Template-Based Multimodal Joint Generative Model of Brain Data

M. Jorge Cardoso[1,2(✉)], Carole H. Sudre[1,2], Marc Modat[1,2], and Sebastien Ourselin[1,2]

[1] Centre for Medical Image Computing (CMIC), University College London, London, UK
m.jorge.cardoso@ucl.ac.uk
[2] Dementia Research Centre (DRC), University College London, London, UK

Abstract. The advent of large of multi-modal imaging databases opens up the opportunity to learn how local intensity patterns covariate between multiple modalities. These models can then be used to describe expected intensities in an unseen image modalities given one or multiple observations, or to detect deviations (e.g. pathology) from the expected intensity patterns. In this work, we propose a template-based multi-modal generative mixture-model of imaging data and apply it to the problems of inlier/outlier pattern classification and image synthesis. Results on synthetic and patient data demonstrate that the proposed method is able to synthesise unseen data and accurately localise pathological regions, even in the presence of large abnormalities. It also demonstrates that the proposed model can provide accurate and uncertainty-aware intensity estimates of expected imaging patterns.

1 Introduction

Neuroimaging studies have become increasingly multimodal, as different imaging techniques (e.g. CT, T1- and T2-weighted MRI, FLAIR, DWI, etc.) contain complementary information about the underlying anatomy, its microstructure and/or its function. However, different modalities also share a large amount of information. As an example, it is possible to predict how a T2 MRI image of a subject should look like given the subject's T1 MRI and a model of image formation [1]. This process, known in the medical image community as image/modality synthesis, has been recently exploited for the purpose of improving multimodal image registration and tissue segmentation [1], synthesising DTI-FA images [2] and PET attenuation-map reconstruction [3], using techniques such as nearest-neighbour patch propagation [1], iterative region-restricted patch-search [2], local similarity-weighted voting [3] and sparse patch-match [4].

One limitation of current synthesis strategies is the lack of knowledge about the uncertainty of the generative process. By generating only one single image estimate, one is assuming that the process of image synthesis is deterministic, or at least that it's uncertainty is non-spatially variant. This is an erroneous assumption that can overestimate the confidence in the synthesis

S. Ourselin et al. (Eds.): IPMI 2015, LNCS 9123, pp. 17–29, 2015.
DOI: 10.1007/978-3-319-19992-4_2

process, possibly propagating large errors to other steps of the image processing pipeline. These uncertainty estimates are especially important in regions where there is no one-to-one intensity correspondence between medical images, e.g. skull/sinus/brain/eye region in CT synthesis from MRI data [3]. In these specific regions, the sinus and skull have highly distinguishable intensities, making the CT probability density function multi-modal. The process of sample averaging seen in [1–4], equivalent to estimating the expectation of the probability density function, is not a good approximation of the MAP estimate, thus producing unrealistic and improbable estimates of the target intensities.

Another limitation of image synthesis methodologies is that one can only reproduce a target modality from a source modality if they share common information or if the missing information can be obtained *a priori* from a population. For example, if a lesion visible in a T2-FLAIR MRI is not visible in a T1 MRI, then this lesion will not appear if one tries to synthesise a T2-FLAIR from a T1 MRI. This limitation has been exploited in the context of pathology detection by synthesising a T2-FLAIR MRI from an observed T1-weighted MRI [2,4], as the pathology is hardly visible in the T1 image. By subtracting the real T2 MRI with a tumour from the T1-derived synthetic T2 MRI, one can obtain a region of interest localising the pathological areas. However, as the synthesis process is considered deterministic (i.e. no estimates of uncertainty), this subtraction process has the same limitations as above - the subtraction process generates edge effects, mis-localising pathological areas when the model does not fit.

From a generative model point-of-view, another strategy to localise intensity outliers in an image is through the process of image clustering [5]. This class of models assume an underlying number of healthy brain tissue classes, normally white-matter (WM), grey-matter and cerebrospinal fluid (CSF), and detects deviations from them. However, any intensity pattern that deviates from these three healthy tissue intensity distributions would be considered as an outlier, even though they commonly appear in healthy brains (e.g. T2 hypo-intensity due to accumulation of cysteine-iron complex in the globus pallidus and substantia nigra). This requires extra empirical pathology-specific post-processing to select a subset of the outliers [5].

With the problem in mind, this work proposes a template-based generative model of brain data and applies it to both the problem image synthesis and inlier/outlier pattern classification for the localisation of image abnormalities. We build on the preliminary idea of [2] but instead of applying a deterministic method, the proposed model describes the full joint probability distribution of a pair of observed images given a set of previously observed templates, making the algorithm robust to edge effects and uncertainty in the model. Furthermore, by describing the process probabilistically, it becomes trivial to further extend the proposed model to an hierarchical segmentation approach.

2 Methods

Image synthesis can be seen as the process of generating an expected image $E[\tilde{y}]$ given an observed image $\tilde{x}$ from a different imaging modality and a database

of N previously observed image pairs $\mathcal{X} = \{y_n, x_n\}$, indexed by n. Note that the technique can be easily extended to more than 2 modalities. The process of collapsing all the available information from $\mathcal{X}$ into one single estimate $E[\tilde{y}]$ ignores both the fact that generating $\tilde{y}$ is an uncertain process an that $p(\tilde{y})$ can be multimodal. Thus, contrary to [2,4], where only the expected image $E[\tilde{y}]$ is generated, we characterise $p(\tilde{y}, \tilde{x}|\mathcal{X}, \theta)$, i.e. the joint probability of observing the image pair $\{\tilde{y}, \tilde{x}\}$ given a set of N previously observed templates $\mathcal{X}$ and the model parameters $\theta = \{\sigma_x^2, \sigma_y^2, \Sigma_{j_n}, \mu_{j_n}, G, w\}$.

From an intuitive point of view, the main difference between the proposed method and previous approaches pertains with the idea that the process of image synthesis is uncertain. Thus, rather than trying to find the expected observation y from an observation x, we estimate what is the probability of observing different values of y when observing x.

2.1 The Observation Model

In order to estimate $p(\tilde{y}, \tilde{x}|\mathcal{X}, \theta)$ one has to first note that, due to the presence of pathology or imaging artefacts, not all intensity pairs $\{\tilde{y}, \tilde{x}\}$ will be likely given a certain set of previously observed images. Thus, it is important to model the possible presence of outliers - intensity pairs that deviate from previous templates. This does not mean that one would be able to predict unseen pathological intensity pairs. Instead, the model assumes that imaging data can fall outside the predictable intensity patterns.

With this aim in mind, the proposed generative model assumes that the observed data is generated from a mixture of $K = 2$ classes labeled by l_i^k, i.e. an inlier class, derived from previous observations, and an outlier class, modelled by a uniform distribution. The probability $p(\tilde{y}_i, \tilde{x}_i, j, l_i|, \mathcal{X}, \theta)$ of observing an intensity pair $\{\tilde{y}_i, \tilde{x}_i\}$, a transformation j and label l at location i, is then defined as

$$p(\tilde{y}_i, \tilde{x}_i, j|l_i, \mathcal{X}, \theta)p(l_i^k) = w \underbrace{p(\tilde{y}_i, \tilde{x}_i, j|l_i^I, \mathcal{X}, \theta)p(l_i^I)}_{\text{inlier model}} + (1-w) \underbrace{p(\tilde{y}_i, \tilde{x}_i|l_i^O)p(l_i^O)}_{\text{outlier model}}$$

In this model, the distribution of the pair $\{\tilde{y}, \tilde{x}\}$ for the outlier class (l^O) is given by an uniform distribution $p(\tilde{y}_i, \tilde{x}_i|l_i^O) = \mathcal{U}$, an inlier class ($l^I$) describing the similarity between the observed pair $\{\tilde{y}_i, \tilde{x}_i\}$ and the previously observed pairs $\mathcal{X} = \{y_n, x_n\}$, a coordinate mapping j_n, a prior distribution $p(l_i^I)$ and $p(l_i^O)$ over the labelling l_i^k, and a global mixing weight w.

In this work, given N previously observed images, the inlier model is defined as a mixture model given by

$$p(\tilde{y}_i, \tilde{x}_i, j|l_i^I, \mathcal{X}, \theta) = \frac{1}{N}\sum_{n=1}^{N} p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n, \mathcal{X}_n, \theta)p(j_n|l_i^I, \theta).$$

i.e., the probability of observing a certain pair of intensities $\{\tilde{y}_i, \tilde{x}_i\}$ is a equally-weighted mixture of the N probabilities of observing the pair given a previously observed pair $\mathcal{X}_n$.

As in the Non-Local STAPLE algorithm [6], we assume that j_n is unknown, or at least uncertain, as the coordinate mapping problem is *ill posed* [7]. As in Simpson *et al.* [7], $p(j_n|l_i^I, \theta)$ is a multivariate Gaussian distribution with parameters $\theta = \{\mu_j^n, \Sigma_j\}$. Here, the expectation of the mapping j_n, i.e. μ_j^n, is estimated using a multimodal pairwise b-spline parameterised registration between the observed image pair and the n-th image pair. A multichannel locally normalised cross correlation is used as an image similarity for registration purposes. In addition, the precision matrix Σ_{ji}^{-1} at location i, represents the inverse of the local directional estimate of registration uncertainty as described in [7]. In this work, Σ_{ji}^{-1} is approximated using the local second-moment matrix, also known as the structure tensor [8], of the observed image $\tilde{x}$ in a cubic convolution region of size $s \times s \times s$ (empirically set to $s = 7$ voxels). This covariance uncertainty approximation assumes that the registration in more uncertain along the edges of the image than across them. Thus, j_n is less likely if it deviates from μ_j^n in a direction orthogonal to the image edges. In future work, this approximation will be replaced by a local covariance estimate as provided by Simpson *et al.* [7].

Similarly to [6], as j_n is unknown, $p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n, \mathcal{X}, \theta)p(j_n|l_i^I, \theta)$ is approximated by its expected value given a multivariate Gaussian distribution on the patch L_2 norm and a multivariate Gaussian distribution over the mapping j_n

$$\begin{aligned} p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n, \mathcal{X}, \theta)p(j_n|l_i^I, \theta) &\approx E[p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n, \mathcal{X}, \theta)p(j_n|l_i^I, \theta)] \\ &= \sum_{j^* \in \mathcal{N}_s} p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n^*, \mathcal{X}, \theta)p(j_n^*|l_i^I, \theta). \end{aligned} \quad (1)$$

Under this approximation and under the assumption of conditional independence between $\tilde{y}$ and $\tilde{x}$, then $p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n^*, \mathcal{X}, \theta)$ is defined as

$$\begin{aligned} p(\tilde{y}_i, \tilde{x}_i|l_i^I, j_n^*, \mathcal{X}, \theta) &= p(\tilde{x}_i|l_i^I, j_n^*, \mathcal{X}, \theta)p(\tilde{y}_i|l_i^I, j_n^*, \mathcal{X}, \theta) \\ &= e^{-\frac{||\mathcal{N}_p(y_i) - \mathcal{N}_p(y_{nj_n^*})||_2^2}{\sigma_{y_{in}}^2} - \frac{||\mathcal{N}_p(x_i) - \mathcal{N}_p(x_{nj_n^*})||_2^2}{\sigma_{x_{in}}^2}} \end{aligned} \quad (2)$$

under the assumption of conditional independence between x and y, and

$$p(j_n^*|l_i^I, \theta) = \frac{1}{Z_{j^*}} e^{-D\Sigma_j^{-1}D^T} \quad (3)$$

In Eq. (1), $\mathcal{N}_s$ is an integration neighbourhood of size $s \times s \times s$ (again with $s = 7$ voxels) of an image similarity component (Eq. 2) and the registration uncertainty distance component (Eq. 3). In Eq. 2, $\mathcal{N}_p(\varkappa)$ is a patch of size $p \times p \times p$ (with $p = 5$ voxels) centred at location $\varkappa$, and in Eq. 3, $D = j_n^* - \mu_n$ is a 1×3 vector characterising the 3-dimensional components of a displacement from μ_n. Both the parameters $\sigma_{y_n}^2$ and $\sigma_{x_n}^2$ denote the sum of a local and a global normally distributed noise model (*iid*) between the observation $\tilde{y}$ and the template y_n, defined as

$$\sigma_{y_{in}}^2 = \sigma(y - y_{n\mu_n})^2 + \sigma(\mathcal{N}_s(y_i) - \mathcal{N}_s(y_{n\mu_n}))^2$$

and equivalently for $\sigma^2_{x_{in}}$, with $\sigma(\varkappa)$ representing the standard deviation of $\varkappa$. Note that the L_2 of the patch can here be used as all images $\mathcal{X}$ have been histogram matched to $\{\tilde{y}_i, \tilde{x}_i\}$ using a 3^{rd} order polynomial fit after non-rigid registration. Finally, Z_{xj^*} is a partition function enforcing

$$\sum_{j^* \in \mathcal{N}_s} p(\tilde{x}_i, j_n^* | l_i^2, x_n, \theta) = 1.$$

The generative model is depicted in Fig. 1. Also of note is the fact that the model extends naturally to more imaging modalities by converting the intensity pairs to K dimensional vectors. In this scenario, the extension of Eq. 2 provides extra complementary information.

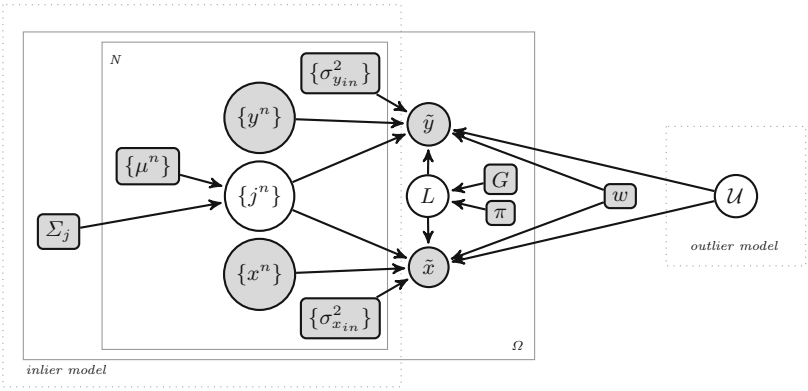

Fig. 1. Generative model of the observed data, with the observed values represented by shaded circles, parameters by shaded rounded-rectangles, and hidden/integrated variables as white circles. Boxes represent replication plates, while the dashed-line boxes annotates the components of the inlier and outlier models.

2.2 Outlier Mixing-Coefficient Prior Estimates

Given the above observation model, one can then estimate the most probable label L, characterising if a certain location belongs to the inlier or outlier class, given by

$$\begin{aligned}
\hat{L} &\approx \underset{L}{\operatorname{argmax}}\, p(l|\tilde{y}, \tilde{x}, j, \mathcal{X}, \theta) \\
&\propto \underset{L}{\operatorname{argmax}}\, p(\tilde{y}, \tilde{x}, j|l, \mathcal{X}, \theta) p(l|\theta) p(\theta) \\
&= \underset{L}{\operatorname{argmax}} \prod_{i \in \Omega} p(\tilde{y}_i, \tilde{x}_i, j|l_i, \mathcal{X}, \theta) p(l_i|\theta) p(\theta),
\end{aligned}$$

where $p(\theta)$ is the prior distribution of the parameters θ, here assumed to be non-informative, and

$$p(l_i^k|\theta) = \pi_k \frac{1}{Z_{MRF}} e^{-\beta U_{MRF}(l_i^k, G)}$$

is the combination of a population prior given the location of the brain and smoothness prior given by a probabilistic extension of a Potts-model-based Markov Random Field (MRF), optimised using a mean field theory approximation, as described in [9]. As we are only interested in outliers within the brain, then

$$\pi_O = 0.5 * (\pi_{GM} + \pi_{WM}) \qquad \pi_I = 1 - \pi_O,$$

i.e. a voxel has a prior probability of 0.5 to be an outlier if it is located within the brain region, defined by non-rigidly registered π_{WM} and π_{GM} ICBM SPM priors (www.fil.ion.ucl.ac.uk/spm/). In the MRF, β=0.5, G is a matrix with the diagonal equal to 0 and the off diagonal equal to 1 and Z_{MRF} is a normalising partition function. The expectation-maximisation algorithm is used to optimise the MRF and the other only free parameter w. The value of w is initialised to 0.9, as a large percentage of the brain should be part of the inlier model, and optimised using the following closed-form update equation

$$w_I^t = \frac{1}{|\Omega|} \sum_i \left(\frac{w_I^{t-1} p(\tilde{y}_i, \tilde{x}_i | l_i^I, \mathcal{X}, \mu_{j_n}, \theta) p(l_i^I)}{\sum_k^K w_k^{t-1} p(\tilde{y}_i, \tilde{x}_i | l_i^k, \mathcal{X}, \mu_{j_n}, \theta) p(l_i^k)} \right)$$

with t representing the current iteration. After optimisation, the voxel-wise estimate of $p(l_i|\tilde{y}_i, \tilde{x}_i, j, \mathcal{X}, \theta)$ provides information about the location of outlier regions as it represents the probability that the voxel in the patch centred at i belong to either the inlier $p(In)_i$ or outlier $p(Out)_i$ classes. This probability can then be used as an estimate of the prior probability that voxel i is an outlier.

2.3 Outlier Segmentation

Given a pair of observed images $\{\tilde{y}, \tilde{x}\}$ and the estimate of $p(l_i|\tilde{y}_i, \tilde{x}_i, j, \mathcal{X}, \theta)$, a voxel-wise segmentation of the outliers of an image can be obtained using a multivariate Gaussian mixture model approach [9] where

$$\pi_i = \{\pi_{WM_i}.p(In)_i, \pi_{GM_i}.p(In)_i, \pi_{CSF_i}.p(In)_i, p(Out)_i\}$$

is a prior distribution for $K = 4$ classes at location i, with π_{WM_i}, π_{GM_i} and π_{CSF_i} representing the healthy population's prior probability to belong to white matter, grey matter and cerebrospinal fluid respectively. Note that any other probabilistic segmentation algorithm can be used at this stage.

2.4 Modality Synthesis

While the proposed generative model enables one to segment and localise out-of-model patterns, a simplified version of the model can also be used for image synthesis. As the process of image synthesis cannot generate more information than the one provided by the observed image $\tilde{x}$ and the priors, we can only look at the inlier part of the described generative model

$$p(\tilde{y}_i, \tilde{x}_i, j_n | l_i^I, \mathcal{X}, \theta).$$

By integrating j_n and assuming that $\tilde{x}_i$ is known but $\tilde{y}_i$ is unknown, one obtains a mixture of mono-modal probability density functions. One can then see this distribution $p(\tilde{y}_i | l_i^I, \mathcal{X}, \theta)$ as a non-parametric probabilistic estimate of $\tilde{y}_i$ using a Gaussian kernel support function, or conversely, a weighted kernel density estimate with the weight given by the $\{\tilde{x}, x_n\}$ patch similarity. Under this model, one can then:

- take random samples from $p(\tilde{y}|\mathcal{X}, \theta)$,
- estimate its expected value $E_{\tilde{y}}[p(\tilde{y}|\mathcal{X}, \theta)]$, which is approximately what is being estimated in [3] but with a different local image similarity,
- calculate the most likely mode of the distribution (i.e. $\hat{y} = \text{argmax}_{\tilde{y}}\ p(\tilde{y}|\mathcal{X}, \theta)$), which provides the most likely intensity given the obtained distribution,
- or make use of the full distribution, providing estimates of the uncertainty of the synthesis process,

all under the assumption of voxel-wise independence.

3 Experiments and Results

3.1 Quantitative Assessment of Image Synthesis in MRI-CT Data

Attenuation correction is an essential requirement for quantification of Positron Emission Tomography (PET) data. In PET/CT acquisition systems, attenuation maps are derived from Computed Tomography (CT) images. However, in hybrid PET/MR scanners, Magnetic Resonance Imaging (MRI) images do not directly provide a patient-specific attenuation map. Current state-of-the-art techniques generate a synthetic CT image from MRI data, either by mapping a tissue segmentation to specific attenuation values, or by directly synthesising the CT image from T1 MRI data [3]. Both these techniques assume that the process of synthesis is deterministic. However, due to the lack of one-to-one mapping between T1 and CT data, some regions can actually have more than one possible cluster of intensities, i.e. the pseudoCT intensity distribution is voxel-wise bimodal. In order to test the advantage of modelling the pseudoCT distribution as a full mixture of Gaussians rather than generating a deterministic intensity, a set of T1-weighted MRI and CT pairs was acquired for 20 elderly subjects with multiple forms of dementia. We compared the state-of-the-art method by Burgos *et al.* [3] to the proposed model. For the proposed model,

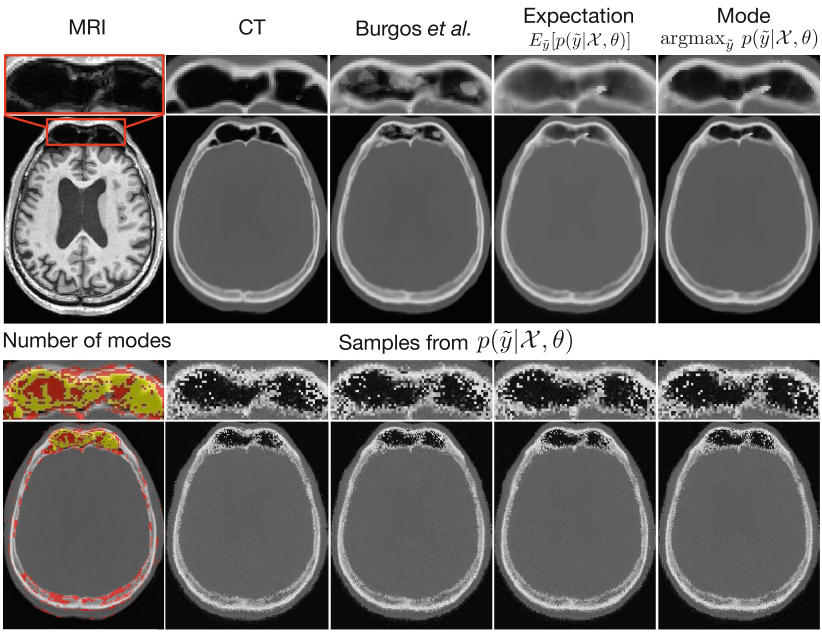

Fig. 2. Top row, from left to right: the original T1-weighted MRI and the patient CT, followed by the output of the method by Burgos et al. and the proposed model's expectation and mode. Bottom row, from left to right: the number of modes of $p(\tilde{y}|\mathcal{X},\theta)$ (1 mode = transparent, 2 modes = red, 3 modes = yellow), followed by four samples of the same distribution. Note the sharp sinus region in the mode image. Also note the number of modes of the distribution increases in uncertain or boundary regions (Colour figure online).

we estimated the expected value $E_{\tilde{y}}[p(\tilde{y}|\mathcal{X},\theta)]$, which should provide a result similar to the averaging process of Burgos *et al.*, and the mode of the distribution (i.e. $\hat{y} = \text{argmax}_{\tilde{y}}\ p(\tilde{y}|\mathcal{X},\theta)$), which should be beneficial in regions where the pseudoCT distribution is bimodal. For the sake of completeness, we also estimate voxel-wisely the number of modes of the estimated distribution, i.e. the number of peaks in the distribution, and also took a few samples from the posterior predictive distribution to demonstrate that the samples vary more in regions with more modes. Results are presented in Fig. 2. Note the sharper intensity in the sinus region, an are which is prone to bimodal intensity distributions. Also note that the four samples of $p(\tilde{y}|\mathcal{X},\theta)$ are less compact in the sinus region. In order to quantitatively assess the accuracy of the proposed methodology, we estimate the mean squared error (MSE) between the real CT and the pseudo CT using [3], and the *per* voxel expectation and the highest mode of $p(\tilde{y}|\mathcal{X},\theta)$. The MSE is defined as $\text{MSE} = \frac{1}{|\Omega|}\sum_i(\text{pseudoCT}_i - \text{realCT}_i)^2$. The mean (std) MSE was found to be 40767.4 (3118.3), 39205.9 (2884.1) and 37216.2 (2179.9) for [3], the expectation and the mode respectively. Due to the non Gaussian nature of the pairwise

differences in MSE, a Wilcoxon Signed-Rank paired test was used for all statistical comparisons. The voxel-wise mode of $p(\tilde{y}|\mathcal{X}, \theta)$ performed statistically significant better ($p < 0.05$) in terms of MSE when compared to [3]. On the other side, no significant different was found between the expectation $E_{\tilde{y}}[p(\tilde{y}|\mathcal{X}, \theta)]$ and [3]. Furthermore, to test the presence of bias in the synthesis process we also estimated the mean error (ME), defined as $\text{ME} = \frac{1}{|\Omega|}\sum_i(\text{pseudoCT}_i - \text{realCT}_i)$. The mean (std) error for the proposed method was 3.2 (25.5) HU and did not significantly differ from zero, supporting the idea that the current synthesis approach is unbiased.

3.2 Quantitative Assessment of the Outlier Segmentation Using Brainweb

The 3 multiple sclerosis (MS) datasets (mild, moderate and severe models) and one normal anatomy dataset were generated using the BrainWeb MR image simulator. Each dataset had both a simulated T1 and T2 MRI images, and for the 3 MS datasets, an associated ground truth probabilistic lesion segmentation. The simulated data was generated using a FLASH sequence with TR = 18 ms, TE = 10 ms, $\alpha = 30°$ for the T1-weighted MRI, and using a spoiled DSE_LATE sequence with TR = 3300 ms, TE = 35, 120 ms, $\alpha = 90°$ for the T2-weighted MRI, both with a 1-mm isotropic voxel size with simulated 3 % noise [10].

As this experiment is solely testing the feasibility of the proposed methodology for the purpose of locating MS lesions under ideal conditions, the normal anatomy T1 and T2 MRIs are used as a single pair $\mathcal{X}$, i.e. N = 1. The method pro-

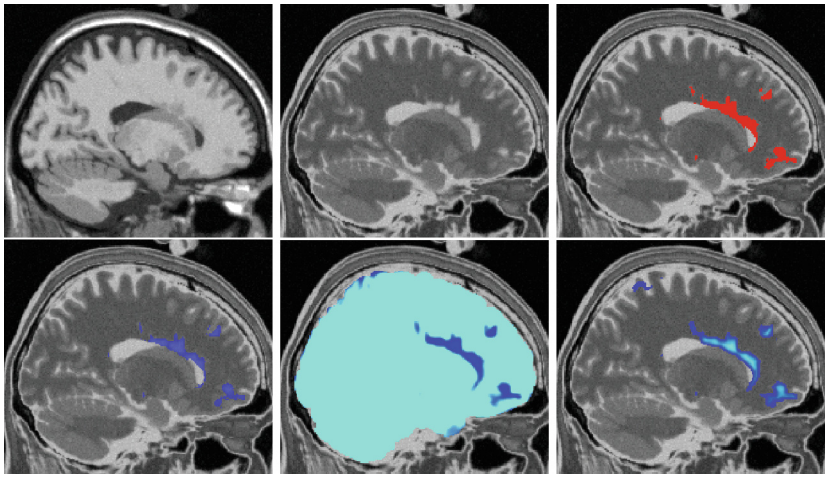

Fig. 3. Brainweb moderate MS model. Left to right) T1 and T2 MRI, groundtruth lesion segmentation and the proposed lesion segmentation, folowed by the inlier observation model $p(\tilde{y}_i, \tilde{x}_i, j_n|l_i^I, \mathcal{X}, \theta)$ and the outlier prior $p(l_i^O|\tilde{y}_i, \tilde{x}_i, j, \mathcal{X}, \theta)$. Note that the outlier prior combines $p(\tilde{y}_i, \tilde{x}_i, j_n|l_i^I, \mathcal{X}, \theta)$ and the label priors.

posed at the end of Sect. 2.2 is here compared to the classical EM-based outlier segmentation method (OSM) [5] with $\kappa = 3$ as recommended. A representative example of both the OSM and proposed methods' results is presented in Fig. 3. OSM obtained a Dice overlap of 3.8, 22.0 and 43.0, for the mild, moderate and severe MS lesion loads respectively, and equivalently, a Dice overlap of 41.8, 51.6 and 65.5 for the proposed method. Note the dramatic increase in accuracy, mainly for the mild MS model. No statistical comparison was performed for this experiment because only three MS models are available in Brainweb.

3.3 Quantitative Assessment of the Outlier Segmentation Using Diabetes Data

This validation aims to determine quantitatively the accuracy of type 2 diabetes white matter lesion (WML) segmentation using the proposed segmentation algorithm and the classical OSM method [5]. For this study, the 20 brain images from the MRBrainS2013 challenge, comprised of both controls and Type 2 diabetes patients (mean age 71 ± 4 years) with WML, were acquired on a 3T Philips scanner with a 3D T1 ($1 \times 1 \times 1$ mm), and fluid attenuated inversion recovery (FLAIR) image ($0.96 \times 0.95 \times 3$ mm) were obtained. Further details about the acquisition and data preprocessing (bias field correction and T1-FLAIR registration) is described in [11] and in the MRBrainS2013 website. Manual WML segmentation was performed on FLAIR images.

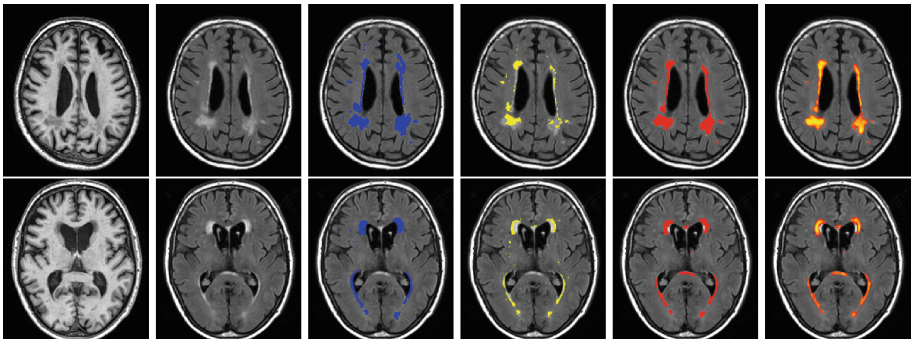

Fig. 4. Subject 4 and 18 of the MRBrainS database. (From left to right) T1 and FLAIR MRI, gold standard lesion segmentation, OMS segmentation, the proposed segmentation and the outlier prior $p(l_i^O|\tilde{y}_i, \tilde{x}_i, j, \mathcal{X}, \theta)$.

With the aim of segmenting only pathological FLAIR hyperintense WML and not the non-pathology-related hyperintense choroid plexus, or the hypointense iron accumulation in the globus pallidus (see the manual segmentation in Fig. 4), the template observations $\mathcal{X}$ should contain some non-pathological intensity outliers but no WML. This was achieved by lesion filling [12] the 20 datasets using the manual WML segmentations, which replaces the WML hyper

intensities with normal WM intensities. A leave-one-out cross validation was then used to segment the lesions, where for each one of the 20 subjects, the remaining $N = 19$ lesion filled T1-FLAIR pairs were used as templates $\mathcal{X}$, avoiding bias due towards a subject's morphology.

Example results are depicted in Fig. 4. Using both the Dice score and lesion volume as accuracy measures, the proposed method obtained a Dice score of 0.45 and a volume R^2 between estimated and the gold standard volume of 0.94, while the OSM [5] method obtained a mean Dice score of 0.38 and a R^2 of 0.55. As the Dice score errors were Gaussian distributed (tested using the one-sample Kolmogorov-Smirnov test on the residuals), a one-tailed T-test was chosen to assess the presence of statistical significant differences in the Dice score. Under this test statistic, the proposed method achieved statistically significantly higher ($p < 10^{-4}$) Dice overlap when compared to the OSM technique.

3.4 Proof-of-Concept Localisation of Anatomically Abnormal Regions in Anatomical Oncology Data

In order to test the feasibility of the proposed algorithm to detect pathologies with extreme presentation in anatomical imaging and to assess the robustness of the algorithm to variations in image quality, structural contrast, and the amount of outliers in the histogram matching approach, the proposed algorithm was applied to the localisation of high grade tumours from the BRATS 2013 database. The 20 lesion filled T1 and FLAIR images from Sect. 3.3 were used as the templates of non-pathological datasets $\mathcal{X}$. As a proof-of-concept, results for the first two subjects (0301 and 0302) of the training database are presented in Fig. 5. Even with large differences in image contrast, the presence of large lesions and large deformations and low contrast in both T1 and FLAIR images,

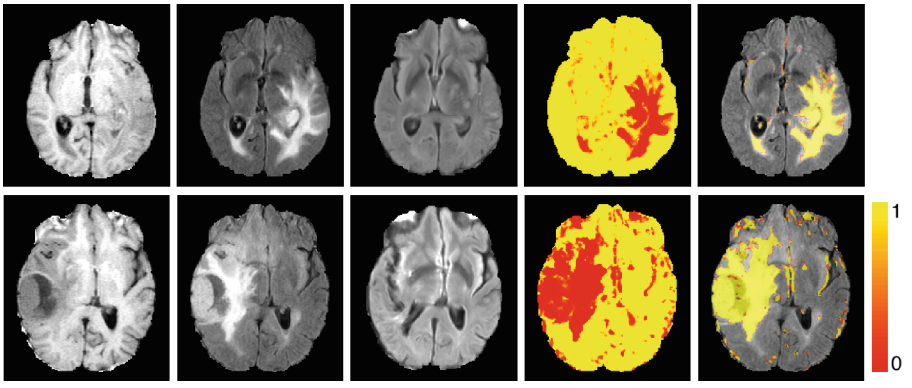

Fig. 5. Tumour localisation using MICCAI BRATS2013 data. Left to right) T1 and FLAIR MRI, synthetic FLAIR using the mode of $p(\tilde{y}_i, \tilde{x}_i, j_n | l_i^I, \mathcal{X}, \theta)$ assuming $\tilde{x}$ is observed, the inlier model $p(\tilde{y}_i, \tilde{x}_i, j_n | l_i^I, \mathcal{X}, \theta)$ and outlier segmentation $p(l_i^O | \tilde{y}_i, \tilde{x}_i, j, \mathcal{X}, \theta)$.

the proposed algorithm was able to localise the non-healthy regions of the image without any pathology specific knowledge about a tumour appearance. Also, for the sake of completeness, a synthetic image was generated using the mode of the distribution. Note that even given a low quality T1 image, the synthetic image is detailed in non-pathological areas.

3.5 Conclusions

This paper proposed a generative model of brain data based from a set of pre-acquired templates. The algorithm enables not only a principled and extensible solution to the problem of modality synthesis, but also a robust way to identify abnormal patterns in medical images. As future work, the proposed framework will be extended to more modalities, allowing it to make use of complementary imaging contrasts to jointly estimate the intensity distributions of multiple modalities. We will also apply the algorithm to the localisation of inhomogeneous pathologies, such as MS and vascular dementia, and other extreme abnormalities, such as traumatic brain injury and cystic lesions. All software will be made available at the time of publication.

Acknowledgments. MJC receives funding from EPSRC (EP/H046410/1). MM and CS are supported by the UCL Leonard Wolfson Experimental Neurology Centre (PR/ylr/18575). SO receives funding from the EPSRC (EP/H046410/1, EP/J020990/1, EP/K005278), the MRC (MR/J01107X/1), the EU-FP7 project VPH-DARE@IT (FP7-ICT-2011-9-601055), the NIHR Biomedical Research Unit (Dementia) at UCL and the National Institute for Health Research University College London Hospitals Biomedical Research Centre (NIHR BRC UCLH/UCL High Impact Initiative - BW.mn.BRC10269).

References

1. Iglesias, J.E., Konukoglu, E., Zikic, D., Glocker, B., Van Leemput, K., Fischl, B.: Is synthesizing MRI contrast useful for inter-modality analysis? In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 631–638. Springer, Heidelberg (2013)
2. Ye, D.H., Zikic, D., Glocker, B., Criminisi, A., Konukoglu, E.: Modality propagation: coherent synthesis of subject-specific scans with data-driven regularization. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 606–613. Springer, Heidelberg (2013)
3. Burgos, N., Cardoso, M.J., Modat, M., Pedemonte, S., Dickson, J., Barnes, A., Duncan, J.S., Atkinson, D., Arridge, S.R., Hutton, B.F., Ourselin, S.: Attenuation correction synthesis for hybrid PET-MR scanners. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 147–154. Springer, Heidelberg (2013)
4. Roy, S., Carass, A., Prince, J.: Magnetic resonance image example based contrast synthesis. IEEE Trans. Med. Imaging **32**(12), 2348–2363 (2013)
5. Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by modeloutlier detection. IEEE Trans. Med. Imaging **20**(8), 677–688 (2001)

6. Asman, A.J., Landman, B.A.: Non-local STAPLE: an intensity-driven multi-atlas rater model. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 426–434. Springer, Heidelberg (2012)
7. Simpson, I.J.A., Woolrich, M.W., Cardoso, M.J., Cash, D.M., Modat, M., Schnabel, J.A., Ourselin, S.: A bayesian approach for spatially adaptive regularisation in non-rigid registration. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part II. LNCS, vol. 8150, pp. 10–18. Springer, Heidelberg (2013)
8. Knutsson, H., Westin, C.-F., Andersson, M.: Representing local structure using tensors II. In: Heyden, A., Kahl, F. (eds.) SCIA 2011. LNCS, vol. 6688, pp. 545–556. Springer, Heidelberg (2011)
9. Cardoso, M.J., Melbourne, A., Kendall, G.S., Modat, M., Robertson, N.J., Marlow, N., Ourselin, S.: AdaPT: an adaptive preterm segmentation algorithm for neonatal brain MRI. NeuroImage **65**, 97–108 (2013)
10. Aubert-Broche, B., Griffin, M., Pike, G.B., Evans, A.C., Collins, D.L.: Twenty new digital brain phantoms for creation of validation image data bases. IEEE Trans. Med. Imaging **25**(11), 1410–1416 (2006)
11. Prados, F., Cardoso, M.J., MacManus, D., Wheeler-Kingshott, C.A.M., Ourselin, S.: A modality-agnostic patch-based technique for lesion filling in multiple sclerosis. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 781–788. Springer, Heidelberg (2014)
12. Battaglini, M., Jenkinson, M., De Stefano, N.: Evaluating and reducing the impact of white matter lesions on brain volume measurements. Hum. Brain Mapp. **33**(9), 2062–2071 (2012)

http://www.springer.com/978-3-319-19991-7

Information Processing in Medical Imaging
24th International Conference, IPMI 2015, Sabhal Mor Ostaig, Isle of Skye, UK, June 28 – July 3, 2015, Proceedings
Ourselin, S.; Alexander, D.C.; Westin, C.-F.; Cardoso, M.J. (Eds.)
2015, XIX, 809 p. 304 illus., Softcover
ISBN: 978-3-319-19991-7