_{i,j}} \right) \quad (12)$$

This version of the fuzzy objective function reverts to hard partitioning when  $\lambda=1$  and the fuzzy membership function is redefined with  $R_{i,j} = \{0, 1\}$ . In this special case, it is possible to make a direct comparison with equation (8) and thus to define:

$$D_{i,j} = -\log p(\mathbf{y}_j|x_j = i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (13)$$

$$P_{i,j} = p(x_j = i|\beta) \quad (14)$$

In this way it is possible to incorporate the MRF multivariate conditional distribution, as well as the MRF priors containing the desired contextual information, into the clustering framework provided by the fuzzy c-means objective function. This hybrid model can be iteratively minimized over the parameters  $\beta$ ,  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  in order to find an optimal solution for  $R$ ; the hard classification field  $\mathbf{x}$  is then defined simply as the class for which  $R$  is maximized at each voxel.

### C. Application To Hyperpolarized $^{13}\text{C}$ Data

The novel approach proposed in this paper is to adapt and apply the hybrid fuzzy clustering MRF framework described above to the problem of segmenting the 5D data produced by hyperpolarized  $^{13}\text{C}$  MRSI. Although the exact form of the data may differ depending on the imaging sequence used, it will in general consist of a set of 2D (single slice) or 3D (multi-slice) images, acquired at multiple time points and at multiple frequencies. The spectral dimension may either be acquired as a continuous spectrum, or at a number of discrete frequencies corresponding to target metabolites. The second case is considered here, which provides a separate series of images for each metabolite.

Four metabolites are considered; pyruvate ( $P$ ), lactate ( $L$ ), alanine ( $A$ ) and bicarbonate ( $B$ ), therefore at each voxel  $j$  in the 3-dimensional imaging volume, the observed intensity vector  $\mathbf{y}_j$  is given by:

$$\mathbf{y}_j(t) = [P_j(t), L_j(t), A_j(t), B_j(t)] \quad (15)$$

To reduce the dimensionality of  $\mathbf{y}_j(t)$  from  $[4 \times T]$  to  $[4 \times 1]$ , where  $T$  is the number of imaging time points, the data was time-averaged. Only images from time points at which the total carbon signal present was greater than the minimum value plus 0.2 times the standard deviation were included in the time average. This protocol effectively removed noise images whilst retaining as much useful signal as possible. Based on the animal data, total carbon SNR was approximately 4 at this cut-off with individual metabolite SNRs considerably lower. The overall objective of the segmentation was to automatically identify regions of differing metabolism, therefore the set of classes  $\Omega = \{1, 2, 3, 4, 5, 6\}$  into which voxels were sorted was defined as follows: 1. background (no pyruvate inflow); 2. no conversion of pyruvate; 3. low conversion rate of pyruvate into lactate; 4. high conversion rate of pyruvate into lactate; 5. conversion of pyruvate into alanine; 6. conversion of pyruvate into bicarbonate. If at any point during the iterative minimization process membership of one of these classes falls to zero, then it is removed from further calculations and  $q \rightarrow q - 1$ .

In conventional color image segmentation, a multivariate Gaussian conditional distribution, as in equation (7), may be used to describe the separate intensities from each of the relevant dimensions of a chosen color space; e.g. red, green and blue. Analogous with this approach, the time-averaged data here is treated as a single image with intensity contributions from four different ‘colors’, i.e. the four different metabolites. The conditional distribution at each voxel is then modelled by (7), where  $\mathbf{y}_j = [P_j, L_j, A_j, B_j]$ ,  $\boldsymbol{\mu}_i$  is the vector of corresponding means for each metabolite in relation to class  $i$  and  $\boldsymbol{\Sigma}_i$  is their covariance matrix. For a given voxel  $j$ , its weighting in the calculation of the noise parameters  $\boldsymbol{\mu}_i$  and  $\boldsymbol{\Sigma}_i$  is dependant on its fuzzy membership to class  $i$  as given by  $R_{i,j}$ .

The energy function  $U(x_j|\beta)$  in equation (9) is comprised of three parts incorporating the spatial, temporal and parametric prior information respectively:

$$U(x_j|\beta) = -\beta \{U_S(x_j) + U_T(x_j) + U_R(x_j)\} \quad (16)$$

It is in general possible to incorporate any number of terms into the energy function; the three included in this work are described below.

1) *Spatial Priors*: The first term in the energy function  $U_S(x_j)$  describes the spatial dependency of the MRF system, and the connectivity between voxels. An 8 or 10-connectivity network is used depending on whether the image to be segmented is 2D or 3D as depicted in figure 1. The spatial energy is calculated as follows:

$$U_S(x_j) = -\frac{\sum_{l \in c_j} \delta(x_j - x_l)}{\text{card}(l \in c_j)} \quad (17)$$

where  $\delta$  is the Dirac delta function defined as follows:

$$\delta = \begin{cases} 1 & \text{if } x_j = x_l \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

The function *card* is simply the count of voxels in the clique  $c_j$ . As  $x_j$  takes each class label  $1, 2, \dots, q$ ,  $U_S$  will take a maximum value of 1 when the labels of all voxels in clique  $c_j$  match that of  $x_j = i$  and a minimum value of 0 when no labels coincide. For the purpose of this calculation, the hard class labels of each voxel in the clique are defined as that which maximizes the current value of the fuzzy membership function:

$$x_l = \arg \max_{i=1}^q (R_{i,l}) \quad (19)$$

This results in imposing a penalty on  $x_j = i$  for not being of the same class as its immediate surroundings, therefore reducing the influence of random noise and fulfilling the requirement that voxels containing similar tissue types or metabolic features are more likely to be adjacent than separate. The spatial priors also prevent small regions of noise being picked up as false positives by making it energetically unfavorable to have many small ROIs.

2) *Temporal Priors*: Although the algorithm uses a time-averaged version of the data, the temporal information is preserved through its incorporation into the prior on each voxel. The changes in the local concentrations of each of the four metabolites of interest can be described by the following set of ordinary differential equations:

$$\frac{dP}{dt} = I(t) - (k_{PL} + k_{PA} + k_{PB})P + k_{LP}L - \rho_P P \quad (20)$$

$$\frac{dL}{dt} = -k_{LP}L + k_{PL}P - \rho_L L \quad (21)$$

$$\frac{dA}{dt} = k_{PA}P - \rho_A A \quad (22)$$

$$\frac{dB}{dt} = k_{PB}P - \rho_B B \quad (23)$$

where  $I(t)$  is the pyruvate inflow profile,  $k_{AB}$  is the rate of metabolic conversion from metabolite  $A$  to metabolite  $B$ , and  $\rho_A$  is the rate of polarization loss of metabolite  $A$ . Values for  $\rho$  can be estimated from the known thermal decay rates and imaging parameters, whereas the  $k$  terms are unknowns to be solved. Nevertheless, the solutions to these equations have distinct expected forms, which are of similar shape regardless of the exact values of the conversion rates.

Measured temporal profiles for the relative concentrations of each metabolite at voxel  $j$  are stored in  $\mathbf{y}_j(t)$ . To assess

whether these time profiles are likely to represent real signal, their shape is compared to a set of trial functions which form the solutions to the above differential equations [11]. Six trial functions are used for each metabolite, varying the  $k_{AB}$  and, for pyruvate, the inflow speed contained in the function  $I(t)$ . Both the measured signal and trial functions are normalized to  $[0,1]$ , the positions of the peaks are aligned and the  $R^2$  is calculated to test goodness of fit to each of the six trial functions. The highest  $R^2$  is then selected for inclusion in the temporal prior function.

For a given voxel  $j$ , the priors for each class  $i$  for metabolite  $X$  can then be defined as follows:

$$P_{X,i,j} = \frac{f_i(R_{X,j}^2)}{\sum_{i=1}^q f_i(R_{X,j}^2)} \quad (24)$$

Here, the function  $f_i(R_{X,j}^2)$  takes two different forms dependant on the expectation of that metabolite in each class. When no signal is expected from a particular metabolite,  $f_{X,j}^- = 1 - (R_{X,j}^2)^3$ , whereas  $f_{X,j}^+ = (R_{X,j}^2)^3$  if signal is required for inclusion of a voxel into a particular class. The cube is used to produce a sharper distinction between signal and noise. Explicitly, where each column in the row vector is a class 1 to 6, then the following apply:

$$P_{P,j} = [f_{P,j}^-, f_{P,j}^+, f_{P,j}^+, f_{P,j}^+, f_{P,j}^+, f_{P,j}^+] \quad (25)$$

$$P_{L,j} = [f_{L,j}^-, f_{L,j}^-, f_{L,j}^+, f_{L,j}^+, f_{L,j}^-, f_{L,j}^-] \quad (26)$$

$$P_{A,j} = [f_{A,j}^-, f_{A,j}^-, f_{A,j}^-, f_{A,j}^+, f_{A,j}^+, f_{A,j}^-] \quad (27)$$

The polarization decay rate of bicarbonate is too fast to define appropriate test functions, therefore only pyruvate, lactate and alanine are included in the temporal prior. Finally, the normalized contribution to the energy function is given by:

$$U_T(x_j = i) = -\frac{P_{P,i,j}P_{L,i,j}P_{A,i,j}}{\sum_{i=1}^q P_{P,i,j}P_{L,i,j}P_{A,i,j}} \quad (28)$$

3) *Parametric Priors*: The final contribution to the energy function considers ratios of the metabolites observed at each voxel and also utilizes the full temporal data set. Using the time curve for each metabolite, it can be shown that the ratio of Areas Under the Curve (AUC) for lactate and pyruvate is equal to [10]:

$$R_{LP} = \frac{AUC_{Lactate}}{AUC_{Pyruvate}} = \frac{k_{PL}}{k_{LP} + \rho_L} \quad (29)$$

What is striking about this result is that it is completely independent of the pyruvate inflow function  $I(t)$ . Furthermore, since  $k_{PL} \gg k_{LP}$  and  $\rho_L$  is expected to be consistent, this simple metric correlates reasonably well with the forward rate conversion constant  $k_{PL}$  [11].

AUC ratio maps were calculated for lactate, alanine and bicarbonate, relative to pyruvate. Values  $>2$  were set to this maximum value, the upper limit expected from a true signal, to avoid random noise in the background region skewing the results. Each map was then normalized by dividing by its maximum value. Individual metabolite priors were then constructed from the ratio maps:

$$P_{XP,i,j} = \frac{f_i(R_{XP,j})}{\sum_{i=1}^q f_i(R_{XP,j})} \quad (30)$$

The following main changes to the functions  $f_i$  used in the temporal case are noted. Firstly, the background class 1 is noise so does not depend on the AUC ratios; this was set such that  $P_{XP,1,j} = 1/6$  after normalization. Secondly, the low lactate class, class 3, requires a negative quadratic response as the lactate/pyruvate AUC increases. The normalized ratiometric contribution to the energy function is given by:

$$U_R(x_j = i) = -\frac{P_{LP,i,j}P_{AP,i,j}P_{BP,i,j}}{\sum_{i=1}^q P_{LP,i,j}P_{AP,i,j}P_{BP,i,j}} \quad (31)$$

Although there are multiple parametric approaches which could be included as prior information, the AUC ratio was chosen because of its rapid calculation speed, insensitivity to pyruvate inflow profile, full usage of temporal data and its dual treatment of metabolite signals, which separates metabolic conversion from perfusion. It is the latter of these which provides the most important additional information for the segmentation process.

#### D. Initialisation

Given that each class in the  $^{13}\text{C}$  segmentation is required to exhibit specific pre-defined metabolic properties, initialization was performed using a Bayesian maximum posterior likelihood method using thresholded, time-averaged metabolite images to estimate initial class means and variances and incorporating previously calculated ratiometric and temporal maps for the Bayesian priors. Estimates for the background class noise parameters  $\mu_1$  and  $\sigma_1$ , the 4-metabolite mean and standard deviation vectors respectively, were obtained from the final metabolite images in the time series, by which time all the signal has decayed. These parameters were then used to determine whether each voxel in the time-averaged data set, with measured signal intensity  $\mathbf{y}_j = [y_{P,j}, y_{L,j}, y_{A,j}, y_{B,j}]$ , contained signal significantly above the noise level.

The following initial thresholding was used to find mean and variance estimates for each class, where  $L_{max}$  is the maximum lactate signal intensity in the time-averaged image:

1. Background class:  $y_{P,j} < \mu_{P,1} + 5\sigma_{P,1}$
2. No exchange:  $j \notin 1, 5, 6$  and  $y_{L,j} < \mu_{L,1} + 5\sigma_{L,1}$
3. Low lactate exchange:  $j \notin 1, 4$  and  $y_{L,j} \geq \mu_{L,1} + 5\sigma_{L,1}$
4. High lactate exchange:  $j \notin 1$  and  $y_{L,j} \geq \mu_{L,1} + 5\sigma_{L,1} + 0.8(L_{max} - \mu_{L,1} + 5\sigma_{L,1})$
5. Conversion to alanine:  $j \notin 1, 3, 4$  and  $y_{A,j} \geq \mu_{A,1} + 5\sigma_{A,1}$
6. Conversion to bicarbonate:  $j \notin 1, 3, 4, 5$  and  $y_{B,j} \geq \mu_{B,1} + 5\sigma_{B,1}$

From these initial voxel groupings, mean  $\mu'_i$  and covariance  $\Sigma'_i$  are calculated for each class as the inputs into a multivariate normal distribution,  $\mathcal{N}(\mu'_i, \Sigma'_i)$ . The final initialization is given by the class which for each voxel maximizes the Bayesian posterior:

$$x'_j = \arg \max_{i=1}^q \frac{P'_{i,j} \mathcal{N}_j(\mu'_i, \Sigma'_i)}{\sum_{k=1}^q P'_{k,j} \mathcal{N}_j(\mu'_k, \Sigma'_k)} \quad (32)$$

Where  $P'_{i,j}$  is an initial prior constructed in the same way as  $P_{i,j}$  but with the spatial term omitted. An initial estimate of 3 was used for the temperature parameter  $\beta$  and the fuzziness parameter  $\lambda$  was set to 3 throughout. This value was chosen

to provide a satisfactory trade-off between underestimating the region size and producing false positives;  $\lambda$  may be optimized further when additional clinical data becomes available.

### E. Iterative Determination Of The Hidden Field $\mathbf{x}$

Having defined the constituent parts of the objective function (12) and the procedure for initialization in the context of the segmentation of  $^{13}\text{C}$  imaging data, it is possible to define the iterative process with which the hidden field  $\mathbf{x}$  is determined. The procedure for the most part follows that described by Chatzis *et al* [20] and is outlined below.

The full objective function to be minimized at each iteration  $w$  is given by:

$$L_\lambda^{(w)} = \frac{1}{2} \sum_{j=1}^s \sum_{i=1}^q R_{i,j}^{(w)} \{n \log(2\pi) + \log(|\Sigma_i^{(w)}|) + (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)})^T \Sigma_i^{(w)-1} (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)})\} + \lambda \sum_{j=1}^s \sum_{i=1}^q R_{i,j}^{(w)} \log \left( \frac{R_{i,j}^{(w)}}{P_{i,j}^{(w)}} \right) \quad (33)$$

Therefore at each step, the priors  $P_{i,j}$ , the noise parameters  $\boldsymbol{\mu}_i$  and  $\Sigma_i$ , the fuzzy membership function  $R_{i,j}$  and the temperature parameter  $\beta$  must all be updated. Given some initial estimate of the hidden field  $\mathbf{x}^{(w)}$  at iteration  $w$ , the priors  $P_{i,j}^{(w)}$  are calculated as follows:

$$P_{i,j}^{(w)} = \frac{e^{\beta^{(w)} \{U_S(x_j=i|c_j^{(w)}) + U_T(x_j=i) + U_R(x_j=i)\}}}{\sum_{k=1}^q e^{\beta^{(w)} \{U_S(x_j=k|c_j^{(w)}) + U_T(x_j=k) + U_R(x_j=k)\}}} \quad (34)$$

Here  $c_j^{(w)}$  denotes the current estimate of the hard class labels for voxels in the clique  $c_j$ . These are determined by choosing the value for each  $x_l \in c_j$  which maximizes the fuzzy membership function  $R_{i,j}^{(w)}$  according to equation (19). Only the spatial term in the energy function  $U(x_j|\beta)$  is updated, along with the temperature parameter  $\beta$ , at each iteration.

Given initial estimates of the noise parameters at  $w$ ,  $D_{i,j}^{(w)}$  is defined as the negative-log of the multivariate Gaussian conditional distribution, equation (7):

$$D_{i,j}^{(w)} = \frac{1}{2} \{n \log(2\pi) + \log(|\Sigma_i^{(w)}|) + (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)})^T \Sigma_i^{(w)-1} (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)})\} \quad (35)$$

The value of the fuzzy membership function can now be computed for the subsequent iteration  $w+1$ :

$$R_{i,j}^{(w+1)} = \frac{P_{i,j}^{(w)} e^{-\frac{D_{i,j}^{(w)}}{\lambda}}}{\sum_{k=1}^q P_{k,j}^{(w)} e^{-\frac{D_{k,j}^{(w)}}{\lambda}}} \quad (36)$$

This function is derived by minimizing the objective function  $L_\lambda$  with respect to  $R_{i,j}$  subject to the constraint:

$$\sum_{i=1}^q R_{i,j} = 1 \quad \forall j \in S \quad (37)$$

This can be achieved by introducing a Lagrange multiplier  $\psi_j$  to enforce the constraint at each voxel  $j \in S$  and then setting the partial derivative of the resulting Lagrangian equal to zero:

$$\frac{\partial}{\partial R_{i,j}} \{L_\lambda - \sum_{j=1}^s \psi_j (\sum_{i=1}^q R_{i,j} - 1)\} = 0 \quad (38)$$

Similarly, setting the partial derivative of  $L_\lambda$  with respect to each of  $\boldsymbol{\mu}_i$  and  $\Sigma_i$  to zero, yields the following update functions for the noise parameters at  $w+1$ :

$$\boldsymbol{\mu}_i^{(w+1)} = \frac{\sum_{j=1}^s R_{i,j}^{(w)} \mathbf{y}_j}{\sum_{j=1}^s R_{i,j}^{(w)}} \quad (39)$$

$$\Sigma_i^{(w+1)} = \frac{\sum_{j=1}^s R_{i,j}^{(w)} (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)}) (\mathbf{y}_j - \boldsymbol{\mu}_i^{(w)})^T}{\sum_{j=1}^s R_{i,j}^{(w)}} \quad (40)$$

The final step is to find the new value for the inverse temperature  $\beta^{(w+1)}$  according to:

$$\beta^{(w+1)} = \arg \max_{\beta} \sum_{j=1}^s \sum_{i=1}^q R_{i,j}^{(w)} \log(P_{i,j}^{(w)}) \quad (41)$$

Maximizing  $\beta$  in this way fine-tunes the system to become maximally sensitive to the value of the energy function  $U(\mathbf{x}|\beta)$ , therefore globally forcing the system into the state of lowest energy based on the definition of  $U(\mathbf{x}|\beta)$ . There is now sufficient information to calculate the updated value of the objective function  $L_\lambda^{(w+1)}$ . This was compared at each step with the value of  $L_\lambda^{(w)}$  and convergence was defined as having reached a stable value of  $\Delta L_\lambda = |L_\lambda^{(w+1)} - L_\lambda^{(w)}|/L_\lambda^{(w)} < c_T$  with  $c_T$ , the convergence threshold, set at 0.1.

In summary, the full iterative procedure for the segmentation of  $^{13}\text{C}$  imaging data is as follows:

- 1: **Input:** 5D  $^{13}\text{C}$  imaging data
- 2: Calculate temporal coherence maps and value of temporal energy function  $U_T(\mathbf{x})$  for each voxel  $j$  in the 3D volume
- 3: Calculate AUC ratio maps and parametric energy function  $U_R(\mathbf{x})$
- 4: Use initialization procedure to establish initial estimates of  $R_{i,j}^{(0)}$ ,  $x_j^{(0)}$ ,  $\boldsymbol{\mu}_i^{(0)}$  and  $\Sigma_i^{(0)}$ , and thus calculate  $L_\lambda^{(0)}$
- 5: Set iteration number  $w = 1$
- 6: **while**  $\Delta L_\lambda > c_T$  **or**  $w < 150$  **do**
- 7: Calculate the new temperature parameter  $\beta^{(w)}$  using  $R_{i,j}^{(w-1)}$  and  $P_{i,j}^{(w-1)}$  with equation (41)
- 8: Update the values of fuzzy membership function  $R_{i,j}^{(w)}$  from  $D_{i,j}^{(w-1)}$  and  $P_{i,j}^{(w-1)}$  using the formula given by (36)
- 9: Calculate the updates to the noise parameters  $\boldsymbol{\mu}_i^{(w)}$  and  $\Sigma_i^{(w)}$  using (39) and (40) respectively
- 10: Calculate the current realisation of  $x_j^{(w)}$  from fuzzy membership function  $R_{i,j}^{(w)}$  using equation (19)
- 11: Calculate the current value of voxel-wise priors  $P_{i,j}^{(w)}$  using  $x_j^{(w)}$  and  $\beta^{(w)}$
- 12: Find the new value of the objective function  $L_\lambda^{(w)}$
- 13: Calculate  $|L_\lambda^{(w)} - L_\lambda^{(w-1)}|/L_\lambda^{(w-1)}$  and test for convergence

14:  $w \rightarrow w + 1$   
 15: **end while**  
 16: **Return:** Hidden field  $\mathbf{x}$  of voxels segmented into  $q$  classes

The fuzzy MRF algorithm was written and implemented in Matlab (MathWorks, Natick, MA, USA). The algorithmic runtime is dependent on the image size, number of image slices and number of classes detected during the initialization process. For  $128 \times 128$  single slice images, the computational runtime is 42-117 s for 4-6 classes as measured with a 2.2 GHz Intel Core i7 processor and 8 MB RAM. This stated time is for processing the raw reconstructed 5D image tensor, including all parametric map calculations. Computational complexity scales linearly with the number of image slices but in practice, this is rarely greater than 4. The algorithm was not optimized for speed, so further reductions in this runtime will be possible.

### III. DATA ACQUISITION AND ANALYSIS

#### A. In Silico Data Generation

In order to quantitatively assess the accuracy of the segmentation algorithm, an *in silico*  $^{13}\text{C}$  dataset consisting of 27 sets of images was generated. Each set had the following properties, chosen to resemble those of the animal data: 2 spatial dimensions, with  $128 \times 128$  voxels in each image; 16 time points and 4 s temporal resolution; 4 metabolites. Each set contained a regions corresponding to classes 1-6. Metabolite signals were generated by solving equations (20-23) at each voxel location over time with the following parameterization;  $\rho_P = \rho_L = \rho_A = 0.064$ ;  $\rho_B = 0.131$ ;  $k_{PL} = 10k_{LP}$ ; and the pyruvate inflow function:

$$I(t) = k_I \left( \frac{1}{1 + e^{-(t-t_s)}} - \frac{1}{1 + e^{-0.4(t-t_e)}} \right) \quad (42)$$

Where  $t_s$  is the inflow start time and  $t_e$  the inflow end time. These were set to increase gradually with distance from a main 'supply vessel' with a fixed difference of 4 s, and the inflow rate  $k_I$  decreased from a maximum value at this same point. To evaluate the performance of the segmentation under different conditions, three parameters were varied across the dataset. Firstly, to assess the detection of differently sized metabolic regions of interest, the mean radius of the high lactate conversion region was ascribed values of 3, 6 and 9 voxels and the surrounding low conversion zone mean radius set to be 2.4 times larger. Alanine and bicarbonate regions had a variable size between these two. Secondly, to assess the effects of variable metabolic conversion rates, each  $k_{PX}$  was varied across the entire image, decreasing outwards from a centre point which took values;  $k_{PL,max} = 0.01, 0.03, 0.05$ ;  $k_{PA,max} = 0.08, 0.012, 0.016$ ;  $k_{PB,max} = 0.01, 0.014, 0.018$ .  $k_{PX}$  maps were set to zero everywhere except within the test regions. For the lactate map, the outer low conversion region had a multiplier of 0.14 applied to the  $k_{PL}$ . Finally, the effect of varying the SNR, qualitatively equivalent to differing initial polarization, was investigated by varying the inflow rate  $k_I$  over arbitrary units of 9000, 12000 and 16000, corresponding to a mean pyruvate SNR of approximately 20, 30 and 40 within the high conversion zone. To mimic the partial voluming and noise corruption characteristic of *in vivo* data, image

resolution was first reduced to  $32 \times 32$ , the nominal matrix resolution of the imaging. Random Gaussian noise with a variable random mean of 220-410 and variance of  $1.1-6.3 \times 10^4$  was then applied to each image before artificially increasing the resolution to  $128 \times 128$ . This procedure produced signal spread and patterns of noise similar to the *in vivo* data.

#### B. In Vivo Data Acquisition

The segmentation algorithm was tested on previously published data from four tumor-bearing adult rats with subcutaneous mammary adenocarcinomas [23], [24]. Briefly, [1- $^{13}\text{C}$ ]Pyruvic acid was hyperpolarized to approximately 25% in a Hypersense hyperpolarizer (Oxford Instruments). The final solution contained 80 mM [1- $^{13}\text{C}$ ]pyruvate at pH 7.6 and physiological temperature and osmolarity. This was injected into a tail vein inside the MRI scanner at a rate of approximately 0.2 ml/s and at a dose of 2.5 ml/kg and dynamic imaging was performed from the time of injection.

Animals were imaged in a clinical 3 T MRI system (Signa HDx, GE Healthcare) using a  $^{13}\text{C}$ - $^1\text{H}$  multinuclear bird-cage coil. IDEAL spiral CSI acquisition [25] was performed through four axial 10 mm slices over 1 min with a temporal resolution of 4 s and a flip angle of  $10^\circ$ . Other parameters were: repeat time (TR) 500 ms; field of view 80 mm and nominal matrix resolution  $32 \times 32$ . Gaussian k-space filtering was applied during post-processing, interpolating from a real resolution of about  $5 \times 5 \text{ mm}^2$  to an effective image resolution of  $0.625 \times 0.625 \text{ mm}^2$ . For anatomical reference, standard gradient echo proton images were acquired from the same slice geometry and FOV (resolution  $256 \times 256$ , slice thickness 3 mm, spacing 7 mm, echo time (TE) 10 ms, TR 500 ms).

### IV. RESULTS

#### A. Statistical Results On In Silico Data

The fuzzy MRF segmentation algorithm was applied to each of the 27 4-dimensional *in silico* image sets described above. The data set was designed to test the limits of the algorithm and assess its ability to distinguish two separate regions: region A was defined as the total area of lactate production (class 3 + class 4); region B, high lactate production region only (class 4). In both cases, all other classes are considered background with the known input  $k_{PL}$  map for each *in silico* image providing the gold standard for comparison. Performance of the algorithm was compared to that of thresholding the time-averaged lactate image at 40, 50, 60, 70 and 80% of the maximum lactate value; a comparison standard chosen for its clinical applicability. The same time-averaging as for the MRF segmentation was applied. To quantify performance, the Precision (positive predictive value) and Recall (sensitivity) for this binary case were calculated; these can be combined to produce the balanced F-score for the segmentation as follows:

$$Precision = \frac{tp}{tp + fp} \quad (43)$$

$$Recall = \frac{tp}{tp + fn} \quad (44)$$

$$F = \frac{2PR}{P + R} \quad (45)$$

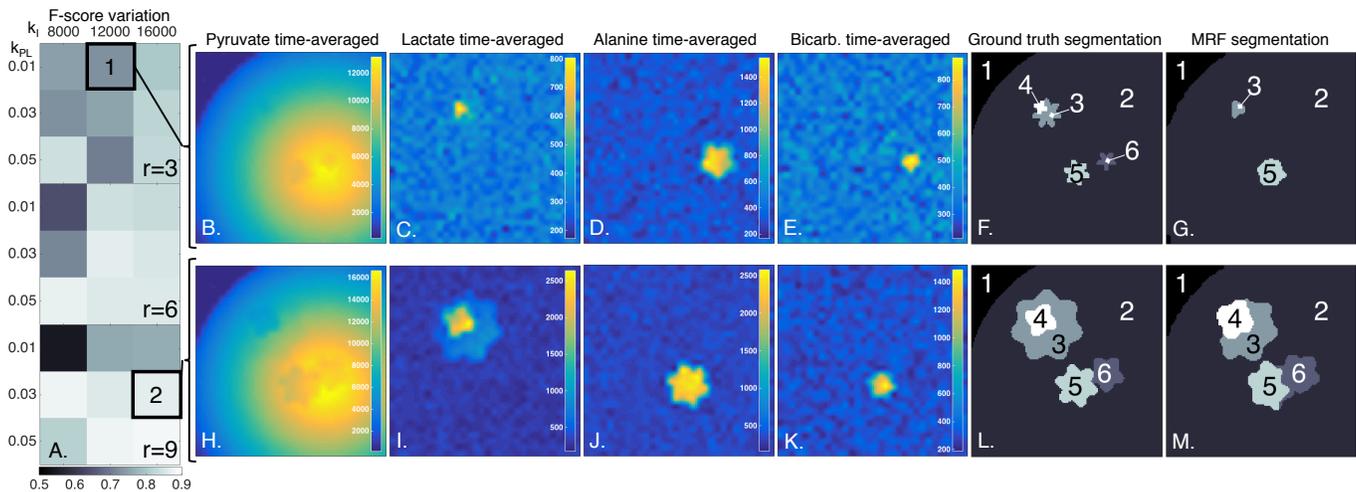


Fig. 2. Performance of the MRF algorithm on *in silico* data. Part A. Variation in F-score as the inflow constant  $k_I$  and rate constant  $k_{PL}$  are increased for a mean inner lesion radius of 3 (top), 6 (middle) and 9 (bottom) image voxels. Further images from two representative sets labelled 1 (B-E) and 2 (F-I) on part A are shown. These are: the time-averaged pyruvate (B, F); time-averaged lactate (C, G); ground-truth segmentation (D-H) and MRF segmentation (E, I). The numbers 1-4 on the latter two of these denote the class labels 1-4.

TABLE I  
COMPARISON OF PERFORMANCE BETWEEN MRF ALGORITHM AND DIFFERENT THRESHOLDING LEVELS (AS % OF MAXIMUM LACTATE) FOR THE DETECTION OF TWO SEPARATE REGIONS A AND B.

		MRF	40%	50%	60%	70%	80%
F-score (A)	Mean	0.694	0.490	0.402	0.329	0.251	0.155
	Min	0.284	0.028	0.048	0.247	0.182	0.081
	Max	0.938	0.840	0.630	0.548	0.462	0.317
F-score (B)	Mean	0.895	0.758	0.864	0.870	0.788	0.571
	Min	0.613	0.006	0.011	0.120	0.692	0.355
	Max	1.000	1.000	0.997	0.982	0.949	0.844
CE (A)	Mean	43.70	497.8	214.4	89.98	85.65	91.54
	Min	12.09	29.87	62.64	63.32	72.73	82.25
	Max	83.47	6795	3662	344.6	90.02	95.80
CE (B)	Mean	27.30	2157	700.5	71.95	34.814	59.42
	Min	0.000	0.000	0.694	3.472	9.722	28.57
	Max	126.5	32030	17230	1463	47.14	78.45

where  $tp$  stands for “true positives”, the count of voxels correctly identified as tumor when compared to the gold standard,  $fp$  are “false positives” and  $fn$  are “false negatives”. The F-score for the full segmentation into six classes was calculated using the averages of the precision and recall over each class, with instances where the algorithm failed to detect a region still included in this score. As an additional measure of accuracy, the classification error (CE) was calculated in addition to the F-score. It is defined as:

$$CE = \frac{100(fp + fn)}{\text{no. pixels in region}} \quad (46)$$

where  $fn$  are “false negatives”, pixels incorrectly labelled as background, and a CE closer to zero denotes a more accurate segmentation. The results from calculating both metrics over each region A and B for all segmentations are shown in table I. The mean values for all 27 images, as well as the maximum and minimum in each set, are given.

TABLE II  
F-SCORES AND CLASSIFICATION ERRORS FOR EACH OF THE SIX CLASSES, EXCLUDING INSTANCES WHERE A CLASS IS EMPTY. EXCLUSIONS BY CLASS; CLASS 4, 8; CLASS 5, 1; AND CLASS 6, 7 OUT OF 27 IMAGES.

F-score	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Mean	0.947	0.975	0.475	0.400	0.583	0.812
Min	0.919	0.932	0.011	0.000	0.527	0.714
Max	0.974	0.991	0.598	0.552	0.639	0.921
CE	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Mean	10.41	5.053	58.84	75.91	43.62	45.31
Min	5.157	1.839	20.14	31.31	15.99	16.92
Max	17.49	14.37	246.9	136.7	79.41	80.23

In every instance, the MRF segmentation performs, on average, better than thresholding for binary segmentation of both tumor regions A and B. In almost every case, the MRF method produced the highest minimum and maximum F-scores, as well as the lowest minimum and maximum classification errors, suggesting that it is a more robust methodology than thresholding the lactate image at any level. Exceptions to this are the maximum lower CEs produced by the high threshold levels for the small region B. In general, both the MRF algorithm and the reference thresholding methods were better able to detect region B than A due to class 3 being very close to the noise floor as evidenced by figure 2C and I. Considering the full segmentation into 6 classes, the mean F-score was 0.798 with a range of 0.547-0.882. Figure 2A illustrates the variation in calculated F-score as the simulation parameters are varied. Low scores indicate that one or more classes failed to be detected. In 8 instances, only the inner high lactate region was detected and marked as class 3 with class 4 empty. Although this is damaging to the statistical results, it is technically not an incorrect result to see low lactate production labelled as such; the  $k_{PL}$  was just 0.01 in all but one of these cases. In a clinical setting, it is advantageous that the algorithm can correctly detect a area of low conversion

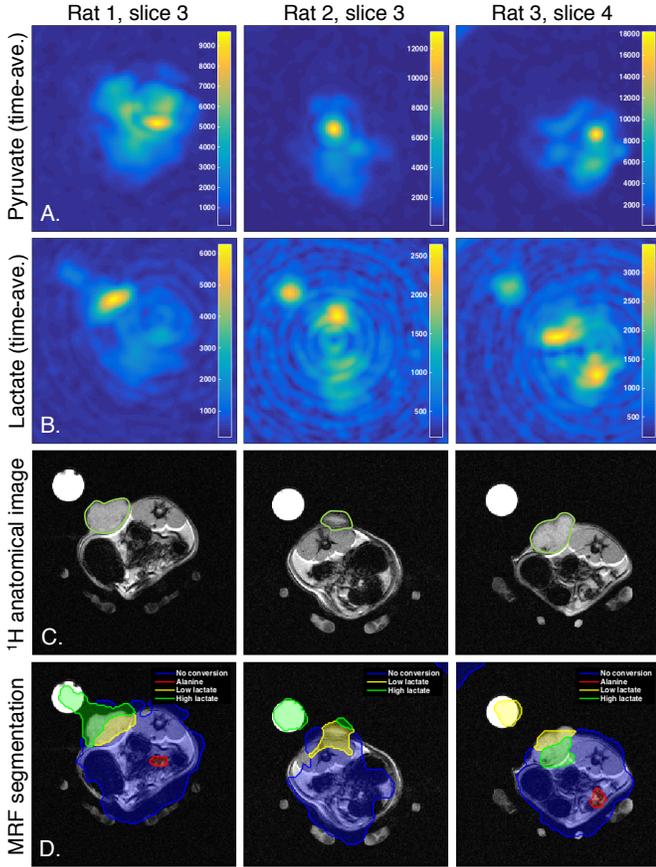


Fig. 3. A single axial image slice centred on the tumor is shown for 3 rats. A. Time-averaged pyruvate over each slice; B. Time-averaged lactate; C. Anatomical reference image with tumor location outlined in green; D. MRF segmentation results showing background class 1 (no color) no-conversion class 2 (blue), low conversion class 3 (yellow), high conversion to lactate class 4 (green) and alanine class 5 (orange). Class 6 is empty. Lactate phantom in upper left corner.

to lactate whilst separating the region from a noisy, very low conversion background. Background lactate is often present, particularly in organs such as the brain or in the muscles, but should be ignored by the segmentation. There were no cases in which region B was not detected at all and no false positives. In 1 case the alanine class and in 7 cases the bicarbonate class failed to be detected. Lower sensitivity to these metabolites has been chosen to avoid false positives which, although not detected in the *in silico* data, are common for *in vivo* data. Table II shows class by class F-score and CE results. Empty classes are omitted from the mean calculations.

### B. In Vivo Results

Application of the MRF segmentation algorithm to *in vivo* imaging was demonstrated on 5D data from four rats with subcutaneous implanted tumors. Figure 3 shows the single-slice segmentation results for three rats alongside the corresponding  $^{13}\text{C}$ -pyruvate,  $^{13}\text{C}$ -lactate and  $^1\text{H}$  anatomical images. The same time-averaging is used for these images as for the segmentation algorithm. In each case, the tumor location is correctly detected and dual-labelled between low and high lactate classes demonstrating the feasibility of the algorithm

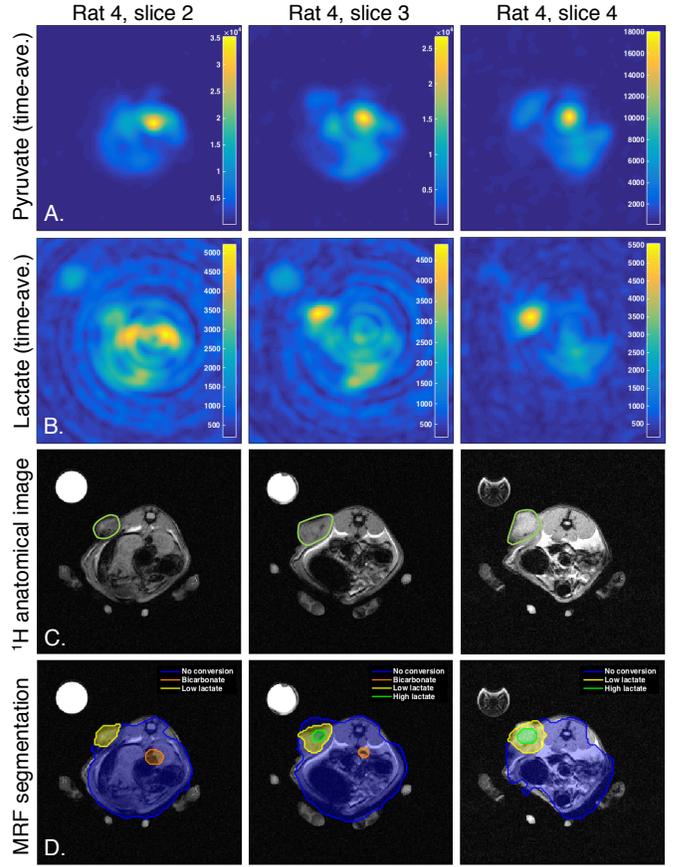


Fig. 4. Three consecutive tumor-bearing slices from rat 4 are shown, with slice 1 cranial in orientation. A. Time-averaged pyruvate over each slice; B. Time-averaged lactate; C. Anatomical reference image with tumor location outlined in green; D. MRF segmentation results showing background class 1 (no color) no-conversion class 2 (blue), low conversion class 3 (yellow), high conversion to lactate class 4 (green) and bicarbonate class 6 (orange). Class 5 is empty. Lactate phantom in upper left corner.

to identify tumor heterogeneity. The only false positives to be detected were the lactate phantoms; the close proximity to the large tumor in rat 1 has caused joining of these regions, highlighting a current limitation of the algorithm. Although the alanine detected in rat 1 is likely to be a false positive from an overlying artery, in rat 3 it points to real signal from the muscle. Four image slices were acquired for each animal, however only one rat had a tumor which spanned multiple slices allowing for a demonstration of the 3D capabilities of the algorithm. The resulting segmentation of the three tumor-containing slices for this rat are shown in figure 4. Poor image quality due to artefacts is a common problem with current  $^{13}\text{C}$  imaging methods; it is therefore crucial to demonstrate the ability of the proposed MRF algorithm to distinguish high contrast artefacts from real tissue metabolism. Despite the low spatial resolution and spiral artefact, particularly in the lactate image where the artefact intensity is comparable to that of the tumor, the algorithm performs well and correctly identifies the tumor in each slice with no false positives. A small area of bicarbonate production was also detected (shown in orange on figure 4D); while this may be real, it may also be artefactual, caused by signal overspill from the very high

pyruvate signal in the overlying artery. Finally, it is worth noting that the lactate phantom, seen as the bright circle in the top left of figure 4C-D, has been correctly classified as part of the background despite the increased lactate signal it generates. Taken together, the *in silico* and *in vivo* results suggest that the model is able to distinguish areas of low from high metabolism. Future correlation of hyperpolarized images with histology will determine how accurately this variation corresponds to true biological variation.

## V. CONCLUSION

In this paper, a fuzzy Markov random field approach has been used as the first example of a segmentation algorithm designed for handling 5D hyperpolarized  $^{13}\text{C}$  data. The method fully utilizes the multi-dimensional format of the data in order to classify each voxel into one of six distinct classes based on its metabolic characteristics. Conversion of  $^{13}\text{C}$ -pyruvate into downstream metabolites  $^{13}\text{C}$ -lactate,  $^{13}\text{C}$ -alanine and  $^{13}\text{C}$ -bicarbonate, in addition to no-conversion and no-perfusion regions, are considered in this work, however the approach could be applied to other hyperpolarized  $^{13}\text{C}$ -labelled molecules in the future. Unlike a hard classification algorithm, the suggested approach uses continuous fuzzy logic to effectively analyse the low resolution, noisy images by retaining information about the true contents of each voxel throughout the segmentation process. Contextual information on each voxel is handled by Bayesian priors describing spatial, temporal and ratiometric dependencies, whilst the use of a multivariate Gaussian conditional distribution allows image contrast from multiple spectral dimensions to be considered concurrently.

When tested on an *in silico* dataset which varied parameters such as the SNR, metabolic conversion rate constant and lesion radius, the MFR approach significantly out-performed the reference thresholding methodology at all five thresholding levels tested. There were no cases in which the algorithm failed to detect the region of high metabolic conversion of pyruvate to lactate and only a small number in which the low conversion region was not detected despite the lactate signal being barely visible above the noise floor in these image sets. Application to *in vivo* data from rats with subcutaneously implanted tumors demonstrated the ability of the algorithm to successfully segment regions of varying metabolic activity in three spatial dimensions. The anatomical tumor location was correctly identified in each occasion, whilst detection of high lactate signal from imaging artefacts was avoided.

Since the algorithm takes imaging parameters including the TR, temporal resolution and flip angle as inputs, it should be robust to changes in these parameters. Future work will look more closely at these and other dependencies by testing the algorithm for robustness against different spatial resolutions and image sequences, as well as developing the capability to effectively handle sequences with multiple or variable flip angles. The capability to automatically detect and exclude phantoms is a further point for development.

Hyperpolarized imaging with  $^{13}\text{C}$ -labelled compounds is an established research field and an emerging clinical technique

allowing metabolism to be characterized *in vivo*. Highly accurate quantification of the multidimensional data produced is an ongoing research objective and the success of this technique as a clinical tool depends on the ability to detect and quantify these changes in metabolism. ROI or VOI analysis is an integral part of this process and the fuzzy MRF approach presented here is capable of efficiently providing this analysis in a way that is repeatable and accurate. This novel methodology for analyzing pyruvate metabolism could be applied to other metabolites and clinical imaging in the future and provides a first solution to the problem of hyperpolarized  $^{13}\text{C}$  image segmentation.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Rolf Schulte (GE Global Research, Munich) and Prof. Markus Schwaiger (Technical University Munich, Munich) for providing the animal data analyzed in this work. CJD is jointly funded by the National Institute for Health Research (NIHR), Cambridge Biomedical Research Centre and GlaxoSmithKline (GSK). The authors acknowledge further research support from Cancer Research UK (C19212/A911376, C19212/A16628), the Cancer Research UK/Engineering and Physical Sciences Research Council Imaging Centre in Cambridge and Manchester, the CRUK Cambridge Centre and Cambridge Experimental Cancer Medicine Centre.

## REFERENCES

- [1] J. H. Ardenkjaer-Larsen, B. Fridlund, A. Gram, G. Hansson, L. Hansson, M. H. Lerche, R. Servin, M. Thaning, and K. Golman, "Increase in signal-to-noise ratio of  $>10,000$  times in liquid-state NMR," *Proceedings of the National Academy of Sciences*, vol. 100, no. 18, pp. 10 158–10 163, Feb. 2003, 00000 PMID: 12930897.
- [2] K. Golman, R. i. . Zandt, and M. Thaning, "Real-time metabolic imaging," *Proceedings of the National Academy of Sciences*, vol. 103, no. 30, pp. 11 270–11 275, Jul. 2006.
- [3] M. L. Zierhut, Y.-F. Yen, A. P. Chen, R. Bok, M. J. Albers, V. Zhang, J. Tropp, I. Park, D. B. Vigneron, J. Kurhanewicz, R. E. Hurd, and S. J. Nelson, "Kinetic modeling of hyperpolarized  $^{13}\text{C}_1$ -pyruvate metabolism in normal rats and TRAMP mice," *Journal of Magnetic Resonance*, vol. 202, no. 1, pp. 85–92, Jan. 2010.
- [4] S. E. Day, M. I. Kettunen, F. A. Gallagher, D.-E. Hu, M. Lerche, J. Wolber, K. Golman, J. H. Ardenkjaer-Larsen, and K. M. Brindle, "Detecting tumor response to treatment using hyperpolarized  $^{13}\text{C}$  magnetic resonance imaging and spectroscopy," *Nature Medicine*, vol. 13, no. 11, pp. 1382–1387, Nov. 2007, 00337.
- [5] S. J. Nelson, J. Kurhanewicz, D. B. Vigneron, P. E. Z. Larson, A. L. Harzstark, M. Ferrone, M. v. Crikkinge, J. W. Chang, R. Bok, I. Park, G. Reed, L. Carvajal, E. J. Small, P. Munster, V. K. Weinberg, J. H. Ardenkjaer-Larsen, A. P. Chen, R. E. Hurd, L.-I. Odegardstuen, F. J. Robb, J. Tropp, and J. A. Murray, "Metabolic Imaging of Patients with Prostate Cancer Using Hyperpolarized  $[1-^{13}\text{C}]$ Pyruvate," *Science Translational Medicine*, vol. 5, no. 198, pp. 198ra108–198ra108, Aug. 2013.
- [6] R. E. Hurd, Y.-F. Yen, J. Tropp, A. Pfefferbaum, D. M. Spielman, and D. Mayer, "Cerebral dynamics and metabolism of hyperpolarized  $[1-^{13}\text{C}]$ pyruvate using time-resolved MR spectroscopic imaging," *Journal of Cerebral Blood Flow and Metabolism*, vol. 30, no. 10, pp. 1734–1741, Oct. 2010.
- [7] D. J. Tyler, "Cardiovascular Applications of Hyperpolarized MRI," *Current Cardiovascular Imaging Reports*, vol. 4, no. 2, pp. 108–115, Apr. 2011.
- [8] C. H. Cunningham, J. Y. Lau, A. P. Chen, B. J. Geraghty, W. J. Perks, I. Roifman, G. A. Wright, and K. A. Connelly, "Hyperpolarized  $^{13}\text{C}$  Metabolic MRI of the Human Heart: Initial Experience," *Circulation Research*, p. CIRCRESAHA.116.309769, Sep. 2016.

- [9] T. Harris, G. Eliyahu, L. Frydman, and H. Degani, "Kinetics of hyperpolarized  $^{13}\text{C}_1$ -pyruvate transport and metabolism in living human breast cancer cells," *Proceedings of the National Academy of Sciences*, vol. 106, no. 43, pp. 18 131–18 136, Oct. 2009, 00075 PMID: 19826085.
- [10] D. K. Hill, M. R. Orton, E. Mariotti, J. K. R. Boulton, R. Panek, M. Jafar, H. G. Parkes, Y. Jamin, M. F. Miniotti, N. M. S. Al-Saffar, M. Belouche-Babari, S. P. Robinson, M. O. Leach, Y.-L. Chung, and T. R. Eykyn, "Model Free Approach to Kinetic Analysis of Real-Time Hyperpolarized  $^{13}\text{C}$  Magnetic Resonance Spectroscopy Data," *PLoS One*, vol. 8, no. 9, Sep. 2013.
- [11] C. J. Daniels, M. A. McLean, R. F. Schulte, F. J. Robb, A. B. Gill, N. McGlashan, M. J. Graves, M. Schwaiger, D. J. Lomas, K. M. Brindle, and F. A. Gallagher, "A comparison of quantitative methods for clinical imaging with hyperpolarized  $^{13}\text{C}$ -pyruvate," *NMR in Biomedicine*, vol. 29, no. 4, pp. 387–399, Apr. 2016.
- [12] U. Nestle, S. Kremp, A. Schaefer-Schuler, C. Sebastian-Welsch, D. Hellwig, C. Rbe, and C.-M. Kirsch, "Comparison of Different Methods for Delineation of  $^{18}\text{F}$ -FDG PET-Positive Tissue for Target Volume Definition in Radiotherapy of Patients with Non-Small Cell Lung Cancer," *Journal of Nuclear Medicine*, vol. 46, no. 8, pp. 1342–1348, Aug. 2005.
- [13] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, Jan. 2001.
- [14] R. S. Alomari, S. Kompalli, and V. Chaudhary, "Segmentation of the Liver from Abdominal CT Using Markov Random Field Model and GVF Snakes," in *International Conference on Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008*, Mar. 2008, pp. 293–298.
- [15] Y. Guo, Y. Feng, J. Sun, N. Zhang, W. Lin, Y. Sa, and P. Wang, "Automatic Lung Tumor Segmentation on PET/CT Images Using Fuzzy Markov Random Field Model," *Computational and Mathematical Methods in Medicine*, vol. 2014, p. e401201, May 2014.
- [16] F. A. Gallagher, M. I. Kettunen, and K. M. Brindle, "Imaging pH with hyperpolarized  $^{13}\text{C}$ ," *NMR in Biomedicine*, vol. 24, no. 8, pp. 1006–1015, Oct. 2011.
- [17] H. Zaidi, M. Diaz-Gomez, A. Boudraa, and D. O. Slosman, "Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging," *Physics in Medicine and Biology*, vol. 47, no. 7, p. 1143, 2002.
- [18] A.-E.-O. Boudraa, J. Champier, L. Cinotti, J.-C. Bordet, F. Lavenne, and J.-J. Mallet, "Delineation and quantitation of brain lesions by fuzzy clustering in Positron Emission Tomography," *Computerized Medical Imaging and Graphics*, vol. 20, no. 1, pp. 31–41, Jan. 1996.
- [19] S. Ruan, B. Moretti, J. Fadili, and D. Bloyet, "Fuzzy Markovian Segmentation in Application of Magnetic Resonance Images," *Computer Vision and Image Understanding*, vol. 85, no. 1, pp. 54–69, Jan. 2002.
- [20] S. P. Chatzis and T. A. Varvarigou, "A Fuzzy Clustering Approach Toward Hidden Markov Random Field Models for Enhanced Spatially Constrained Image Segmentation," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 5, pp. 1351–1361, Oct. 2008.
- [21] J. Zhang, "The mean field theory in EM procedures for Markov random fields," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2570–2583, Oct. 1992.
- [22] H. Ichihashi, K. Miyagishi, and K. Honda, "Fuzzy c-means clustering with regularization by K-L information," in *The 10th IEEE International Conference on Fuzzy Systems, 2001*, vol. 2, Dec. 2001, pp. 924–927 vol.3.
- [23] O. Khagai, R. F. Schulte, M. A. Janich, M. I. Menzel, E. Farrell, A. M. Otto, J. H. Ardenkjaer-Larsen, S. J. Glaser, A. Haase, M. Schwaiger, and F. Wiesinger, "Apparent rate constant mapping using hyperpolarized  $[1-^{13}\text{C}]$ pyruvate," *NMR in Biomedicine*, vol. 27, no. 10, pp. 1256–1265, Oct. 2014.
- [24] R. F. Schulte, J. I. Sperl, E. Weidl, M. I. Menzel, M. A. Janich, O. Khagai, M. Durst, J. H. Ardenkjaer-Larsen, S. J. Glaser, A. Haase, M. Schwaiger, and F. Wiesinger, "Saturation-recovery metabolic-exchange rate imaging with hyperpolarized  $[1-^{13}\text{C}]$  pyruvate using spectral-spatial excitation," *Magnetic Resonance in Medicine*, vol. 69, no. 5, pp. 1209–1216, May 2013.
- [25] F. Wiesinger, E. Weidl, M. I. Menzel, M. A. Janich, O. Khagai, S. J. Glaser, A. Haase, M. Schwaiger, and R. F. Schulte, "IDEAL spiral CSI for dynamic metabolic MR imaging of hyperpolarized  $[1-^{13}\text{C}]$ pyruvate," *Magnetic Resonance in Medicine*, vol. 68, no. 1, pp. 8–16, Jul. 2012.