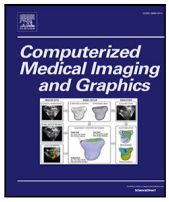
Computerized Medical Imaging and Graphics 113 (2024) 102333

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag



Automatic detection of myocardial ischaemia using generalisable spatio-temporal hierarchical Bayesian modelling of DCE-MRI

Yalei Yang [a,d], Dirk Husmeier [a], Hao Gao [a], Colin Berry [b], David Carrick [c], Aleksandra Radjenovic [b,*]

[a] School of Mathematics & Statistics, University of Glasgow, University Place, Glasgow, G12 8QQ, United Kingdom
[b] School of Cardiovascular & Metabolic Health, University of Glasgow, BHF Glasgow Cardiovascular Research Centre (GCRC), 126 University Place, Glasgow, G12 8TA, United Kingdom
[c] University Hospital Hairmyres, 218 Eaglesham Rd, East Kilbride, Glasgow G75 8RG, United Kingdom
[d] Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Key Laboratory for Research and Evaluation of Radiopharmaceuticals (National Medical Products Administration), Department of Nuclear Medicine, Peking University Cancer Hospital & Institute, Beijing, China

ARTICLE INFO

Keywords:
Hierarchical Bayesian models
Spatio-temporal Markov random fields
First pass myocardial perfusion DCE-MRI
Myocardial hypoperfusion
Myocardial ischaemia
Gibbs sampling
Model selection
Watanabe Akaike information criterion

ABSTRACT

Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) can be used as a non-invasive method for the assessment of myocardial perfusion. The acquired images can be utilised to analyse the spatial extent and severity of myocardial ischaemia (regions with impaired microvascular blood flow). In the present paper, we propose a novel generalisable spatio-temporal hierarchical Bayesian model (GST-HBM) to automate the detection of ischaemic lesions and improve the in silico prediction accuracy by systematically integrating spatio-temporal context information. We present a computational inference procedure with an adequate trade-off between accuracy and computational efficiency, whereby model parameters are sampled from the posterior distribution with Gibbs sampling, while lower-level hyperparameters are selected using model selection strategies based on the Watanabe Akaike information criterion (WAIC). We have assessed our method on both synthetic (in silico) data with known gold-standard and 12 sets of clinical first-pass myocardial perfusion DCE-MRI datasets. We have also carried out a comparative performance evaluation with four established alternative methods: Gaussian mixture model (GMM), opening and closing operations based on Gaussian mixture model ($GMM^{max}_{C\&O}$), Markov random field constrained Gaussian mixture model (GMM-MRF) and model-based hierarchical Bayesian model (M-HBM). Our results show that the proposed GST-HBM method achieves much higher in silico prediction accuracy than the established alternative methods. Furthermore, this method appears to provide a more robust delineation of ischaemic lesions in datasets affected by spatially variant noise.

1. Introduction

Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) is increasingly being used to assess myocardial perfusion (Khalifa et al., 2014). This is based on the seminal article (Jerosch-Herold, 2010), where model-based and model-independent methods were introduced to quantify myocardial perfusion. Most of these methods focus on quantifying physiological parameters, especially the myocardial blood flow (MBF) (Jerosch-Herold, 2010; Jerosch-Herold et al., 2002; Larsson et al., 1996; Jerosch-Herold et al., 1998). MBF is commonly reported using a standardised reporting grid (17-segment AHA model (Cerqueira et al., 2002)). The advantage of this approach lies in its standardisation and computational simplicity (Fenchel et al., 2005; Tamarappoo et al., 2010). However, the definition of discrete myocardial segments is subject to operator bias and there is an inevitable loss of granularity caused by signal averaging. Pixel-based methods (Kellman et al., 2017) avoid these drawbacks, and preserve source DCE-MRI spatial resolution, although at a substantially increased computational cost. Recently, Bayesian spatial modelling methods have been applied to improve the accuracy of MBF estimates (Yang et al., 2022a; Scannell et al., 2020; Schmid and Yang, 2009; Lehnert et al., 2019) by introducing pixel (voxel) neighbourhood information using Markov random field priors (Bishop, 2006). The Bayesian spatial regression model (Genkin et al., 2007; Metzner et al., 2019) has been applied to classify myocardial pixels into ischaemic and healthy groups using the estimated MBF values in Lehnert et al. (2019). However, the main goal of the work in Lehnert et al. (2019) is to quantify MBF values, and a separate

* Corresponding author.
E-mail addresses: dirk.husmeier@glasgow.ac.uk (D. Husmeier), aleksandra.radjenovic@glasgow.ac.uk (A. Radjenovic).

https://doi.org/10.1016/j.compmedimag.2024.102333
Received 18 August 2023; Received in revised form 15 December 2023; Accepted 26 December 2023
Available online 11 January 2024

classification step is subsequently performed based on the estimated MBF assays. Specifically, all pixels with MBF values lower than 60% of the median MBF values are assigned to the ischaemic group and other pixels are assigned to the healthy group. The classification methods for the myocardial pixels using myocardial perfusion DCE-MRI have been proposed by the authors' group in recent publications (Yang et al., 2019, 2022a,b). In Yang et al. (2019), a spatially variant finite mixture model fitted with an expectation–maximisation (EM) algorithm has been introduced to improve the classification accuracy of the standard Gaussian mixture model (GMM). Specifically, this method has made use of spatial information to reduce the number of singular and spurious small clusters. In Yang et al. (2022b), a Markov random field constrained Gaussian mixture model (GMM-MRF) has been developed to classify myocardial pixels. This method aims to use an iterated conditional modes (ICM) algorithm for a maximum a posteriori probability (MAP) estimation of the classification map. A model-based hierarchical Bayesian model (HBM) classification approach has been introduced in Yang et al. (2022a). This method has combined a parametric model of the MBF, the Fermi model, with spatial prior information using an HBM to obtain classification and estimation maps of the MBF that have higher accuracy than the method proposed in Scannell et al. (2020). However, the model-based HBM proposed in Yang et al. (2022a) has turned out to be unstable when exposed to spatially variant noise. This is because the parametric model – the Fermi model – tends to misestimate the MBF when the early phase of the DCE-MRI time series (pre-contrast, or baseline phase) has spatially variant noise. Errors in derived MBF estimates caused by the spatially variant baseline noise have a negative impact on the accuracy of classification.

Classification methods for myocardial pixels have also been widely developed based on late gadolinium enhancement (LGE) MRI data. A thorough comparison among different methods has been reported in Karim et al. (2016). The authors have compared support vector machines and level sets (Sethian, 1999), region growing and morphology (Hojjatoleslami and Kittler, 1998), conditional random fields (Lafferty et al., 2001), watershed transformations (Hennemuth et al., 2008) and graph cuts with EM-algorithm (Dai et al., 2015). Except for the rule-based methods, learning-based methods, e.g. Chen et al. (2020), de la Rosa et al. (2021) and Ukwatta et al. (2014), have also been developed for the classification of LGE data. Moreover, classification methods have also been developed for cardiac cine MRI data (Xu et al., 2020; Zhang et al., 2019).

The work proposed in the present paper aims to delineate areas of hypoperfusion (myocardial ischaemia) from dynamic contrast-enhanced MRI datasets (DCE-MRI). It differs from algorithms that focus on infarct delineation based on LGE or cine MRI both in terms of the classification target (ischaemic versus scar tissue), and in terms of the nature of the source data. Whilst cine MRI is dynamic, it is not enhanced by exogenous contrast. LGE is contrast enhanced, but it is not dynamic. Due to the nature of these differences between data types and classification targets, a direct comparison with cine and LGE methods is not possible.

In the present paper, we apply and evaluate a novel generalisable spatio-temporal hierarchical Bayesian model (GST-HBM) using Markov random field priors to the DCE-MRI myocardial perfusion data. We aim to classify the myocardial tissues into two categories, healthy tissue and lesion based on the signal intensity and the spatio-temporal information of the myocardial pixels. Specifically, given the original images, the proposed method first generates an initial classification map. Then, based on the generated classification map, it generates denoised images. The proposed method repeats this procedure until it converges. Spatio-temporal Markov random field priors are used to introduce the spatio-temporal information for each pixel. We have derived a posterior inference scheme for the parameters in the GST-HBM model, using a Markov chain Monte Carlo (MCMC) variant (Gibbs sampling) to approximately draw samples of the parameters from their posterior distributions. For an adequate trade-off between accuracy and computational efficiency, the hyperparameters are selected using model selection techniques based on the Watanabe Akaike information criterion (WAIC) (Watanabe, 2013). The proposed GST-HBM method is generalisable and therefore applicable to a wide range of dynamic (time series) contrast enhanced imaging data. We demonstrated its application to one specific example of such data (DCE-MRI of myocardial perfusion). Moreover, GST-HBM is applicable to lesion detection and characterisation using DCE-MRI in other diseases, most notably in oncology (e.g. brain, breast, prostate), where a range of tracer kinetic models can be deployed (Ingrisch and Sourbron, 2013). In addition to DCE-MRI, the proposed GST-HBM method is readily applicable to other imaging modalities, such as dynamic computed tomography (CT) and positron emission tomography (PET).

Starting with a review of the data and a brief description of the four established methods that we included in our comparative evaluation study in Section 2, we propose our new method in Section 3. The comparative evaluation, with an assessment of our method and a comparison with the established alternative methods, is presented in Section 4. Section 5 contains a discussion and an outlook on future work. Our paper concludes in Section 6.

2. Material and methods

2.1. Myocardial perfusion DCE-MRI

Myocardial perfusion DCE-MRI is a method to noninvasively obtain the images of the tissue of interest (e.g. myocardium) before, during and after intravenous administration of exogenous MRI contrast agent (see Fig. 1). The clinical data used in this work are two-dimensional mid-ventricular short axis DCE-MRI slices (Carrick et al., 2015).

The contrast agents shorten local T1 (spin-lattice or longitudinal) relaxation times in proportion to their local tissue concentration. As a result, in T1 weighted MRI (the data we used in this work), the signal intensity (SI) for normally perfused regions is higher than for the hypoperfused (ischaemic) regions during the first-pass of the contrast agents. The rate of SI increase for relatively hypoperfused regions is also much slower than the normally perfused regions. The signal curves in Fig. 2 illustrate this phenomenon. To detect and delineate myocardial ischaemic lesions, we exploited the differences in patterns of SI enhancement between normally perfused and hypoperfused tissue regions.

2.2. Clinical data

The clinical data have been derived from a BHF MR-MI clinical imaging cohort study (ClinicalTrials.gov, NCT02072850). The patients from this study were diagnosed with ST-segment elevation myocardial infarction (STEMI) (Carrick et al., 2015). The cardiac MRI protocol included the assessment of resting perfusion using a T1-weighted, fast gradient echo DCE-MRI sequence. The dynamic sequences of the MR images recorded the signal changes within the heart before, during and immediately following intravenous administration of contrast agent.

The myocardial perfusion DCE-MRI data (mid-ventricular short-axis slice) from 12 datasets (data 1–12) were analysed in this work. For one dataset (data 1), we show all results (model selection, MCMC convergence test, classification maps) in Section 4. For the other datasets, we present classification maps in the supplementary materials to illustrate the robustness of the proposed method. The cardiac MRI exams were performed with a Siemens MAGNETOM Avanto (Erlangen, Germany) 1.5-Tesla scanner with a 12-element phased array cardiac surface coil. The assessment of resting myocardial perfusion was performed during intravenous administration of 0.075 mmol/kg of contrast agent (gadoterate meglumine, Dotarem, Guebert S.A.). The myocardial tissue (contouring of the epicardial and endocardial boundary, 500 - 800 pixels in total) was manually delineated by an experienced clinician using the cardiac image analysis software QMASS 8.1 (Leiden, The Netherlands) (Carrick et al., 2015).

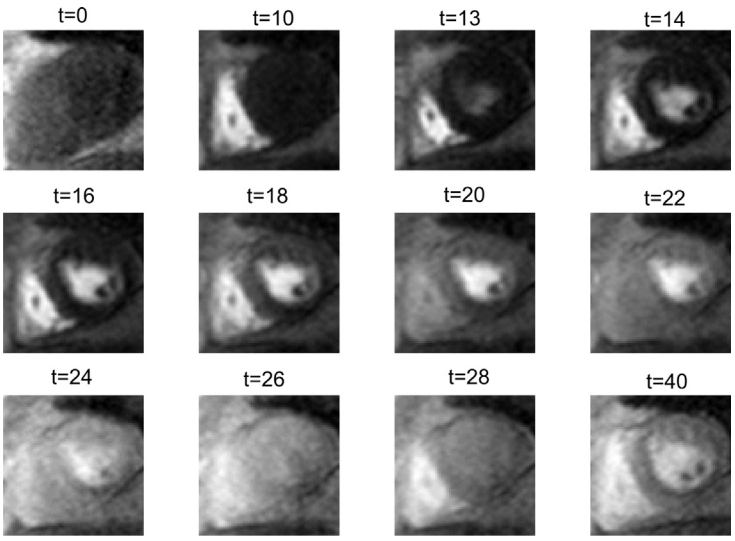


Fig. 1. The original images of the myocardial perfusion given different enhancement time points. The temporal resolution of the DCE-MRI datasets is equal to the duration of a heartbeat (typically around 1 s).

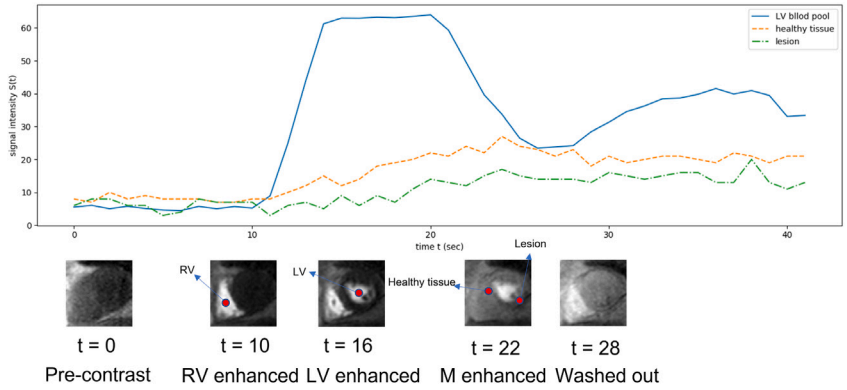


Fig. 2. The line chart shows DCE-MRI signal intensity S(t) before, during and after administration of MRI contrast agent. The three curves represent signal intensity changes of randomly selected pixels in three regions of interest (ROIs): left ventricle (LV) blood pool, lesion and contralateral "healthy" tissue. The 5 images below the line chart show 5 steps of myocardial perfusion with time, i.e. pre-contrast, the signal intensity enhancement inside the right ventricle (RV), followed by LV blood pool and myocardium (M) followed by LV blood pool and the wash out of contrast from the blood pool (t = 28).

2.3. Synthetic data

The synthetic data was designed to mimic clinically observed lesion. A double exponential curve was used to model the signal intensity with time (see details in Radjenovic (2003), Chapter 2):

$$s(t) = \frac{p_2 p_3}{(p_2 - p_1)} \times (e^{-p_1 t} - e^{-p_2 t}). \quad (1)$$

Rician noise is usually assumed for MR images according to the work in Gudbjartsson and Patz (1995). In this way, the synthetic data was generated based on Eq. (1) by adding Rician noise with different scales. Specifically, we used different variances of Rayleigh distribution (a special case of Rician distribution) to design the degrees of noise. The standard deviation values were 1.5 (low noise), 2 (moderate noise), 2.5 (high noise) and 3 (very high noise). Fig. 3 shows the MBF estimates obtained by the Fermi method (Jerosch-Herold et al., 1998) with different noise variances and noise-free ground truth.

In the proposed work, we have designed synthetic images with three types of lesions.

- The first type is a non-transmural lesion with a standard annulus shape (see Fig. 4(a)). For this case, the region close to the endocardial boundary was assumed to be a lesion and we also assumed that there were 50% of lesion pixels and 50% of healthy pixels. In other words, 180° of the myocardium was assumed to be a lesion.
- The second type is a fully transmural lesion with a standard annulus shape (see Fig. 4(b)). Similarly, 180° of the myocardium was assumed to be a lesion.
- The third type is a fully transmural lesion with a realistic ventricular wall obtained from one clinical dataset (see Fig. 4(c,d,e)). For this case, there were three sizes of lesions in the myocardium. A small lesion spanning 60°, a medium-size lesion spanning 120°, and a large lesion spanning 180° circumferentially. The main reason for designing different angular spans is that the ratio of healthy pixels and lesion pixels will also affect the classification results. Let r denote the ratio of healthy pixels and lesion pixels. This implies: $r = 1$ when the angle is 180°, $r = 2$ when the angle is 120° and $r = 5$ when the angle is 60°.

There were a total of 20 combinations based on different lesion types and values of noise variance. For each combination, 10 separate sets of synthetic images were simulated to improve the robustness of our evaluation and reduce the effect of potential outliers.

The parameters in Eq. (1) for healthy tissues and lesions are given based on empirical experiences. To be specific, we use Eq. (1) to fit a model to the clinical DCE-MRI data (serial 1) by least-squares estimation to obtain the values of the parameters. In detail, the values of parameters in Eq. (1) are set to $p_1 = 0.01$, $p_2 = 0.4$, $p_3 = 25$ for the healthy tissues. For the lesions, the parameters are set to $p_1 = 0.02$, $p_2 = 0.3$, $p_3 = 20$.

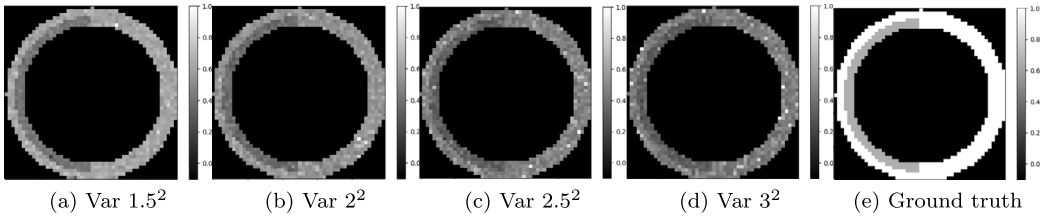


Fig. 3. Panels (a) - (d) show the MBF estimates obtained by Fermi method with Rician noise variance 1.5^2, 2^2, 2.5^2 and 3^2 respectively. The MBF estimates have been standardised within the range $[0, 1]$. Panel (e) shows the ground truth of the MBF estimates without any noise.

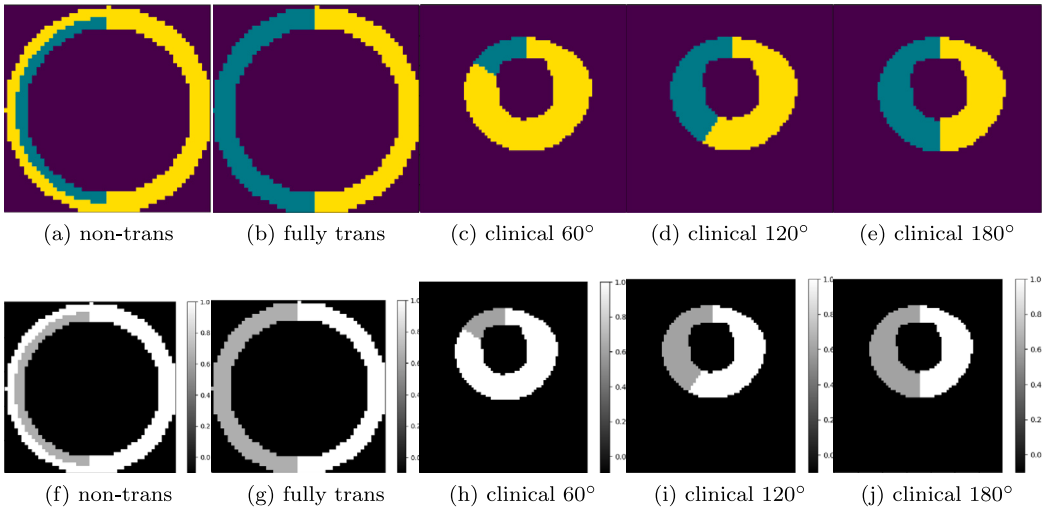


Fig. 4. Panels (a) - (e) show the design of non-transmural, fully transmural, 60°, 120° and 180° lesions respectively. The dark blue region denotes the lesion, and the yellow region denotes the healthy tissues. Panels (f) and (j) show the MBF estimates without any noise for the five different types of lesions respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Benchmark methods

We used four alternative methods as the benchmark methods in this paper. The first benchmark method was the Gaussian mixture model classification method (see chapter 9, Bishop (2006)). This method can separate all myocardial tissues (pixels) into several (two in our work) different Gaussian distributions based on the SI of the pixels and return the labels of the respective cluster. The maximum likelihood parameter estimates of the Gaussian mixture model can be obtained with the EM algorithm (Bishop, 2006; Moon, 1996). As mentioned in Section 3.1, this classification method is susceptible to spurious singleton clusters. This method is named as GMM (see Table 1).

The second benchmark method was an image processing approach called the "opening and closing operations" (Chen and Haralick, 1995), referred to as "$\text{GMM}^{\max}_{\text{C\&O}}$" (see Table 1). The opening and closing operations can potentially improve the GMM classification results by reducing spurious singleton clusters. This approach applies mathematical morphology (erosion and dilation) to reduce the occurrence of spurious clusters. Both closing and opening operations have been derived from two fundamental operations called "erosion" and "dilation". In general, the opening operation can remove small lesion clusters inside healthy tissues and the closing operation can remove small healthy clusters inside lesions. However, this method requires a user to specify the kernel size for the opening and closing operators. Since there is no objective way to set this parameter, we tried different kernel sizes and chose the one that achieved the closest agreement with the ground truth (which happened to be a 2×2 kernel). We also discarded all but the largest lesion, in line with the ground truth. Note that this procedure gives the $\text{GMM}^{\max}_{\text{C\&O}}$ method an unfair competitive advantage over the GST-HBM method that we have proposed, as it uses ground truth information that would not be available in real applications.

Table 1
Descriptions of all methods compared within this paper.

Abbreviation of methods	Description
GMM	Gaussian mixture model classification
$\text{GMM}^{\max}_{\text{C\&O}}$	GMM based opening and closing operations with the largest lesion kept
GMM-MRF	The method proposed in Yang et al. (2022b)
M-HBM	The method proposed in Yang et al. (2022a)
GST-HBM	The method proposed in this paper

The third benchmark method was the GMM-MRF model (see Table 1) proposed in Yang et al. (2022b). This method is a modification of the standard GMM which aims to improve the classification accuracy by suppressing spurious singleton clusters through the use of spatial context information. The drawback of this method is that it was fitted by a greedy optimisation method, the ICM algorithm, which is susceptible to entrapment in local minima. As a consequence, it failed to eliminate some misclassified pixels that were generated by spatially variant noise (Yang et al., 2022b).

The fourth benchmark method was the model-based HBM classification method proposed in Yang et al. (2022a). This method combines physiological and spatial information to accurately estimate and classify the MBF using myocardial perfusion DCE-MRI. This method is referred to as M-HBM (see Table 1).

We refer to the method proposed in the paper as GST-HBM (see Table 1).

3. Theory

3.1. Spatio-temporal information

Spatial methods have been widely applied for quantitative assessment of myocardial perfusion DCE-MRI (Scannell et al., 2020; Schmid and Yang, 2009; Lehnert et al., 2019; Cordero-Grande et al., 2011). According to the idea that adjacent tissues show similar properties, we incorporate spatial information. We also introduce temporal information because the myocardial perfusion DCE-MRI data are time-series images. The signal intensities for a pixel at different times are related, which is illustrated in Fig. 2. Specifically, there are three stages for the change of signals of myocardial pixels. In the beginning, the signal intensity stays relatively low because the contrast agent has not yet reached the myocardium. Then, it increases steadily because the contrast agent is flowing into the myocardium. Finally, it begins to decrease steadily because the contrast agent is washed out from the myocardium. Therefore, we introduce spatio-temporal information to the DCE-MR SI images using Markov random field priors, which are explicitly described in Section 3.3.2.

For some naive applications of standard clustering methods, like GMM classification, only the signal information can be utilised. Specifically, GMM segmentation assigns all pixels into several groups (two groups in this case) based on the values of these signals. In this way, GMM segmentation is susceptible to spurious singleton clusters in some low signal-to-noise ratio (SNR) cases. This phenomenon is physiologically unrealistic because the blood supply to myocardial tissues is from three main coronary arteries. This means that the lesions are highly likely to be located in continuous areas. The spurious singleton clusters can be reduced or even eliminated by introducing neighbourhood information (Yang et al., 2019). We therefore applied Markov random field priors to the label of each pixel. The details of this application can be found in Section 3.3.1.

3.2. Hierarchical Bayesian model

Let $y^i(t)$ denote the logarithm of the signal intensity of an MR image pixel ($i = 1, 2, \dots, N$) at time point t ($t = 1, 2, \dots, M$), where N is the number of pixels and M is the number of time points. $k^i \in \{0, 1\}$ is defined as the state (healthy and lesion) for the ith pixel.

In this work, a GST-HBM was used to classify the tissues based on myocardial perfusion DCE-MRI. This model can be found in Fig. 5. Table 2 shows the description of the parameters in Fig. 5. In Fig. 5, if there is a link from node A to node B, then node A is the parent of node B, or in other words, node B is the child of node A. The child node is conditional on the parent node. The nodes and links indicate conditional dependence relations between different parameters. Specifically, the observation $y^i(t)$ is conditionally dependent on the true signal intensity $\psi^i(t)$ and variance σ^2. The variable $\psi^i(t)$ is assumed to be the denoised signal, and σ^2 is the variance of the white noise. Furthermore, $\psi^i(t)$ follows a conditional Gaussian distribution. To be specific, the true signal $\psi^i(t)$ given different labels k^i is Gaussian distributed with mean $\mu_{k^i}(t)$ and variance $\sigma^2_{k^i}(t)$. Moreover, $\psi^i(t)$ is constrained by its spatial neighbour $\psi^{-i}(t)$ and temporal neighbour $\psi^i(t+j)$, $a_{k^i}(t+j)$ and $\mathbf{b}_{k^i}(t+j)$ where $a_{k^i}(t+j)$ and $\mathbf{b}_{k^i}(t+j)$ are autoregressive parameters (coefficients) between $\psi^i(t)$and $\psi^i(t+j)$.

In this way, Fig. 5 explicitly illustrates the relationships between the parameters in this GST-HBM model. According to the factorisation rule, which is specified in Bishop (2006) chapter 8, the joint distribution for this hierarchical Bayesian model is the product of conditional distributions determined by their respective parent nodes. In this way, the joint distribution is:

$$
\begin{aligned}
&P\left(y^i(t), \psi^i(t), \sigma^2, \psi^i(t+j), \psi^{-i}(t), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j), \mu_{k^i}(t), \sigma^2_{k^i}(t), \boldsymbol{\Gamma}, k^i, k^{-i}\right) \\
=&P\left(y^i(t)|\psi^i(t), \sigma^2\right) P\left(\psi^i(t)|\psi^i(t+j), \psi^{-i}(t), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j), \mu_{k^i}(t), \sigma^2_{k^i}(t), k^i, k^{-i}\right) \\
&P\left(\sigma^2|\alpha_{**}, \beta_{**}\right) P\left(\psi^i(t+j)\right) P\left(a_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}\right) P\left(\mathbf{b}_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}\right)
\end{aligned}
$$

Table 2
Descriptions of the parameters for the GST-HBM model.

Parameter	Description
$y^i(t)$	The logarithm (log) of the observed signal for pixel i at time t
$\psi^i(t)$	The log of the true signal for pixel i at time t
σ^2	The variance of the IID Gaussian noise
k^i	The label for pixel i
k^{-i}	The labels for the neighbours of pixel i
$\psi^{-i}(t)$	The log of the true signals for the neighbours of pixel i at time t
$\psi^i(t+j)$	The log of the true signals for pixel i at time $t+j$
$a_{k^i}(t+j)$	Autoregressive parameter (constant) between $\psi^i(t)$and $\psi^i(t+j)$
$\mathbf{b}_{k^i}(t+j)$	Autoregressive parameter (coefficient) between $\psi^i(t)$and $\psi^i(t+j)$
$\mu_{k^i}(t)$	Mean for the true signals conditional on the label at time t
$\sigma^2_{k^i}(t)$	Variance for the true signals conditional on the label at time t
$*$ and $**$	The symbols with $*$ and $**$ are hyperparameters

$$
P\left(\mu_{k^i}(t)|\mu_*, \sigma^2_*\right) P\left(\sigma^2_{k^i}(t)|\alpha_*, \beta_*\right) P\left(\psi^{-i}(t)\right) P\left(k^i|k^{-i}\right) P\left(k^{-i}\right) \tag{2}
$$

where $\boldsymbol{\Gamma} = \{\mu_*, \sigma^2_*, \alpha_*, \beta_*, \mu_{**}, \sigma^2_{**}, \alpha_{**}, \beta_{**}\}$ is a vector containing all hyperparameters. This equation shows how the joint probability of all variables of the proposed GST-HBM model can be expanded into a product of less complex conditional probabilities. The analysis of the full probabilistic model can thus be simplified, in terms of less-complex conditional distributions of specific parameters. The statistical inference of the conditional distributions of all parameters in Fig. 5 is presented in Section 3.4. The distributions of the parameters in Table 2 and their corresponding prior distributions are introduced in the remaining parts of Sections 3.2 and 3.3.

In this GST-HBM, the logarithm of signal intensity $y^i(t)$ is assumed to be Gaussian distributed with mean $\psi^i(t)$ and variance σ^2, which can be written as:

$$
P\left(y^i(t)|\psi^i(t), \sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\left(y^i(t) - \psi^i(t)\right)^2}{2\sigma^2}\right) \tag{3}
$$

where $\psi^i(t)$ indicates the true signal intensity of pixel i at time t. σ^2 is assumed to be the variance of the IID Gaussian noise. $\alpha_{**} = 0.1$ and $\beta_{**} = 0.1$ are fairly uninformative hyperparameters for variance σ^2. In this way, the prior distribution for σ^2, which is conjugate, is

$$
P(\sigma^2|\alpha_{**}, \beta_{**}) = \frac{\beta^{\alpha_{**}}_{**}}{\Gamma(\alpha_{**})} (\sigma^2)^{-\alpha_{**}-1} \exp\left(-\frac{\beta_{**}}{\sigma^2}\right). \tag{4}
$$

The myocardial blood flows for healthy tissues and lesions are different. Moreover, the myocardial blood flows are positively correlated with the signal intensities for myocardial tissues. In this way, $\psi^i(t)$ is conditionally dependent on the label k^i. Specifically, the probability density of $\psi^i(t)$ conditional on the label k^i is

$$
P\left(\psi^i(t)|\mu_\phi(t), \sigma^2_\phi(t)\right) = \frac{1}{\sqrt{2\pi}\sigma_\phi(t)} \exp\left(\frac{-\left(\psi^i(t) - \mu_\phi(t)\right)^2}{2\sigma^2_\phi(t)}\right) \tag{5}
$$

with $\phi \in \{0, 1\}$. $\mu_\phi(t)$ and $\sigma^2_\phi(t)$ are the simple forms of $\mu_{k^i=\phi}(t)$ and $\sigma^2_{k^i=\phi}(t)$. Let $\boldsymbol{\psi}_\phi(t) = \{\psi^i(t)\}_{i|k^i=\phi}$, we have:

$$
P\left(\boldsymbol{\psi}_\phi(t)|\mu_\phi(t), \sigma^2_\phi(t)\right) = \left(2\pi\sigma^2_\phi(t)\right)^{-\frac{N_\phi}{2}} \exp\left(\frac{-\sum_{i|k^i=\phi}\left(\psi^i(t) - \mu_\phi(t)\right)^2}{2\sigma^2_\phi(t)}\right) \tag{6}
$$

with $\phi \in \{0, 1\}$. N_ϕ is the number of pixels which satisfy label $k^i = \phi$. In a nutshell, all pixels in the myocardium are separated in two categories, healthy group and lesion group. The pixels in different groups have different means and standard deviations. $\mu_* = 0$, $\sigma^2_* = 10$, $\alpha_* = 0.1$ and $\beta_* = 0.1$ are fairly uninformative hyperparameters for $\mu_{k^i}(t)$ and

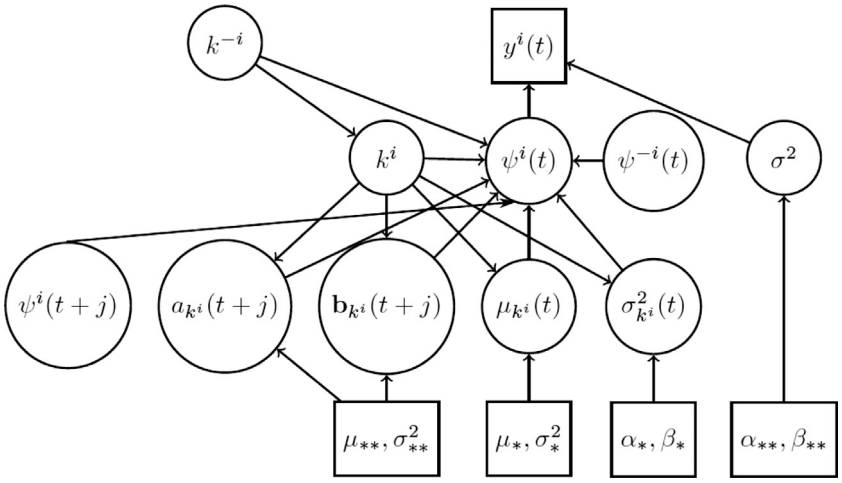


Fig. 5. This figure shows the structure of the hierarchical Bayesian model proposed in this work. The circle nodes denote variables and the rectangle nodes denote fixed values.

$\sigma^2_{k^i}(t)$. In this way, the prior distributions for $\mu_{k^i}(t)$ and $\sigma^2_{k^i}(t)$, which are conjugate, are

$$P\left(\mu_{k^i}(t)|\mu_*, \sigma^2_*\right) = \frac{1}{\sqrt{2\pi}\sigma_*} \exp\left(\frac{-\left(\mu_{k^i}(t) - \mu_*\right)^2}{2\sigma^2_*}\right) \tag{7}$$

and

$$P\left(\sigma^2_{k^i}(t)|\alpha_*, \beta_*\right) = \frac{\beta^{\alpha_*}_*}{\Gamma(\alpha_*)} \left(\sigma^2_{k^i}(t)\right)^{-\alpha_*-1} \exp\left(-\frac{\beta_*}{\sigma^2_{k^i}(t)}\right). \tag{8}$$

The true signal $\psi^i(t)$ and label k^i are dependent on their spatial neighbours $\psi^{-i}(t)$ and k^{-i} using Markov random fields. We further applied $\psi^i(t+j)$ as the temporal neighbours of $\psi^i(t)$. The definitions of $\psi^{-i}(t)$, $\psi^i(t+j)$ and k^{-i} are explicitly illustrated in Section 3.3.

3.3. Markov random fields

3.3.1. Markov random fields for labels

A Markov random field method was used to introduce the spatial neighbourhood information in this paper. According to the Hammersley–Clifford theorem, a Markov random field can be written as a Gibbs distribution (Li, 2009),

$$P(k^i|k^{-i}) = \frac{1}{\mathrm{Q}} \exp\left(-U(k^i|k^{-i})\right) \tag{9}$$

where Q is a normalisation constant to ensure the integral of $P(k^i|k^{-i})$ is 1. k^{-i} is the set of all other labels excluding k^i, $k^{-i} = \{k^m\}_{m\neq i}$. We defined $U(k^i|k^{-i})$ to be:

$$U(k^i|k^{-i}) = \frac{1}{\mathrm{T}} \sum_{i\sim m} u(k^i|k^m), \tag{10}$$

where T is a weight parameter. $i \sim m$ denotes the spatial neighbours of pixel i. The definition of $u(k^i|k^m)$ is

$$u(k^i|k^m) = \begin{cases} -\left(\frac{1}{2}\right)^{o-1} & k^i = k^m, \\ \left(\frac{1}{2}\right)^{o-1} & k^i \neq k^m \end{cases} \tag{11}$$

where o indicates the degree of neighbouring. Fig. 6 shows the definitions of spatial and temporal neighbours in this work. The definition for temporal neighbours will be used in Section 3.3.2. We chose $o = 1$ in this work because it is computationally cheaper than higher degrees, e.g. $o = 2$. However, our method provides a straightforward mechanism for the extension to higher degrees.

3.3.2. Markov random fields for true signals

As mentioned in Section 3.1, we introduced spatial Markov random fields to the true signal $\psi^i(t)$. A Markov random field can be written as a Gibbs distribution, we therefore have:

$$P\left(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i}\right) = \frac{1}{Q_{\mathrm{ps}}} \exp\left(-U(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i})\right) \tag{12}$$

where $\mathrm{Q_{ps}}$ is a normalisation constant. $\psi^{-i}(t)$ is the set of all other true signals excluding $\psi^i(t)$, $\psi^{-i}(t) = \{\psi^m(t)\}_{m\neq i}$. We defined $U\left(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i}\right)$ to be:

$$U\left(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i}\right) = \frac{1}{T_{\mathrm{ps}}} \sum_{i\sim m} v_{i,m} \left(\psi^i(t) - \psi^m(t)\right)^2 \tag{13}$$

where T_{ps} is a weight parameter and $v_{i,m}$ is the edge-preservation parameter defined by

$$v_{i,m} = \begin{cases} 0 & k^i \neq k^m(t) \\ 1 & k^i = k^m(t). \end{cases} \tag{14}$$

The edge-preservation parameter $v_{i,m}$ is used to prevent the smoothing effect of the neighbourhood information extend beyond the boundaries of different tissues, i.e. healthy tissues and lesions (Scannell et al., 2020; Bardsley, 2012).

In the proposed method, we used an auto-regressive model to describe the relationship for the true signals between different times. j is a non-zero integer to indicate the neighbouring time points. To be specific, the Markov random fields prior given temporal neighbour $\psi^i(t+j)$ is

$$P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right) = \frac{1}{Q_{\mathrm{pt}}} \exp\left(-U(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j))\right) \tag{15}$$

where

$$U\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right) = \frac{1}{T_{\mathrm{pt}}} \sum_{t\sim j} \left(\psi^i(t) - \sum_{p=0}^{q} b^p_{k^i}(t+j)\psi^i(t+j) - a_{k^i}(t+j)\right)^2 \tag{16}$$

where $\mathrm{Q_{pt}}$ is a normalisation constant. $\mathrm{T_{pt}}$ is a weight parameter and $\mathbf{b}_{k^i}(t+j) = \{b^0_{k^i}(t+j), \ldots, b^q_{k^i}(t+j)\}$. In our work, we assumed that the auto-regressive parameters, $\mathbf{b}_{k^i}(t+j)$ and $a_{k^i}(t+j)$ are dependent on the label k^i. This is physiologically realistic. For myocardial perfusion MRI data, the lesion tissues are hypoperfused and the healthy tissues are normally perfused. This means that not only are the signal intensities for the lesion pixels lower than for the healthy tissues but also the growth rate of the signal intensities for the lesion pixels is lower than for the healthy tissues. In our work, we chose $j = -1$ and $q = 0$. Given $q = 0$ and $j = -1$, Gaussian prior distributions are assumed for both $b^0_{k^i}(t-1)$ and $a_{k^i}(t-1)$ with fairly uninformative hyperparameters

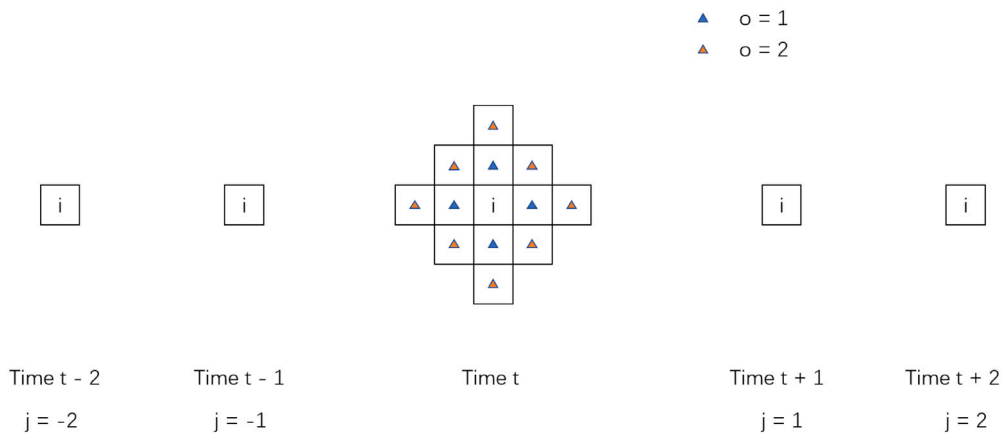


Fig. 6. This figure shows the definitions of spatial and temporal neighbouring pixels in this paper. The blue triangles show the first order of spatial neighbours and the red triangles show the second order of spatial neighbours. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$\mu_{**} = 0$ and $\sigma^2_{**} = 10$. Specifically, the prior distributions for $b^0_{k^i}(t-1)$ and $a_{k^i}(t-1)$, which are conjugate, are

$$P\left(b^0_{k^i}(t-1)|\mu_{**},\sigma^2_{**}\right) = \frac{1}{\sqrt{2\pi}\sigma_{**}}\exp\left(\frac{-(b^0_{k^i}(t-1)-\mu_{**})^2}{2\sigma^2_{**}}\right) \tag{17}$$

and

$$P\left(a_{k^i}(t-1)|\mu_{**},\sigma^2_{**}\right) = \frac{1}{\sqrt{2\pi}\sigma_{**}}\exp\left(\frac{-(a_{k^i}(t-1)-\mu_{**})^2}{2\sigma^2_{**}}\right). \tag{18}$$

3.4. Posterior inference

The hierarchical Bayesian model in Fig. 5 is a directed acyclic graph. In a directed acyclic graph, the probability of a selected parameter conditional on all other parameters is given by the probability of this parameter conditional on its Markov blanket (its parents, children and co-parents nodes) (see details in Bishop (2006), Chapter 8). For example, for parameter σ^2, its parent node is (α_{**},β_{**}). Its child node is $y^i(t)$. Its co-parent node is $\psi^i(t)$. In this section, the conditional distributions for $\psi^i(t)$, $\mu_{k^i}(t)$, $\sigma^2_{k^i}(t)$, σ^2, $a_{k^i}(t+j)$, $\mathbf{b}_{k^i}(t+j)$ and k^i are derived using their corresponding Markov blankets. Furthermore, given the derived conditional distributions, these parameters can be sampled using a specific MCMC method, i.e. Gibbs sampling.

The Markov blanket for $\psi^i(t)$ is $\{y^i(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j), k^i, \psi^{-i}(t), k^{-i}, \mu_{k^i}(t), \sigma^2_{k^i}(t), \sigma^2\}$. In this way, its conditional distribution is

$$\begin{aligned} &P\left(\psi^i(t)|y^i(t),\psi^i(t+j),a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j),k^i,\psi^{-i}(t),k^{-i},\mu_{k^i}(t),\sigma^2_{k^i}(t),\sigma^2\right)\\ \propto &P\left(y^i(t)|\psi^i(t),\sigma^2\right)P\left(\psi^i(t)|\mu_{k^i}(t),\sigma^2_{k^i}(t)\right)P\left(\psi^i(t)|\psi^{-i}(t),k^i,k^{-i}\right)\\ &P\left(\psi^i(t)|\psi^i(t+j),a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j)\right). \end{aligned} \tag{19}$$

$P(y^i(t)|\psi^i(t),\sigma^2)$ can be found in Eq. (3). Given $k^i = \phi$ with $\phi \in \{0,1\}$, $P(\psi^i(t)|\mu_{k^i=\phi}(t),\sigma^2_{k^i=\phi}(t))$ can be found in Eq. (5). $P(\psi^i(t)|\psi^{-i}(t),k^i,k^{-i})$ can be found in Eqs. (12), (13), (14) and $P(\psi^i(t)|\psi^{-i}(t),a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j))$ can be found in Eqs. (15), (16). By substituting these equations into Eq. (19), we have:

$$\begin{aligned} &P\left(\psi^i(t)|y^i(t),\psi^i(t+j),a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j),k^i,\psi^{-i}(t),k^{-i},\mu_{k^i}(t),\sigma^2_{k^i}(t),\sigma^2\right)\\ &\propto \exp\left(\frac{-A\times(\psi^i(t))^2+2B\times\psi^i(t)}{2C}\right) \end{aligned} \tag{20}$$

where

$$A = \sigma^2_\phi(t)T_{\mathrm{ps}}T_{\mathrm{pt}} + \sigma^2 T_{\mathrm{ps}}T_{\mathrm{pt}} + 2\sigma^2\sigma^2_\phi(t)T_{\mathrm{pt}}\sum_{i\sim m}\nu_{i,m} + 2\sigma^2\sigma^2_\phi(t)T_{\mathrm{ps}}, \tag{21}$$

$$\begin{aligned} B = {}&\sigma^2_\phi(t)T_{\mathrm{ps}}T_{\mathrm{pt}}y^j(t) + \sigma^2 T_{\mathrm{ps}}T_{\mathrm{pt}}\mu_\phi(t) + 2\sigma^2\sigma^2_\phi(t)T_{\mathrm{pt}}\sum_{i\sim m}\nu_{i,m}\psi^m(t)\\ &+ 2\sigma^2\sigma^2_\phi(t)T_{\mathrm{ps}}\left(b^0_\phi(t-1)\psi^i(t-1) + a_\phi(t-1)\right) \end{aligned} \tag{22}$$

and

$$C = \sigma^2\sigma^2_\phi(t)T_{\mathrm{ps}}T_{\mathrm{pt}} \tag{23}$$

with $\phi \in \{0,1\}$, $q = 0$ and $j = -1$. $a_\phi(t-1)$ and $b^0_\phi(t-1)$ are the simple forms of $a_{k^i=\phi}(t-1)$ and $b^0_{k^i=\phi}(t-1)$. Therefore, the conditional distribution for the true signal $\psi^i(t)$ is

$$\begin{aligned} &P\left(\psi^i(t)|y^j(t),\psi^i(t+j),a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j),k^i,\psi^{-i}(t),k^{-i},\mu_{k^i}(t),\sigma^2_{k^i}(t),\sigma^2\right)\\ &= N\left(\frac{B}{A},\frac{C}{A}\right) \end{aligned} \tag{24}$$

with $\phi \in \{0,1\}$, $q = 0$ and $j = -1$.

The Markov blanket for the parameter $\mu_{k^i}(t)$ is $\{\mu_*, \sigma^2_*, \psi^i(t), \sigma^2_{k^i}(t), k^i, \psi^{-i}(t), \psi^i(t+j), k^{-i}, a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}$. According to Eq. (19), the parameter $\mu_{k^i}(t)$ is independent of $\psi^{-i}(t)$, $\psi^i(t+j)$, k^{-i}, $a_{k^i}(t+j)$ and $\mathbf{b}_{k^i}(t+j)$. Therefore, its conditional distribution is

$$\begin{aligned} &P\left(\mu_\phi(t)|\{\psi^i(t)\}_{i|k^i=\phi},\sigma^2_\phi(t),\mu_*,\sigma^2_*,\psi^{-i}(t),\psi^i(t+j),k^{-i},a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j)\right)\\ &\propto P\left(\boldsymbol{\psi}_\phi(t)|\mu_\phi(t),\sigma^2_\phi(t)\right)P\left(\mu_{k^i}(t)|\mu_*,\sigma^2_*\right) \end{aligned} \tag{25}$$

with $\phi \in \{0,1\}$. $P(\boldsymbol{\psi}_\phi(t)|\mu_\phi(t),\sigma^2_\phi(t))$ can be found in Eq. (6). Meanwhile, $P(\mu_{k^i}(t)|\mu_*,\sigma^2_*)$ can be found in Eq. (7). Since we choose a conjugate prior for $\mu_{k^i}(t)$, its conditional distribution is

$$\begin{aligned} &P\left(\mu_\phi(t)|\{\psi^i(t)\}_{i|k^i=\phi},\sigma^2_\phi(t),\mu_*,\sigma^2_*\right) =\\ &\mathrm{N}\left(\left(\frac{1}{\sigma^2_*}+\frac{N_\phi}{\sigma^2_\phi(t)}\right)^{-1}\left(\frac{\mu_*}{\sigma^2_*}+\frac{\sum_{i|k^i=\phi}\psi^i(t)}{\sigma^2_\phi(t)}\right),\left(\frac{1}{\sigma^2_*}+\frac{N_\phi}{\sigma^2_\phi(t)}\right)^{-1}\right) \end{aligned} \tag{26}$$

with $\phi \in \{0,1\}$ and N_ϕ is the number of pixels which satisfy label $k^i = \phi$.

The Markov blanket for the parameter $\sigma^2_{k^i}(t)$ is $\{\alpha_*, \beta_*, \psi^i(t), \mu_{k^i}(t), k^i, \psi^{-i}(t), \psi^i(t+j), k^{-i}, a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}$. Similarly, according to Eq. (19), the parameter $\sigma^2_{k^i}(t)$ is independent of $\psi^{-i}(t)$, $\psi^i(t+j)$, k^{-i}, $a_{k^i}(t+j)$ and $\mathbf{b}_{k^i}(t+j)$. In this way, its conditional distribution is

$$\begin{aligned} &P\left(\sigma^2_\phi(t)|\{\psi^i(t)\}_{i|k^i=\phi},\mu_\phi(t),\alpha_*,\beta_*,\psi^{-i}(t),\psi^i(t+j),k^{-i},a_{k^i}(t+j),\mathbf{b}_{k^i}(t+j)\right)\\ &\propto P\left(\boldsymbol{\psi}_\phi(t)|\mu_\phi(t),\sigma^2_\phi(t)\right)P\left(\sigma^2_{k^i}(t)|\alpha_*,\beta_*\right) \end{aligned} \tag{27}$$

with $\phi \in \{0,1\}$. $P\left(\boldsymbol{\psi}_\phi(t)|\mu_\phi(t),\sigma^2_\phi(t)\right)$ can be found in Eq. (6). Meanwhile, $P\left(\sigma^2_{k^i}(t)|\alpha_*,\beta_*\right)$ can be found in Eq. (8). Similarly, we choose

a conjugate prior for $\sigma^2_{k^i}(t)$, i.e. an inverse-gamma (IG) distribution IG(α_*, β_*), so that its conditional distribution is

$$P\left(\sigma^2_{k^i}(t)|\{\psi^i(t)\}_{i|k^i=\phi}, \mu_\phi(t), \alpha_*, \beta_*\right) = \mathrm{IG}\left(\alpha_* + \frac{N_\phi}{2}, \beta_* + \frac{\sum_{i|k^i=\phi}(\psi^i(t) - \mu_\phi(t))^2}{2}\right) \tag{28}$$

with $\phi \in \{0, 1\}$ and N_ϕ is the number of pixels which satisfy label $k^i = \phi$.

The Markov blanket for the parameter σ^2 is $\{y^i(t), \psi^i(t), \alpha_{**}, \beta_{**}\}$. Its conditional distribution is

$$P\left(\sigma^2|\{y^i(t)\}_{i,t}, \{\psi^i(t)\}_{i,t}, \alpha_{**}, \beta_{**}\right) \propto \prod_{i=1}^{N}\prod_{t=1}^{M} P\left(y^i(t)|\psi^i(t), \sigma^2\right) P\left(\sigma^2|\alpha_{**}, \beta_{**}\right) \tag{29}$$

where $P(y^i(t)|\psi^i(t), \sigma^2)$ can be found in Eq. (3) and $P(\sigma^2|\alpha_{**}, \beta_{**})$ can be found in Eq. (4). Since we choose a conjugate prior for σ^2, its conditional distribution is

$$P\left(\sigma^2|\{y^i(t)\}_{i,t}, \{\psi^i(t)\}_{i,t}, \alpha_{**}, \beta_{**}\right) = \mathrm{IG}\left(\alpha_{**} + \frac{N}{2}, \beta_{**} + \frac{\sum_{i=1}^{N}\sum_{t=1}^{M}\left(y^i(t) - \psi^i(t)\right)^2}{2}\right). \tag{30}$$

The Markov blanket for the parameter $a_{k^i}(t+j)$ is $\{\psi^i(t), \psi^i(t+j), \mathbf{b}_{k^i}(t+j), \mu_{**}, \sigma^2_{**}, k^i, \psi^{-i}(t), k^{-i}, \mu_{k^i}(t), \sigma^2_{k^i}(t)\}$. According to Eq. (19), $a_{k^i}(t+j)$ is independent of $\psi^{-i}(t)$, k^{-i}, $\mu_{k^i}(t)$ and $\sigma^2_{k^i}(t)$. Therefore, its conditional distribution is

$$P\left(a_\phi(t+j)|\{\psi^i(t), \psi^i(t+j)\}_{i|k^i=\phi}, \mathbf{b}_\phi(t+j), \mu_{**}, \sigma^2_{**}, \psi^{-i}(t), k^{-i}, \mu_{k^i}(t), \sigma^2_{k^i}(t)\right) \propto \prod_{i|k^i=\phi} P(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)) P(a_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}) \tag{31}$$

where $P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right)$ can be found in Eqs. (15) and (16). $P\left(a_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}\right)$ can be found in Eq. (18). Given $j = -1$ and $q = 0$, the conditional distribution for $a_\phi(t-1)$ is

$$P\left(a_\phi(t-1)|\{\psi^i(t)\}_{i|k^i=\phi}, \{\psi^i(t-1)\}_{i|k^i=\phi}, b^0_\phi(t-1), \mu_{**}, \sigma^2_{**}\right) = \mathrm{N}(D, E) \tag{32}$$

where

$$D = \left(\frac{1}{\sigma^2_{**}} + \frac{2N_\phi}{T_{\mathrm{pt}}}\right)^{-1}\left(\frac{\mu_{**}}{\sigma^2_{**}} + \frac{2\sum_{i|k^i=\phi}\left(\psi^i(t) - b^0_{k^i=\phi}(t-1)\psi^i(t-1)\right)}{T_{\mathrm{pt}}}\right) \tag{33}$$

and

$$E = \left(\frac{1}{\sigma^2_{**}} + \frac{2N_\phi}{T_{\mathrm{pt}}(t-1)}\right)^{-1} \tag{34}$$

with $\phi \in \{0, 1\}$.

The Markov blanket for $\mathbf{b}_{k^i}(t+j)$ is $\{\psi^i(t), \psi^i(t+j), a_{k^i}(t+j), \mu_{**}, \sigma^2_{**}, k^i, \psi^{-i}(t), k^{-i}, \mu_{k^i}(t), \sigma^2_{k^i}(t)\}$. According to Eq. (19), $\mathbf{b}_{k^i}(t+j)$ is independent of $\psi^{-i}(t)$, k^{-i}, $\mu_{k^i}(t)$ and $\sigma^2_{k^i}(t)$. Therefore, its conditional distribution is

$$P\left(\mathbf{b}_\phi(t+j)|\{\psi^i(t), \psi^i(t+j)\}_{i|k^i=\phi}, a_\phi(t+j), \mu_{**}, \sigma^2_{**}, \psi^{-i}(t), k^{-i}, \mu_{k^i}(t), \sigma^2_{k^i}(t)\right) \propto \prod_{i|k^i=\phi} P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right) P\left(\mathbf{b}_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}\right) \tag{35}$$

where $P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right)$ can be found in Eqs. (15) and (16). Given $j = -1$ and $q = 0$, $P\left(b^0_{k^i}(t+j)|\mu_{**}, \sigma^2_{**}\right)$ can be found in Eq. (17). Therefore, the explicit form of the conditional distribution for $b^0_{k^i=\phi}(t-1)$ is

$$P\left(b^0_\phi(t-1)|\{\psi^i(t)\}_{i|k^i=\phi}, \{\psi^i(t-1)\}_{i|k^i=\phi}, a_\phi(t-1), \mu_{**}, \sigma^2_{**}\right) = \mathrm{N}\left(\frac{B'}{A'}, \frac{C'}{A'}\right) \tag{36}$$

where

$$A' = T_{\mathrm{pt}} + 2\sigma^2_{**}\sum_{i|k^i=\phi}\left(\psi^i(t-1)\right)^2 \tag{37}$$

$$B' = \mu_{**}T_{\mathrm{pt}} + 2\sigma^2_{**}\sum_{i|k^i=\phi}\psi^i\left(\psi^i(t) - a_{k^i}(t-1)\right) \tag{38}$$

and

$$C' = \sigma^2_{**}T_{\mathrm{pt}} \tag{39}$$

The Markov blanket for k^i is $\{\psi^i(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j), \mu_{k^i}(t), \sigma^2_{k^i}(t), k^{-i}, \psi^{-i}(t), \mu_*, \sigma^2_*, \mu_{**}, \sigma^2_{**}, \alpha_*, \beta_*\}$. According to Eqs. (7), (8), (17) and (18), k^i is independent of parameters μ_*, σ^2_*, μ_{**}, σ^2_{**}, α_*, β_*. In this way, its conditional distribution is

$$\begin{aligned} &P\left(k^i = \phi|k^{-i}, \{\psi^i(t), \mu_\phi(t), \sigma^2_\phi(t), \psi^{-i}(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}_t\right) \\ &\propto P(k^i|k^{-i})\prod_{t=1}^{M} P\left(\psi^i(t)|\mu_\phi(t), \sigma^2_\phi(t)\right) P\left(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i}\right) \\ &\prod_{t=1}^{M} P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right) \end{aligned} \tag{40}$$

where $P(k^i|k^{-i})$ can be found in Eqs. (9), (10) and (11). Meanwhile, $P\left(\psi^i(t)|\mu_\phi(t), \sigma^2_\phi(t)\right)$ can be found in Eq. (5). $P\left(\psi^i(t)|\psi^{-i}(t), k^i, k^{-i}\right)$ can be found in Eq. (12), (13), (14). $P\left(\psi^i(t)|\psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\right)$ can be found in Eqs. (15) and (16). Since k^i can only be either 0 or 1, we set it to be a Bernoulli distribution conditional on its Markov blanket. Given $j = -1$ and $q = 0$, when $k^i = 0$, we have

$$\begin{aligned} q \propto &P\left(k^i = 0|k^{-i}, \{\psi^i(t), \mu_{k^i}(t), \sigma^2_{k^i}(t), \psi^{-i}(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}_t\right) \\ &\propto \exp\left(-\frac{1}{T}\sum_{i\sim m} u\left(k^i|k^m\right)\right)\exp\left(-\sum_{t=1}^{M}\frac{\left(\psi^i(t) - \mu_{k^i}(t)\right)^2}{2\sigma^2_{k^i}(t)}\right) \\ &\exp\left(-\sum_{t=1}^{M}\frac{1}{T_{\mathrm{ps}}}\sum_{i\sim m} v_{i,m}\left(\psi^i(t) - \psi^m(t)\right)^2\right) \\ &\exp\left(-\sum_{t=1}^{M}\frac{1}{T_{\mathrm{pt}}}\left(\psi^i(t) - b^0_{k^i}(t-1)\psi^i(t-1) - a_{k^i}(t-1)\right)^2\right). \end{aligned} \tag{41}$$

Similarly, given $j = -1$ and $q = 0$, when $k^i = 1$, we have

$$\begin{aligned} p \propto &P\left(k^i = 1|k^{-i}, \{\psi^i(t), \mu_{k^i}(t), \sigma^2_{k^i}(t), \psi^{-i}(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}_t\right) \\ &\propto \exp\left(-\frac{1}{T}\sum_{i\sim m} u\left(k^i|k^m\right)\right)\exp\left(-\sum_{t=1}^{M}\frac{\left(\psi^i(t) - \mu_{k^i}(t)\right)^2}{2\sigma^2_{k^i}(t)}\right) \\ &\exp\left(-\sum_{t=1}^{M}\frac{1}{T_{\mathrm{ps}}}\sum_{i\sim m} v_{i,m}\left(\psi^i(t) - \psi^m(t)\right)^2\right) \\ &\exp\left(-\sum_{t=1}^{M}\frac{1}{T_{\mathrm{pt}}}\left(\psi^i(t) - b^0_{k^i}(t-1)\psi^i(t-1) - a_{k^i}(t-1)\right)^2\right). \end{aligned} \tag{42}$$

Therefore, the explicit conditional distribution for k^i is

$$P(k^i|k^{-i}, \{\psi^i(t), \mu_{k^i}(t), \sigma^2_{k^i}(t), \psi^{-i}(t), \psi^i(t+j), a_{k^i}(t+j), \mathbf{b}_{k^i}(t+j)\}_t) = \left(\frac{p}{p+q}\right)^{k^i}\left(\frac{q}{p+q}\right)^{1-k^i} \tag{43}$$

The conditional distributions of $\psi^i(t)$, $\mu_{k^i}(t)$, $\sigma^2_{k^i}(t)$, σ^2, $a_{k^i}(t+j)$, $\mathbf{b}_{k^i}(t+j)$ and k^i has been derived in this section, and these conditional distributions will be used for MCMC simulations.

3.5. MCMC simulations

In the present work, we applied the proposed GST-HBM method to both clinical and synthetic data. In practice, the samples of the

GST-HBM parameters can be approximately drawn using an MCMC method, the Gibbs sampling method, from their posterior distributions. Specifically, we have already derived the conditional distributions of all GST-HBM parameters in Section 3.4, which is used in the Gibbs sampling scheme described in detail in Algorithm 1.

Our algorithm has been implemented in Python. Some basic libraries, e.g. NumPy, pandas, and Scipy, have been used to implement the Gibbs sampling routine. The benchmark method "GMM" has been implemented using the "sklearn.mixture.GaussianMixture" package. All remaining code has been developed by the first author. The code has been shared on GitHub https://github.com/legendfuture/GST-HBM/ in line with editorial policy.

Algorithm 1: Gibbs sampling for the GST-HBM method

Input: Hyperparameters α_*, β_*, α_{**}, β_{**}, μ_*, σ_*^2, μ_{**}, σ_{**}^2, T, T_{ps}, T_{pt}
Output: $\{\psi^i(t)\}_{i,t}$, $\{\mu_{k^i}(t)\}_t$, $\{\sigma_{k^i}^2(t)\}_t$, σ^2, $\{a_{k^i}(t+j)\}_t$, $\{\mathbf{b}_{k^i}(t+j)\}_t$ and $\{k^i\}_i$
Data: $\{y^i(t)\}_{i,t}$ where $i = 1, 2, ..., N$ and $t = 1, 2, ..., M$
Set the initial values of $\psi^i(t)$ to be $y^i(t)$; set the initial value of σ^2 to be 1;
for $i \leftarrow 1$ **to** N **do**
  Based on the image at t_{m} (the time that the signals achieve to peak values), obtain initial values of k^i, using the EM algorithm (Gaussian mixture model) (see details from Chapter 9 in Bishop (2006));
end
for $t \leftarrow 1$ **to** M **do**
  Given k^i and $\psi^i(t)$, obtain the initial values of $\mu_{k^i}(t)$ and $\sigma_{k^i}^2(t)$ based on maximum likelihood estimation (MLE);
  Given k^i and $\psi^i(t)$, obtain the initial values of $a_{k^i}(t+j)$ and $\mathbf{b}_{k^i}(t+j)$ using least-squares fitting;
end
for $p \leftarrow 1$ **to** M_0 *(The number of MCMC samples)* **do**
  for $t \leftarrow 1$ **to** M **do**
    for $i \leftarrow 1$ **to** N **do**
      Given $y^i(t)$, $\psi^i(t+j)$, $a_{k^i}(t+j)$, $\mathbf{b}_{k^i}(t+j)$, $\mu_{k^i}(t)$, $\sigma_{k^i}^2(t)$, k^i, k^{-i}, $\psi^{-i}(t)$, T_{ps}, T_{pt} and σ^2, draw $\psi^i(t)$ based on equation (24);
    end
    Given $\{\psi^i(t)\}_i$, $\sigma_{k^i}^2(t)$, $\{k^i\}_i$, μ_* and σ_*^2, draw $\mu_{k^i}(t)$ based on equation (26);
    Given $\{\psi^i(t)\}_i$, $\mu_{k^i}(t)$, $\{k^i\}_i$, α_* and β_*, draw $\sigma_{k^i}^2(t)$ based on equation (28);
    Given $\{\psi^i(t)\}_i$, $\{\psi^i(t+j)\}_i$, $\mathbf{b}_{k^i}(t+j)$, T_{pt}, k^i, μ_{**} and σ_{**}^2, draw $a_{k^i}(t+j)$ based on equation (32);
    Given $\{\psi^i(t)\}_i$, $\{\psi^i(t+j)\}_i$, $a_{k^i}(t+j)$, T_{pt}, k^i, μ_{**} and σ_{**}^2, draw $\mathbf{b}_{k^i}(t+j)$ based on equation (36);
  end
  for $i \leftarrow 1$ **to** N **do**
    Given $\{\psi^i(t)\}_t$, $\{\mu_{k^i}(t)\}_t$, $\{\sigma_{k^i}^2(t)\}_t$, k^{-i}, $\{\psi^{-i}(t)\}_t$, $\{\psi^i(t+j)\}_t$, $\{a_{k^i}(t+j)\}_t$, $\{\mathbf{b}_{k^i}(t+j)\}_t$, T, T_{ps} and T_{pt}, draw k^i based on equation (43);
  end
  Given $\{y^i(t)\}_{i,t}$, $\{\psi^i(t)\}_{i,t}$, α_{**} and β_{**}, draw σ^2 based on equation (30);
end

4. Results

4.1. Results for synthetic data

4.1.1. Model selection for hyperparameters

The values of the hyperparameters of GMM-MRF, M-HBM and GST-HBM need to be selected based on the training data. For GMM-MRF and M-HBM, the details are presented in Yang et al. (2022b,a). For the GST-HBM model proposed in this work, the hyperparameters are selected based on WAIC, as discussed in Section 4.2.1. The hyperparameters of different combinations, i.e. lesion type and noise variance, are independently selected by WAIC.

4.1.2. Visual inspection

The classification results for the synthetic data with high noise (noise variance 2.5^2) are presented in Fig. 7. They can be summarised as follows:

- The classification maps obtained with the GMM are always substantially different from the ground truth. The reason for this poor performance is that the GMM only makes use of the SI of the image, but unlike the other methods included in our study, it does not exploit spatial information to suppress singular and spurious small clusters.
- The classification maps of methods $\text{GMM}_{\text{C\&O}}^{\max}$ and GMM-MRF are visually similar to the ground truth for fully transmural and 180° cases, but noticeably different from the ground truth for non-transmural, 60° and 120° cases. This can be explained by the fact that the method $\text{GMM}_{\text{C\&O}}^{\max}$ cannot make use of the SI from the image, and its classification is therefore solely dependent on the classification map of the GMM. Since the GMM method had high numbers of misclassified pixels, as seen in Figs. 7(h), (n) and (t), the corresponding maps generated by $\text{GMM}_{\text{C\&O}}^{\max}$, i.e. Figs. 7(i), (o) and (u), turned out to be also highly inaccurate. The method GMM-MRF, on the other hand, can make use of the SI from the image. However, its MAP was obtained using the ICM algorithm, which is an approximate greedy optimisation method that is susceptible to entrapment in local optima.
- The classification maps of the two HBM based methods (M-HBM and GST-HBM) were quite similar to the ground truth visually. However, the GST-HBM outperformed the M-HBM, with the latter still suffering from singular and spurious small clusters in Figs. 7(q) and (w). Panel (w) has a larger version of Fig. 8(a) for easier inspection.

The accuracy of the different methods for other degrees of noise was found to be consistent with the high-noise scenario presented above. The visual inspection can provide subjective and intuitive insight into the classification performance. For objective quantification, further comprehensive analysis of the in silico prediction accuracy is presented in Section 4.1.3.

4.1.3. Prediction accuracy

To quantify the classification accuracy, we first counted the number of misclassified pixels and then calculated the cross-entropy between the ground truth labels and the predicted posterior probabilities:

$$H(p_i, q_i) = -\sum_i [p_i \log(q_i) + (1 - p_i) \log(1 - q_i)]. \tag{44}$$

In Eq. (44), the subscript i indicates the pixel index. $p_i \in \{0, 1\}$ indicates the true label given the ground truth. q_i indicates the predicted posterior probabilities. The reason we used the cross-entropy as an alternative quantitative criterion, in addition to the misclassification rate, is to assess if the methods can properly deal with uncertainty quantification, i.e. correctly predict the uncertainty of their predictions. This shifts the focus from the classification label to the posterior probability of finding a pixel in a given class. For example, predicting the posterior probability of a pixel to be lesion to be 0.51 indicates a considerably lower level of confidence than predicting a posterior probability of, say, 0.99. This difference is duly captured by the cross-entropy, but not the misclassification rate, which would not distinguish between the above two scenarios. When the noise level is high, e.g. a noise variance of 2.5^2 or 3^2, the methods may fail to delineate the lesion. The calculated number of misclassified pixels and the average

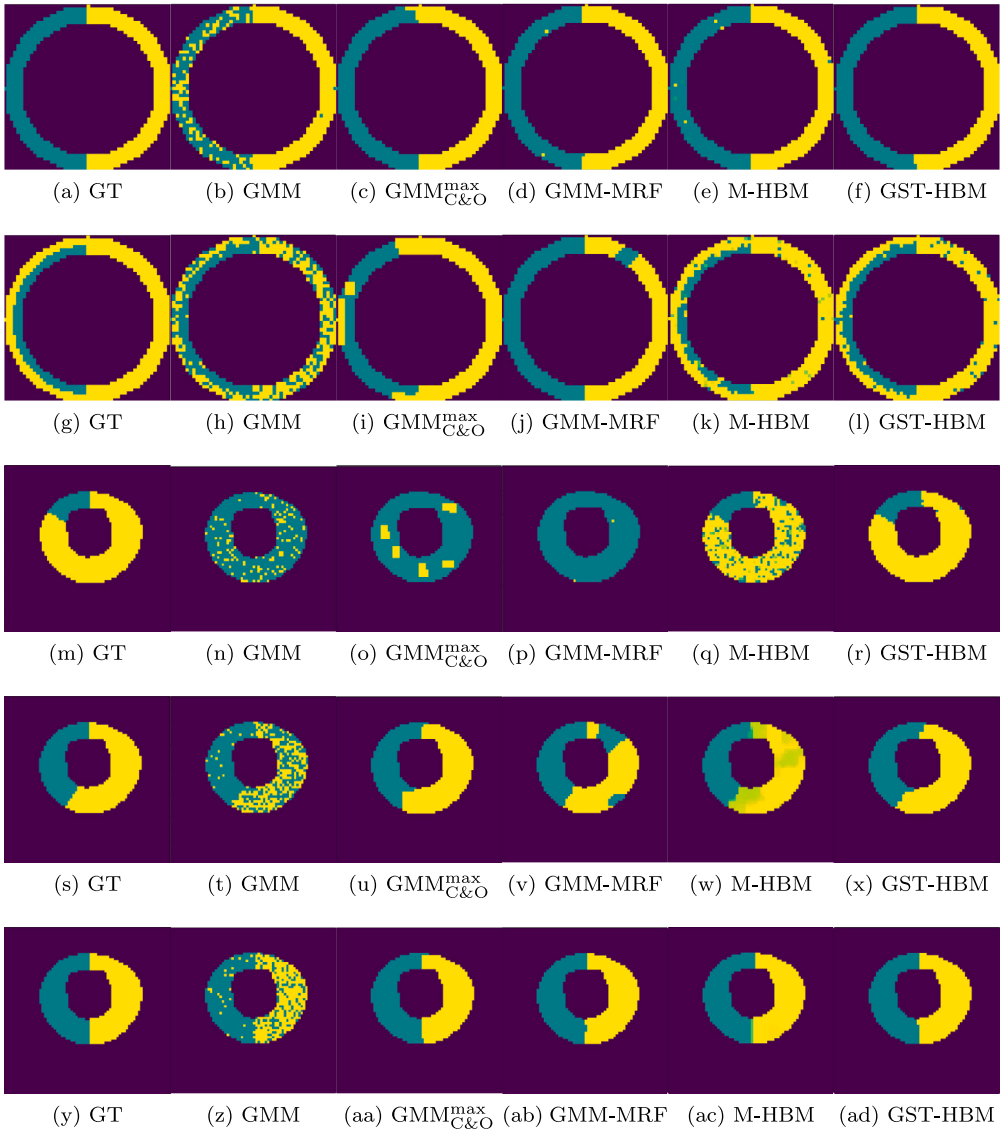


Fig. 7. These figures show lesion delineations for the synthetic examples with additive i.i.d. Rician noise with variance 2.5^2 (high noise). The pixels inside and outside of the myocardium are background (indicated by black). The five rows denote the classification results for fully transmural, non-transmural, 60°, 120° and 180° respectively. Panels (a), (g), (m), (s) and (y) show the ground truth (GT). Panels (b) - (e), (h) - (k), (n) - (q), (t) - (w) and (z) - (ac) show the lesion delineation using alternative methods, which are explicitly explained in Section 2.4 and Table 1. Panels (f), (l), (r), (x) and (ad) show the lesion delineation using the proposed GST-HBM. The abbreviations and their corresponding descriptions of the methods can be found in Table 1.

cross-entropy do not provide a complete picture for quantifying method performance, as they have been averaged over several images. For example, consider three images and two classification methods, method A and method B. Assume that for method A, we have obtained the cross-entropy values 0.1, 0.1 and 10, for each of the three images in turn, while for method B, the corresponding cross-entropy values are 3, 3 and 3. The average cross-entropy for method A is higher than for method B. However, now assume that a cross-entropy higher than 1 is an indicator of failure. In that case, method A has achieved two successful classifications (for images 1 and 2), whereas method B has none. This motivates the introduction of the number of failure cases (defined as a cross-entropy higher than 1) as a new performance score. Since there are 10 separate clinical data sets for each type of design, a score of 0 means that this method has successfully delineated the lesions for all cases, whereas a score of 10 indicates that the method has completely failed to delineate the lesions.

Table 3 shows the average number of misclassified pixels, the cross-entropy and the number of failure cases. It can be seen that the performance of the studied methods varies across the synthetic data sets. For the non-transmural case, the in silico prediction accuracy ranks (from worst to best) are: GMM$^{max}_{C\&O}$, GMM, GMM-MRF, M-HBM, GST-HBM. For the other four types of synthetic data, according to the results in Table 3, the in silico prediction accuracy ranks (from worst to best) are: GMM, GMM-MRF, GMM$^{max}_{C\&O}$, M-HBM, GST-HBM. Here is a summary of some interesting patterns:

- The GMM always performed poorly (large number of misclassified pixels, high cross-entropy and high number of failure cases), and this conclusion is consistent with the visual inspection.
- The method GMM$^{max}_{C\&O}$ performed poorly for the non-transmural case, even when the noise was low (var 1.5^2 and 2^2). The reason is that the myocardial wall is very thin, and the C&O operations transformed the non-transmural lesion in the GMM classification map to a spurious fully transmural lesion.
- Although overall the in silico prediction accuracy of GMM-MRF trailed behind that of GMM$^{max}_{C\&O}$ for four types of synthetic data

Table 3
The three values in the table are the average number of misclassified pixels, the cross-entropy between true labels and predicted posterior probabilities (smaller values mean better results) and the number of failure cases (cross-entropy higher than 1) respectively. The entry "NA" means that none of the methods can generate visually correct classification maps. The smallest value in each column has been marked by bold fonts, with the following priority order: number of failure cases, cross-entropy and average number of misclassified pixels.

Fully transmural	VAR (Rician noise)			
	1.5^2	2^2	2.5^2	3^2
GMM	8.7/0.41/0	63.5/3.06/10	100.6/4.86/10	239.5/11.69/10
$GMM^{max}_{C\&O}$	0.3/0.01/0	3.8/0.17/0	7.4/0.34/0	118.6/5.70/10
GMM-MRF	**0.2/0.00/0**	1.3/0.04/0	12.8/0.61/2	213.9/10.34/6
M-HBM	0.6/0.01/0	2.16/0.03/0	12.1/0.40/0	156/6.10/4
GST-HBM	0.4/0.02/0	**0.6/0.03/0**	**2.4/0.10/0**	**21.6/1.04/4**
Non-transmural	VAR (Rician noise)			
	1.5^2	2^2	2.5^2	3^2
GMM	10.7/0.5/0	21.4/1.04/7	190.1/9.21/10	504.8/24.4/10
$GMM^{max}_{C\&O}$	105.6/5.11/10	134.9/6.53/10	150.5/7.27/10	188.7/9.10/10
GMM-MRF	2.4/0.09/0	6.9/0.31/0	209.6/10.13/10	537.8/26.02/10
M-HBM	7.4/0.09/0	8.7/0.27/0	**132.9/2.86/8**	**421.1/20.39/8**
GST-HBM	**2.4/0.08/0**	**5.8/0.27/0**	58.1/2.81/10	155.7/3.70/10
60°	VAR (Rician noise)			
	1.5^2	2^2	2.5^2	3^2
GMM	6.9/0.27/0	147.7/5.41/8	698.4/25.60/10	NA
$GMM^{max}_{C\&O}$	4.0/0.15/0	9.5/0.36/1	550.9/20.4/10	NA
GMM-MRF	**0.6/0.01/0**	11.1/0.41/1	792.1/29.1/10	NA
M-HBM	2.5/0.08/0	63.2/0.27/0	139.1/4.00/10	NA
GST-HBM	1.4/0.04/0	**5.7/0.22/0**	**30.1/1.10/4**	NA
120°	VAR (Rician noise)			
	1.5^2	2^2	2.5^2	3^2
GMM	10.8/0.41/0	43.2/1.60/10	219.6/8.04/10	379.9/13.94/10
$GMM^{max}_{C\&O}$	2.3/0.06/0	11.5/0.43/0	12.5/0.46/0	27.2/0.98/5
GMM-MRF	0.7/0.01/0	1.8/0.04/0	30.9/1.12/2	427.7/15.70/10
M-HBM	**0.6/0.00/0**	2.0/0.02/0	**14.3/0.05/0**	163.1/2.29/10
GST-HBM	0.7/0.02/0	**1.1/0.02/0**	16.2/0.60/0	**20.8/0.77/2**
180°	VAR (Rician noise)			
	1.5^2	2^2	2.5^2	3^2
GMM	11.4/0.42/0	48.1/1.75/10	175.0/6.43/10	291.7/10.70/10
$GMM^{max}_{C\&O}$	0.00/0.00/0	1.4/0.05/0	9.5/0.35/0	44.1/1.63/7
GMM-MRF	0.5/0.00/0	1.4/0.04/0	**2.0/0.08/0**	202.5/7.40/5
M-HBM	0.9/0.02/0	**0.6/0.01/0**	10.0/0.31/0	**79.2/0.63/5**
GST-HBM	**0.00/0.00/0**	0.9/0.04/0	6.3/0.24/0	28.7/1.28/5

(all except the non-transmural case), GMM-MRF showed a decent performance when the noise was low (noise variance with 1.5^2 and 2^2). In fact, this method performed as well as the two HBM methods when the noise was low.

- M-HBM has sometimes a large number of misclassified pixels, e.g. for the non-transmural and 60° cases in Table 3, but achieves a low cross-entropy. The reason is that this method sometimes predicted large areas of uncertain classification, e.g. in Fig. 7(w). While this increases the number of misclassified pixels, it leads to a decrease in cross-entropy.

Note that with the sole exception of GMM, all methods included in our comparative evaluation depend on various regularisation parameters. A disadvantage of the mathematical morphology method, $GMM^{max}_{C\&O}$, is that the regularisation parameters for this method are "user-defined", meaning that they have to be set by the user based on intuition, experience, and less objective criteria. On the other hand, all the other methods, i.e. GMM-MRF, M-HBM and GST-HBM, are probabilistic models. The regularisation parameters can be set objectively, based on the available data and established statistical information criteria. A more detailed discussion of the disadvantages of the $GMM^{max}_{C\&O}$ can be found in the online supplementary materials.

Both the qualitative (visual inspection) and quantitative (in silico prediction accuracy) analysis of the performance of the different methods on different types of synthetic data lead to the following recommendations:

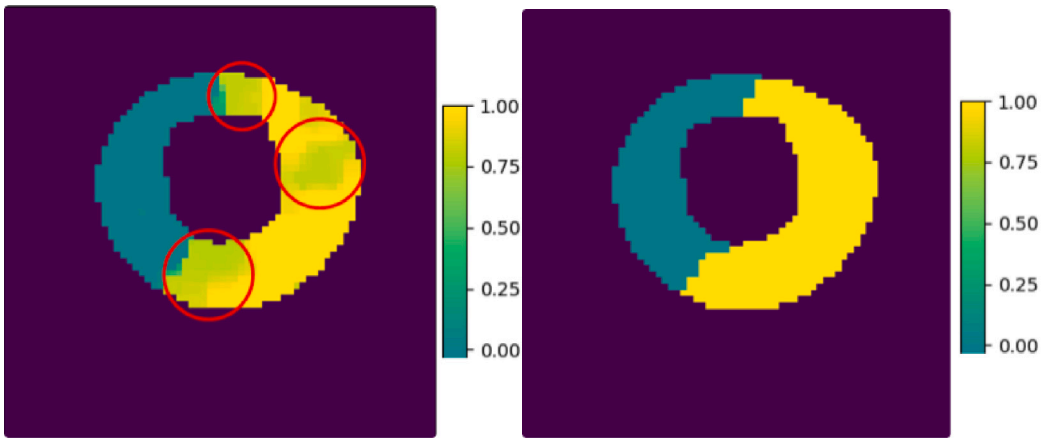
- For fully transmural data (standard shape or clinical shape), if the noise is low, GMM-MRF is recommended because it is as accurate as the GST-HBM method, but computationally much faster (running time). If the noise is high, GST-HBM is recommended due to its highest accuracy.
- For non-transmural data, GST-HBM is recommended because it has the highest in silico prediction accuracy.
- If physiological parameters, like the MBF, are needed, then M-HBM is recommended, as it has the second highest in silico prediction accuracy and provides the option to estimate the physiological parameters.

Both M-HBM and GST-HBM methods can provide quantified uncertainty of the classification maps. The quantified uncertainty, i.e. calculated posterior probability, can be used to measure the degree of certainty of the method. In the present work, the GST-HBM has been proven to give more certain results compared with the M-HBM. A certain result means that the posterior probability of a pixel is located in $[0, 0.05]$ (lesions) or $[0.95, 1]$ (healthy tissues). On the other hand, an uncertain result means that the posterior probability of a pixel is located in the intermediate range $(0.05, 0.95)$. Figs. 8(a) and 8(b) show the classification maps to the same synthetic dataset using M-HBM and GST-HBM respectively. It can be seen that the model-based HBM gave uncertain classifications in the red circles, and the GST-HBM gave very certain classification.

4.2. Results for clinical data

4.2.1. Model selection for hyperparameters

In the proposed GST-HBM method, there are three user-defined hyperparameters T, T_{ps} and T_{pt}, which can be found in Eqs. (10), (13), (16). Firstly, we used WAIC values to explore the best combination of the hyperparameters. In principle, they could be sampled from the posterior distribution with MCMC techniques, along with the other parameters.To achieve this, they could be given exponential distributions as prior. These prior distributions are conjugate for their likelihoods. Therefore, the posteriors are also in the family of exponential distribution and can be directly sampled. However, this would lead to a substantial increase in the computational complexity, given that convergence and mixing of hyperparameters tends to be happening at much lower rates than for parameters; see e.g. Neal (1996). In this work, we ran an simulation that T, T_{ps} and T_{pt} were sampled. The convergence speed of the sampling case is at least 5 times slower than the fixed case. Moreover, the values of traceplots are rather unstable, with several continually occurring outliers at extreme values. This will dramatically reduce the stability of the MCMC method and increase the convergence time. The specific details of this simulation can be found in the supplementary materials. For clinical decision making, excessive computational costs need to be avoided, and we also want to better exploit parallel computing resources for computational cost reduction. This can be achieved by computing advanced information criteria, like WAIC, for a set of candidate values in parallel, using high-performance computer clusters, and then selecting the results corresponding to those hyperparameters that have obtained the lowest WAIC score. In our work, we have processed Algorithm 1 separately for different combinations of the hyperparameters on different processors. In our simulations, the hyperparmeters highly influenced the final classification maps. Some typical classification maps are further presented in the supplementary materials. Table 4 shows the WAIC values for different combinations of the hyperparameters. We find that the combination $T = 1000$, $T_{ps} = 0.001$ and $T_{pt} = 100$ achieves the



(a) M-HBM (b) GST-HBM

Fig. 8. Panels (a) and (b) show the posterior probability of M-HBM and GST-HBM (they are the same as Figs. 7(w) and 7(x) with added annotations). The light green pixels inside the red circles indicate areas of uncertain M-HBM classifications in panel (a). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lowest (i.e. best) WAIC value, and these hyperparameter settings were used for this set of data. For other sets of data, the hyperparameters are selected by their WAIC values.

4.2.2. Convergence test

All parameters were sampled from their posterior distributions using Gibbs sampling according to Algorithm 1. Fig. 9 inspects the convergence of the parameter σ^2. The reason we chose this parameter to assess the convergence of the Gibbs sampling is that it can reveal the uncertainty of differences between observed and generated signals. The traceplot for the MCMC samples look reasonable, without any signs of insufficient convergence. We also use Geweke's test (Geweke et al., 1991) to test for convergence. If the Geweke scores consistently stay within the range (−1.96,1.96), there is no significant evidence for insufficient convergence. Therefore, given Fig. 9, there is no significant evidence that the parameter σ^2 has not converged. We also repeat the convergence explorations for other parameters ($\psi^i(t)$, $\mu_{k^i}(t)$, ...). There is no significant evidence for lack of convergence here either. To further assess the MCMC convergence diagnoses, we have run 5 additional independent MCMC simulations on the same data with different random seeds. We found that the independent MCMC simulations present visually very similar classification maps to the presented classification maps in Fig. 10(e). The average Dice score between the proposed classification map and other independent simulations is 0.986 with standard deviation 1.1×10^{-3}. This suggests sufficient convergence of the MCMC simulations. The visualisation of the comparisons can be found in the supplementary materials.

4.2.3. Classification results

Fig. 10 shows the classification maps obtained with all the methods included in our study, i.e. GMM, $\text{GMM}^{\text{max}}_{\text{C\&O}}$, GMM-MRF, M-HBM and GST-HBM. The LGE image in panel (f) is used as the ground truth to indicate the location of the lesion. Although the two types of data, i.e. the dynamic perfusion sequence and the LGE image, cannot be overlayed pixel by pixel, the myocardial defects predicted by these two types of images are positioned in the same part of the myocardium. For example, in Fig. 10(f), the injured myocardium in the red circle is located in the inferior (bottom) and inferolateral (right bottom) segments of the myocardium. Therefore, for the classification maps in panels (a) - (e), the correct classification should label the inferior and inferolateral segments of the myocardium as lesion.

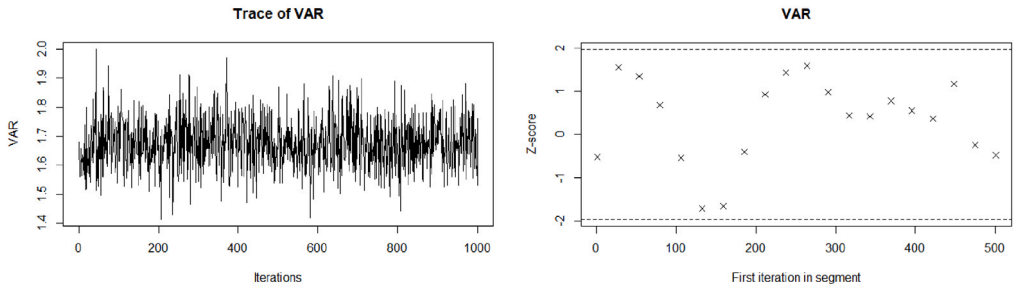
Panels (a) - (e) in Fig. 10 show the classification maps predicted by the various methods. GMM, $\text{GMM}^{\text{max}}_{\text{C\&O}}$, GMM-MRF and M-HBM have all delineated spurious segments in the anterior (top) of the myocardium. However, according to the LGE image in panel (f), there are no lesions located in the anterior of the myocardium (red circle). This suggests that the classification map in panel (e), obtained with GST-HBM, is the most realistic one. The poorer performance of GMM, $\text{GMM}^{\text{max}}_{\text{C\&O}}$ and GMM-MRF is consistent with the performance evaluation on the synthetic data, as discussed in Section 4.1.

The physiological model that forms part of M-HBM is based on the central volume principle (see details in Zierler (1965)), and the solution of the equation underlying the central volume principle is highly affected by the SNR of the baseline pre-contrast phase of the DCE-MRI time series. The reason for the failure of M-HBM to correctly delineate the hypoperfused lesion (Fig. 10(d)) is that the noise distribution is spatially variant, resulting in localised regions with extremely low SNR (e.g. red ROI in Fig. 11(a)). Such regions are vulnerable to adverse error propagation that in turn results in severe misestimation of MBF. As can be seen in Fig. 11(c), M-HBM yields relatively low estimates of MBF in this location (red ROI) despite any apparent abnormality in the corresponding location on reference LGE image (Fig. 10(f)), or on maximum enhancement image in Fig. 11(b).

However, the proposed GST-HBM method overcomes the problem posed by the presence of the spatially variant noise by efficient use of spatio-temporal information that results in denoising of the original images and consequently achieves excellent agreement between the final classification map and reference LGE images (Figs. 10(e) and (f) respectively).

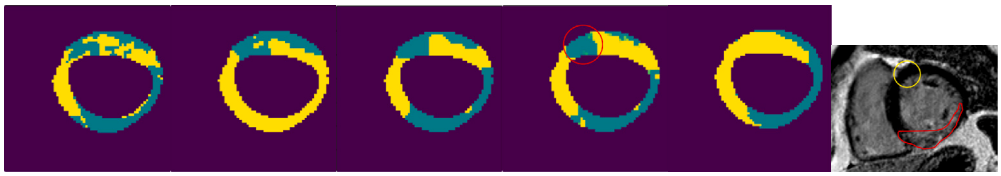
5. Discussion

In the present work, we have proposed a classification approach based on an HBM with spatio-temporal Markov random field prior for automatic lesion detection in myocardial perfusion DCE-MRI scans. This method is a fully data driven approach, meaning that all parameters and hyperparameters can be consistently inferred from the data, without the need for any heuristic user-defined tuning parameters. While the effectiveness of our classification method is demonstrated using raw DCE-MRI myocardial data, its formulation is applicable to any form of calibrated (transformed) datasets, where blood pool signal is used to derive absolute values of MBF (Kellman et al., 2017). Furthermore, the spatial and temporal constraints proposed in this method are absolute-value norm distance and autoregressive model of order 1 respectively. Any other definitions of distance or temporal model can be directly applied to replace the ones given in this paper. This is the



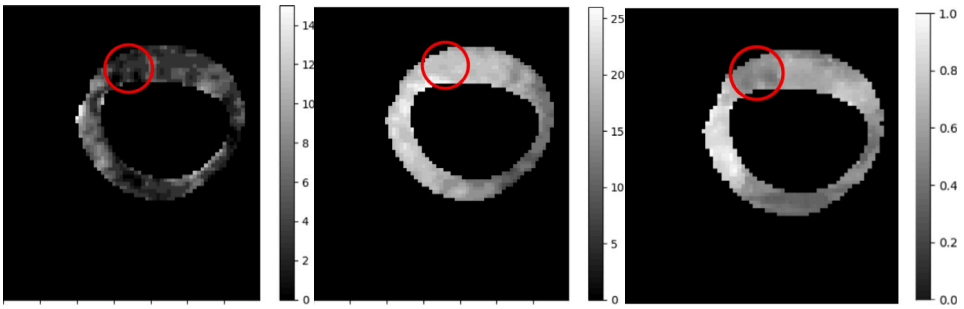
(a) Traceplot of the posterior distribution of σ^2 (b) Geweke test

Fig. 9. These figures inspect the convergence of the MCMC simulation for variance σ^2. (a) shows the traceplot of the MCMC samples for parameter σ^2 after burn-in. (b) shows Geweke scores based on the MCMC samples of σ^2. The dashed horizontal lines show the value 1.96 and −1.96. If the Geweke scores are within the range $(-1.96, 1.96)$, there is no significant evidence that this parameter has not converged.



(a) GMM (b) $\text{GMM}^{\max}_{\text{C\&O}}$ (c) GMM-MRF (d) M-HBM (e) GST-HBM (f) LGE

Fig. 10. Panels (a) - (e) show the classification maps predicted by the methods compared in our study: GMM, $\text{GMM}^{\max}_{\text{C\&O}}$, GMM-MRF, M-HBM and GST-HBM respectively. Yellow pixels indicate healthy tissues, while dark green pixels indicate ischaemic lesions. Panel (f) represents the corresponding LGE image. The region inside the red ROI indicates the location of the infarct. The red circle in panel (d) indicates the misclassification region using the M-HBM method. The yellow circle in panel (f) shows the corresponding region in the LGE image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) Original Pre-contrast (b) Original Maximum Contrast Enhancement (c) MBF Estimated by M-HBM

Fig. 11. Panels (a) and (b) show the original images at the pre-contrast phase and maximum contrast enhancement of the DCE-MRI time series respectively. The grey scale bars indicate the original signal intensities of pixels. Panel (c) shows the map of the MBF estimation (standardised to $[0, 1]$) using M-HBM method. The ROI inside the red circle corresponds to the misclassification region using M-HBM in Fig. 10(d).

reason we named this model "generalisable spatio-temporal" HBM. The only parts of the proposed method that need to be adjusted when applying it to other types of data are the modelling of the observations $\psi^i(t)$, spatial dependence ψ^{-i} and temporal dependence $\psi^i(t+j)$, $a_{ki}(t+j)$ and $\boldsymbol{b}_{ki}(t+j)$ in Table 2. For example, the model of the temporal constraints we used in this work has a form of auto-regressive model. This module of the GST-HBM method could be replaced by a different model of myocardial perfusion DCE-MRI (Biglands et al., 2015). Four alternative established methods, GMM, $\text{GMM}^{\max}_{\text{C\&O}}$ (GMM with opening and closing operations), GMM-MRF and model-based HBM, have been applied for comparison, using both synthetic and clinical data. For the former, the ground truth is known. This allowed us to compute three objective quantitative criteria, the number of misclassified pixels, the cross-entropy between true and predicted labels and the number of failure cases. All quantitative criteria demonstrate that the proposed GST-HBM clearly outperforms the established alternative methods used for comparison, which we further corroborated with a standard statistical hypothesis test.

For the clinical myocardial perfusion DCE-MRI images, we evaluated the performance of the methods based on the LGE image from the same patient. The LGE image has been used to indicate the location of the myocardial infarction that is expected to be the focus of the fixed perfusion defect. We have tested our method using 12 datasets. In some datasets, there is spatially variant noise at the pre-contrast

Table 4
WAIC (smaller value means better result). The smallest one is marked in bold font and superscript ∗.

T = 0.01		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	104 252	104 124	103 842	98 532
	0.001	101 027	100 302	100 305	98 502
	0.01	98 807	98 514	98 216	100 917
	0.1	102 736	101 210	101 380	105 226
	1	111 700	113 505	113 669	123 724
T = 0.1		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	97 661	97 304	97 293	97 217
	0.001	99 637	99 083	98 708	98 578
	0.01	98 677	98 326	98 422	100 556
	0.1	102 381	101 644	101 540	105 437
	1	111 029	113 386	113 428	124 787
T = 1		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	97 329	96 769	97 019	96 149
	0.001	96 051	94 841	94 890	95 485
	0.01	96 923	96 267	96 158	96 572
	0.1	101 281	100 084	99 906	103 518
	1	110 814	113 004	113 248	123 764
T = 10		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	95 198	95 586	95 002	94 510
	0.001	96 177	94 804	94 974	95 249
	0.01	102 746	96 357	96 163	96 477
	0.1	110 980	102 810	101 778	103 518
	1	132 716	128 545	134 199	135 595
T = 100		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	95 545	95 314	95 357	94 601
	0.001	95 800	94 837	94 739	95 176
	0.01	103 013	96 318	95 900	96 394
	0.1	110 922	103 495	101 823	103 036
	1	132 517	129 165	133 887	163 825
T = 1000		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	94 930	94 803	94 795	94 752
	0.001	95 653	94 582	**94487**∗	95 006
	0.01	101 552	95 941	96 051	96 553
	0.1	107 653	99 990	98 736	103 363
	1	131 629	128 897	137 598	166 143
T = 10000		T_{pt}			
		1	10	100	1000
T_{ps}	0.0001	95 121	94 982	94 843	94 599
	0.001	97 135	95 485	95 780	95 677
	0.01	103 827	97 643	96 796	96 796
	0.1	112 721	105 003	103 935	103 762
	1	134 024	128 067	142 491	143 790

phase of the DCE-MRI time series, and this noise can strongly affect the performance of any physiological method that is based on the central volume principle. Consequently, model-based HBM methods that incorporate this principle fail to correctly delineate the lesion. However, the GST-HBM method proposed in this work has proven to succeed in denoising the original image and correctly delineating the lesion.

For inference, we have sampled the parameters from the posterior distribution with Gibbs sampling, while the lower level hyperparameters were selected based on an advanced information criterion (WAIC), which can be directly computed from the Gibbs samples. The motivation for this combined approach over a full sampling scheme is faster mixing and convergence and the facilitation of the exploitation of parallel computer clusters (see further below).

In summary, there are two main advantages of the proposed GST-HBM over all other alternative established methods. Firstly, the in silico prediction accuracy of the proposed GST-HBM is significantly better than for the alternative established methods using the synthetic data, based on the three performance criteria of classification error, cross-entropy and the number of failure cases. Secondly, the proposed GST-HBM is the only method that is able to deal with the myocardial perfusion DCE-MRI images that are subject to spatially varying noise at the pre-contrast phase of the DCE-MRI time series. All alternative methods included in our comparative evaluation study failed to delineate the lesions properly; the reason is that they depend to some extent on either the MBF map generated by the physiological models or the physiological models themselves, which are sensitive to the spatially varying noise. The proposed GST-HBM method has some further advantages over the alternative methods. Compared with the GMM, it can make use of spatial information to remove spurious singleton clusters. It is more robust and consistent than the GMM based opening and closing operations, in that it does not depend on any subjective user-defined tuning parameters. Compared with the GMM-MRF method, its parameters are sampled using MCMC. This is less susceptible to local minima than the greedy optimisation on which GMM-MRF is based.

A further advantage of the proposed GST-HBM method is that it can provide uncertainty quantification, given by the posterior probability that a pixel belongs to one of the two classes (lesion versus healthy tissue). These posterior probabilities are obtained with the Gibbs sampling scheme described in Section 3.5. Quantifying uncertainty is a substantial advantage over a crisp classification of the form “healthy versus lesion”, as it equips a clinician with an important decision support tool for risk estimation and treatment recommendation.

One disadvantage of the proposed GST-HBM is its higher computational complexity. The three methods, i.e. GMM, GMM based opening and closing operations and GMM-MRF, used for comparison are computationally cheap: it typically takes less than a minute to obtain the results. However, the Gibbs sampling simulations for the proposed GST-HBM method, like for the model-based HBM, are computationally expensive. We typically required about 1000 Gibbs sampling steps to reach an acceptable level of convergence, based on the Geweke convergence test, which was equivalent to computational costs of about one hour on the hardware we were using (Intel(R) Core(TM) i9-7900X CPU @ 3.30 GHz processor with 64 GB memory) The main factors that influence the computational costs are both number of time points and image pixels. They are linearly correlated with the product of number of time points and pixels. The tests information can be found in the supplementary materials.

A further disadvantage of the proposed GST-HBM is that it assumes only two types of tissue (healthy tissue versus lesion). In its current version, the GST-HBM cannot subdivide lesions into different subclasses corresponding to different degrees of lesion severity. Lesion severity can be described by two main attributes: lesion morphology (including location, size and transmurality), and the amplitude of pathological changes (in this case, the degree of ischaemia in the affected tissue). Our GST-HBM method currently addresses the first (morphological) component, and provides the framework for assessing the second. We derived GST-HBM classification maps from uncalibrated DCE-MRI data, but these maps could be applied to calibrated MBF data to derive the

absolute amplitude component of ischaemia and add it to the morphological assessment. This would provide a comprehensive assessment of lesion severity using GST-HBM. An extension of the GST-HBM to deal with more than two classes is conceptually straightforward. Specifically, the two groups proposed in the GST-HBM method can be further extended to three groups that contain an additional microvascular obstruction lesion (MVO) class (Bekkers et al., 2010).

To reduce the computational complexity, we can draw on parallelisation: both the pixel-level computations on which the Gibbs sampling scheme is based as well as the WAIC computations for hyperparameter selection can easily be run in parallel. While for the present proof-of-concept study this parallelisation has not been implemented yet, it is conceptually straightforward and can be expected to lead to a substantial reduction in the computational costs.

All methods mentioned in the presented work are dependent on manually contoured endo- and epicardial boundaries (see Section 2.2). The proposed method can make use of automatic as well as manual endo- and epicardial contours without any changes to the algorithm structure. Significant improvement has been made in the domain of automated contour detection for quantitative cardiac magnetic resonance (CMR) analysis in recent years. Automatic contour detection for cardiac DCE-MRI presents additional challenges compared to other components of CMR (Kim et al., 2020; Chen et al., 2008), and forms an essential component of motion-correction. The impact of contouring errors on quantitative analysis of myocardial perfusion has been investigated by Biglands et al. (2011). By virtue of reducing the adverse impact of baseline noise on classification accuracy, our proposed algorithm may have a higher tolerance to contouring errors compared to other benchmark methods. We aim to investigate this issue further in our future work.

6. Conclusion

In the present paper, we have proposed a novel GST-HBM classification method for lesion detection in DCE-MRI images based on a hierarchical Bayesian model with a spatial Markov random field prior. Our method is a fully data driven approach that achieves automatic detection of myocardial ischaemia from DCE-MRI scans. It has been proven to outperform several state-of-the-art benchmark methods on both clinical and synthetic data, and we have quantified the degree of improvement with three different performance criteria.

Funding

This work was funded by the UK Engineering and Physical Sciences Research Council (EPSRC), grant reference numbers EP/S020950/1, EP/S030875/1 and EP/T017899/1. Yalei Yang and Aleksandra Radjenovic were funded by a grant from GlaxoSmithKline R&D. Colin Berry is supported by the British Heart Foundation (RE/18/6134217). This work was done during Yalei Yang's PhD at the University of Glasgow. Now, Yalei Yang is funded by United Imaging Healthcare Co., Ltd and working as a Postdoc at Peking University Cancer Hospital & Institute.

CRediT authorship contribution statement

Yalei Yang: Conceptualization, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Dirk Husmeier:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Hao Gao:** Conceptualization, Software, Supervision, Writing – review & editing. **Colin Berry:** Data curation, Funding acquisition, Resources. **David Carrick:** Investigation, Resources. **Aleksandra Radjenovic:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Synthetic data will be made available upon request. We are unable to authorise access to patient data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compmedimag.2024.102333.

References

Bardsley, J.M., 2012. Laplace-distributed increments, the Laplace prior, and edge-preserving regularization. J. Inverse Ill-Posed Probl. 20 (3), 271–285.
Bekkers, S.C., Yazdani, S.K., Virmani, R., Waltenberger, J., 2010. Microvascular obstruction: underlying pathophysiology and clinical diagnosis. J. Am. Coll. Cardiol. 55 (16), 1649–1660.
Biglands, J., Magee, D., Boyle, R., Larghat, A., Plein, S., Radjenović, A., 2011. Evaluation of the effect of myocardial segmentation errors on myocardial blood flow estimates from DCE-MRI. Phys. Med. Biol. 56 (8), 2423.
Biglands, J.D., Magee, D.R., Sourbron, S.P., Plein, S., Greenwood, J.P., Radjenovic, A., 2015. Comparison of the diagnostic performance of four quantitative myocardial perfusion estimation methods used in cardiac MR imaging: CE-MARC substudy. Radiology 275 (2), 393–402.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer.
Carrick, D., Haig, C., Rauhalammi, S., Ahmed, N., Mordi, I., McEntegart, M., Petrie, M.C., Eteiba, H., Hood, S., Watkins, S., et al., 2015. Prognostic significance of infarct core pathology revealed by quantitative non-contrast in comparison with contrast cardiac magnetic resonance imaging in reperfused ST-elevation myocardial infarction survivors. Eur. Heart J. 37 (13), 1044–1059.
Cerqueira, M.D., Weissman, N.J., Dilsizian, V., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., et al., 2002. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. Circulation 105 (4), 539–542.
Chen, T., Babb, J., Kellman, P., Axel, L., Kim, D., 2008. Semiautomated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI. IEEE Trans. Med. Imaging 27 (8), 1084–1094.
Chen, S., Haralick, R.M., 1995. Recursive erosion, dilation, opening, and closing transforms. IEEE Trans. Image Process. 4 (3), 335–345.
Chen, Z., Lalande, A., Salomon, M., Decourselle, T., Pommier, T., Perrot, G., Couturier, R., 2020. Myocardial infarction segmentation from late gadolinium enhancement MRI by neural networks and prior information. In: 2020 International Joint Conference on Neural Networks. IJCNN, IEEE, pp. 1–8.
Cordero-Grande, L., Vegas-Sánchez-Ferrero, G., Casaseca-de-la Higuera, P., San-Román-Calvar, J.A., Revilla-Orodea, A., Martín-Fernández, M., Alberola-López, C., 2011. Unsupervised 4D myocardium segmentation with a Markov random field based deformable model. Med. Image Anal. 15 (3), 283–301.
Dai, S., Lu, K., Dong, J., Zhang, Y., Chen, Y., 2015. A novel approach of lung segmentation on chest CT images using graph cuts. Neurocomputing 168, 799–807.
de la Rosa, E., Sidibé, D., Decourselle, T., Leclercq, T., Cochet, A., Lalande, A., 2021. Myocardial infarction quantification from late gadolinium enhancement MRI using top-hat transforms and neural networks. Algorithms 14 (8), 249.
Fenchel, M., Helber, U., Kramer, U., Stauder, N.I., Franow, A., Claussen, C.D., Miller, S., 2005. Detection of regional myocardial perfusion deficit using rest and stress perfusion MRI: a feasibility study. Am. J. Roentgenol. 185 (3), 627–635.
Genkin, A., Lewis, D.D., Madigan, D., 2007. Large-scale Bayesian logistic regression for text categorization. Technometrics 49 (3), 291–304.
Geweke, J., et al., 1991. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, Vol. 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN.
Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. Magn. Reson. Med. 34 (6), 910–914.
Hennemuth, A., Seeger, A., Friman, O., Miller, S., Klumpp, B., Oeltze, S., Peitgen, H.-O., 2008. A comprehensive approach to the analysis of contrast enhanced cardiac MR images. IEEE Trans. Med. Imaging 27 (11), 1592–1610.
Hojjatoleslami, S., Kittler, J., 1998. Region growing: a new approach. IEEE Trans. Image Process. 7 (7), 1079–1084.
Ingrisch, M., Sourbron, S., 2013. Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: a primer. J. Pharmacokinet. Pharmacodyn. 40, 281–300.

Jerosch-Herold, M., 2010. Quantification of myocardial perfusion by cardiovascular magnetic resonance. J. Cardiovasc. Magn. Reson. 12 (1), 57.

Jerosch-Herold, M., Swingen, C., Seethamraju, R.T., 2002. Myocardial blood flow quantification with MRI by model-independent deconvolution. Med. Phys. 29 (5), 886–897.

Jerosch-Herold, M., Wilke, N., Stillman, A.E., Wilson, R.F., 1998. Magnetic resonance quantification of the myocardial perfusion reserve with a Fermi function model for constrained deconvolution. Med. Phys. 25 (1), 73–84.

Karim, R., Bhagirath, P., Claus, P., Housden, R.J., Chen, Z., Karimaghaloo, Z., Sohn, H.-M., Rodríguez, L.L., Vera, S., Albà, X., et al., 2016. Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images. Med. Image Anal. 30, 95–107.

Kellman, P., Hansen, M.S., Nielles-Vallespin, S., Nickander, J., Themudo, R., Ugander, M., Xue, H., 2017. Myocardial perfusion cardiovascular magnetic resonance: optimized dual sequence and reconstruction for quantification. J. Cardiovasc. Magn. Reson. 19 (1), 1–14.

Khalifa, F., Soliman, A., El-Baz, A., Abou El-Ghar, M., El-Diasty, T., Gimel'farb, G., Ouseph, R., Dwyer, A.C., 2014. Models and methods for analyzing DCE-MRI: A review. Med. Phys. 41 (12), 124301.

Kim, Y.-C., Kim, K.R., Choe, Y.H., 2020. Automatic myocardial segmentation in dynamic contrast enhanced perfusion MRI using Monte Carlo dropout in an encoder-decoder convolutional neural network. Comput. Methods Programs Biomed. 185, 105150.

Lafferty, J., McCallum, A., Pereira, F.C., 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

Larsson, H.B., Fritz-Hansen, T., Rostrup, E., Søndergaard, L., Ring, P., Henriksen, O., 1996. Myocardial perfusion modeling using MRI. Magn. Reson. Med. 35 (5), 716–726.

Lehnert, J., Kolbitsch, C., Wubbeler, G., Chiribiri, A., Schaeffter, T., Elster, C., 2019. Large-scale Bayesian spatial-temporal regression with application to cardiac MR-perfusion imaging. SIAM J. Imaging Sci. 12 (4), 2035–2062.

Li, S.Z., 2009. Markov Random Field Modeling in Image Analysis. Springer Science & Business Media.

Metzner, S., Wübbeler, G., Elster, C., 2019. Approximate large-scale Bayesian spatial modeling with application to quantitative magnetic resonance imaging. AStA Adv. Stat. Anal. 103 (3), 333–355.

Moon, T.K., 1996. The expectation-maximization algorithm. IEEE Signal Process. Mag. 13 (6), 47–60.

Neal, R., 1996. Bayesian Learning for Neural Networks. Springer.

Radjenovic, A., 2003. Measurement of Physiological Variables by Dynamic Gd-DTPA Enhanced MRI (Ph.D. thesis). University of Leeds.

Scannell, C.M., Chiribiri, A., Villa, A.D., Breeuwer, M., Lee, J., 2020. Hierarchical Bayesian myocardial perfusion quantification. Med. Image Anal. 60, 101611.

Schmid, V.J., Yang, G.-Z., 2009. Spatio-temporal modelling of first-pass perfusion cardiovascular MRI. In: World Congress on Medical Physics and Biomedical Engineering. September 7–12, 2009, Munich, Germany, Springer, pp. 45–48.

Sethian, J.A., 1999. Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Vol. 3. Cambridge University Press.

Tamarappoo, B.K., Dey, D., Nakazato, R., Shmilovich, H., Smith, T., Cheng, V.Y., Thomson, L.E., Hayes, S.W., Friedman, J.D., Germano, G., et al., 2010. Comparison of the extent and severity of myocardial perfusion defects measured by CT coronary angiography and SPECT myocardial perfusion imaging. JACC Cardiovasc. Imaging 3 (10), 1010–1019.

Ukwatta, E., Yuan, J., Qiu, W., Wu, K.C., Trayanova, N., Vadakkumpadan, F., 2014. Myocardial infarct segmentation and reconstruction from 2D late-gadolinium enhanced magnetic resonance images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 554–561.

Watanabe, S., 2013. A widely applicable Bayesian information criterion. J. Mach. Learn. Res. 14, 867–893.

Xu, C., Howey, J., Ohorodnyk, P., Roth, M., Zhang, H., Li, S., 2020. Segmentation and quantification of infarction without contrast agents via spatiotemporal generative adversarial learning. Med. Image Anal. 59, 101568.

Yang, Y., Gao, H., Berry, C., Carrick, D., Radjenovic, A., Husmeier, D., 2022a. Classification of myocardial blood flow based on dynamic contrast-enhanced magnetic resonance imaging using hierarchical Bayesian models. J. R. Stat. Soc. Ser. C. Appl. Stat. 71 (5), 1085–1115.

Yang, Y., Gao, H., Berry, C., Radjenovic, A., Husmeier, D., 2019. Quantification of myocardial perfusion lesions using spatially variant finite mixture modelling of DCE-MRI. In: Proceedings of the International Conference on Statistics: Theory and Applications. ICSTA.

Yang, Y., Gao, H., Berry, C., Radjenovic, A., Husmeier, D., 2022b. Myocardial perfusion classification using a Markov random field constrained Gaussian mixture model. In: Proceedings of the International Conference on Statistics: Theory and Applications. ICSTA.

Zhang, N., Yang, G., Gao, Z., Xu, C., Zhang, Y., Shi, R., Keegan, J., Xu, L., Zhang, H., Fan, Z., et al., 2019. Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine MRI. Radiology 291 (3), 606–617.

Zierler, K.L., 1965. Equations for measuring blood flow by external monitoring of radioisotopes. Circ. Res. 16 (4), 309–321.