# An iterative multi-atlas patch-based approach for cortex segmentation from neonatal MRI

Carlos Tor-Díez, Nicolas Passat, Isabelle Bloch, Sylvain Faisan, Nathalie Bednarek, François Rousseau

# An Iterative Multi-Atlas Patch-Based Approach for Cortex Segmentation from Neonatal MRI

Carlos Tor-Díez[a], Nicolas Passat[b], Isabelle Bloch[c], Sylvain Faisan[d], Nathalie Bednarek[b,e], François Rousseau[a]

[a]*IMT Atlantique, LaTIM U1101 INSERM, UBL, Brest, France*
[b]*Université de Reims Champagne-Ardenne, CReSTIC, Reims, France*
[c]*LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France*
[d]*ICube UMR 7357, Université de Strasbourg, CNRS, FMTS, Strasbourg, France*
[e]*Service de médecine néonatale et réanimation pédiatrique, CHU de Reims, France*

**Abstract**

Brain structure analysis in the newborn is a major health issue. This is especially the case for preterm neonates, in order to obtain predictive information related to the child development. In particular, the cortex is a structure of interest, that can be observed in magnetic resonance imaging (MRI). However, neonatal MRI data present specific properties that make them challenging to process. In this context, multi-atlas approaches constitute an efficient strategy, taking advantage of images processed beforehand. The method proposed in this article relies on such a multi-atlas strategy. More precisely, it uses two paradigms: first, a non-local model based on patches; second, an iterative optimization scheme. Coupling both concepts allows us to consider patches related not only to the image information, but also to the current segmentation. This strategy is compared to other multi-atlas methods proposed in the literature. Experiments on dHCP datasets show that the proposed approach provides robust cortex segmentation results.

*Keywords:* multi-atlas, patch-based segmentation, iterative segmentation, neonatal MRI, cortex

## 1. Introduction

During the last decades, advances in medical imaging have led to their intensive use in clinical routine. This is particularly true for magnetic resonance imaging (MRI), that allows for the acquisition of high-quality images, in a non-ionizing and non-invasive way. At the same time, image processing methods ded-

icated to MRI have increased the informative potential of these data, particularly in the study of brain structures [1].

Most efforts invested in the development of image analysis methods, particularly in segmentation, have focused on brain MRI data of adult subjects. However, MRI data acquired on children and newborns have been the subject of growing interest over the last few years. This interest is motivated in particular by the study of early human brain development mechanisms. Thus, it gave rise to the development of specific fields in brain image analysis, on the one hand *in utero* (fetal data) [2, 3], and on the other hand *ex utero* (neonatal data) [4, 5].

Indeed, standard brain segmentation methods and tools, deeply validated for adult MRI data, may not be adapted to neonatal brain images. In addition to issues related to acquiring MRI data in good conditions (reduced acquisition time, risk of subject motion...), interpretation difficulties are induced by immaturity of some brain structures (e.g. partial myelination of the white matter, leading to contrast changes) and, of course, the reduced size of structures of interest.

Brain morphometry is a key tool to assess early brain development and to compute relevant biomarkers [6] that can be further used for diagnosis (for instance through quantitative evaluation of lesions), prediction of motor and cognitive development, or therapy (e.g. neuroprotection by melatonin). Neonatal morphometry can also be used to implement a personalized rehabilitative protocol to stimulate brain plasticity [7]. In particular, the cortex is a region of interest for many recent works focusing on brain folding [8, 9, 10], cortical connectivity [11] and cortical development [12, 13]. However, the cerebral cortex is a thin surface object and remains difficult to segment in neonatal MRI data.

Various methods were recently proposed for neonatal brain MRI segmentation. From a technical point of view, different approaches were explored: mathematical morphology [14, 15], deformable models [16, 17], classification [18, 19, 20], patch-driven level set approach [21] or probabilistic modelling [22].

Beyond the relative relevance of each of these image processing / machine learning paradigms, another approach, the so-called multi-atlas framework [23], has demonstrated its interest in the context of brain segmentation. Multi-atlas based techniques are usually performed by registering an anatomy textbook [24] (consisting of one or several labelled images) onto the image to be segmented. The labels are then mapped and fused using classification rules to produce a new label map of the input image. The main items of these registration-based label propagation approaches are the accuracy of the non-rigid registration, the fusion rules, the selection of the labelled images and the labelling errors in primary manual segmentation. Multi-atlas methods provide a spatial prior that restricts the result

to a more realistic anatomy, with the assumption that intensities and segmentation labels are locally correlated. Indeed, by using prior knowledge provided by an appropriate dedicated anatomy textbook, multi-atlas segmentation techniques can cope with the lower quality of neonates MRI and the complexity (in terms of geometry and topology) of the structures of interest [25, 26].

The work presented in this article relies on this multi-atlas strategy. Our purpose is to segment the cortex from 3D MRI data in newborns, considering not only prior appearance information (a given intensity is associated with a given semantics) but also prior spatial information. However, the point-wise mapping conventionally used in multi-atlas approaches, together with the uncertainties induced by the registration step between data, tend to make these prior spatial information less robust.

A possible solution to this issue consists in using the notion of image patches, successfully used for non-local image denoising [27]. It allows one to take into account not only the value at a point, but also the profile of the values in its neighbourhood, for comparison purpose with the points or neighbourhoods in the other images. This strategy, which can compensate for errors and approximations in registration, was used in [28, 29, 30, 31] for brain segmentation. Such an algorithm makes use of local similarities between the image to be segmented and the images contained in the anatomy textbook. The principle is similar to a fuzzy block matching approach which avoids the constraint of a strict one-to-one mapping.

In this article, we propose an iterative way of using patch-based approaches to perform multi-atlas segmentation of the cortex in 3D MRI. Indeed, we propose to use jointly: (1) a patch-based non-local model, and (2) an iterative optimization scheme. The coupled use of these two concepts enables to consider patches not only linked to the image (as in the literature) but also linked to an estimation of its current segmentation. These two kinds of information can be seen similar to the use of data-fidelity and regularization terms, respectively. This latter regularization ensures a homogeneous behaviour of the segmented structure with regard to its geometry, and the progressive refinement of the result over iterations.

This article is organized as follows. In Section 2 the main principles of multi-atlas segmentation are recalled. Section 3 provides a survey of the main multi-atlas methods for cortical segmentation. Our method is presented in Section 4. A comparative study of the different methods is presented in Section 5. A discussion in Section 6 concludes the article.

## 2. Multi-atlas segmentation

In contrast to atlas-based segmentation methods, that use a unique probabilistic atlas computed from $n$ images, a *multi-atlas* segmentation method uses $n$ segmentation examples. Each such example is obtained from an image, in order to guide the segmentation process of the given input image.

### *2.1. General approach*

The $n$ example images are generally chosen similar (in terms of modality, resolution, observation...) to the input image. Indeed, multi-atlas approaches are implicitly based on a smoothness assumption that if two input images are close, then so should be the corresponding segmentation outputs. Usually, the current image segmentation is then defined as a linear combination of the segmentation information derived from these $n$ examples.

The set $\mathcal{E}$ of examples is defined as:

$$\mathcal{E} = \{E_i = (I_i, S_i), i = 1 \dots n\} \tag{1}$$

where each example $E_i$ is composed of one image $I_i$ and its associated segmentation $S_i$. In our case, $I_i$ is a (grey-level) 3D neonatal brain MR image. In the associated segmentation $S_i$, each point $y$ has a value in $[0, 1]$ corresponding to its degree of membership to the chosen structure of interest (in our case, the cortex).

Let $I$ be the input image, and $S$ be the segmentation to be computed from $I$, thanks to the examples of $\mathcal{E}$. Then, the segmentation $S$ is not simply defined as a function $f(I)$, but more generally as:

$$S = f(I, E_1, \dots, E_n) \tag{2}$$

Indeed, $S$ depends not only on intrinsic information related to the image $I$, but also on the way the images $I_i$ are themselves segmented for the same purpose.

This function $f$ is often defined as a linear operator. More precisely, the segmentation $S$ of $I$ is estimated, at each voxel $x$, as a linear combination of the information gathered in different points $y$ within all other images. In this context, these points $y$ and the weights $w_i(x, y)$ are defined according to the similarity between $x$ in $I$ and $y$ in $I_i$. The most intuitive way of defining such weights consists in considering them as reversely proportional to the distance between $x$ and $y$. For instance, this distance can be defined with respect to the intensity information at $x$ and $y$ within their respective images, that is:

$$w_i(x, y) = g(\|I(x) - I_i(y)\|) \tag{3}$$

where $g$ is a function of the distance between $I(x)$ and $I_i(y)$.

The weights $w_i(x, y)$ are computed from intensity-based similarity, but are used for segmentation purposes. Under the isometry assumption that weights computed in the image space are valid in the segmentation space, the segmentation $S$ of $I$ at $x$ can be defined as a (normalized) linear combination $f$ of the segmentations $S_i$ obtained from the $n$ examples in $\mathcal{E}$:

$$\begin{aligned} S(x) &= f(I, E_1, \ldots, E_n)(x) \\ &= \sum_{i=1}^{n} \sum_{y \in \Omega_i} w_i(x, y) S_i(y) \end{aligned} \tag{4}$$

where $\Omega_i$ is the support of the images $I_i$ and $S_i$.

However, if the weights are computed without considering the spatial context, the output segmentation result may not be robust, in particular to noise and registration errors. Thus, instead of considering the voxel as comparison unit, some multi-atlas segmentation methods rely on the notion of image patch, that is a region within the neighbourhood of each point of interest $x$. The similarity related to $x$ in the images $I_i$ can then be considered with respect to a richer context.

### *2.2. Patch-based approaches*

For multi-atlas segmentation purposes, using patches as comparison unit requires to define a distance between $x$ in $I$ and $y$ in $I_i$. This distance depends on $I(x)$ and $I_i(y)$, but also on the set of values of $I$ and $I_i$ in the neighbourhoods $P(x)$ and $P(y)$, respectively.

Patches are often defined as isotropic volumes $P(.)$ (squares or cubes of size $2k + 1$) centred on the points of interest. In this framework, the weights $w_i(x, y)$ (Equation (3)) depend on the information carried by the sub-images restricted to the supports of patches:

$$w_i(x, y) = g\left(\|P_I(x) - P_{I_i}(y))\|\right) \tag{5}$$

In other words, the distance considered for the computation of $w_i(x, y)$ involves the set of point-wise distances on the two patches $P_I(x)$ in $I$ and $P_{I_i}(y)$ in $I_i$.

This calculus may be costly, depending on the size and number of patches $P(.)$. In order to reduce this computational cost, it is generally chosen to restrict the space of patches for a given point $x$ of $I$. We only consider the points $y$ located in the neighbourhood $\mathcal{N}(x)$ of $x$ in $I_i$ (of course, a spatial mapping has to

be carried out beforehand between $I$ and the images $I_i$ of $\mathcal{E}$; this can be done for instance via a registration step). Then, the segmentation $S$ of $I$ rewrites as:

$$S(x) = \sum_{i=1}^{n} \sum_{y \in \mathcal{N}(x)} w_i(x, y) S_i(y) \tag{6}$$

In the next section, we will focus on the computation of these weights $w_i(x, y)$.

## 3. Weights computation for multi-atlas segmentation

Multi-atlas methods are composed of three main steps: (1) registration of the learning dataset $\mathcal{E}$ on the input image $I$, (2) segmentation propagation; and (3) segmentation fusion. In the sequel, we focus on the fusion step (3). In particular, by assuming that segmentation fusion is performed based on Equation (6) (with a normalization coefficient used for making the sum of all weights equal to 1, in each $x$), the discussion mainly deals with the way of computing the weights $w_i(x, y)$.

### *3.1. Non-local means*

In the pioneering article [27] on Non-Local Means (NLM), the weights considered are calculated according to a Gaussian noise hypothesis, for denoising purposes. This kind of weights, denoted by $w_{NLM}$, corresponds to a similarity function inversely proportional to the point-wise intensity distance between the patches, following a normal distribution:

$$w_{NLMi}(x, y) = \exp - \frac{\|P_I(x) - P_{I_i}(y))\|^2}{h^2} \tag{7}$$

where $h$ is a regularization constant that can be automatically tuned as $h^2 = 2\sigma^2 \beta p$ ($p$, the size of patches), with usually $\beta = 1$, while $\sigma$ corresponds to the standard deviation of the Gaussian noise in images [32]. By construction, we have $w_{NLMi}(x, y) \in ]0, 1]$. Such weighting functions have been used for multi-atlas segmentation purpose [28, 29], where $\sigma$ is estimated on the input image to segment.

### *3.2. Joint label fusion*

The Joint Label Fusion (JLF) method [30] (which was applied in MICCAI 2012 Grand Challenge on Multi-Atlas Labeling and finished in the first place) also relies on image patches. In contrast to NLM, only one patch is selected within the

search area $\mathcal{N}(x)$, in order to determine the contribution of an image $I_i$ to the segmentation of $I$ at $x$. In other words, Equation (6) rewrites as:

$$S(x) = \sum_{i=1}^{n} \hat{w}_i(x) S_i(\hat{y}_i) \tag{8}$$

where, for each image $I_i$, $\hat{y}_i$ is the unique point chosen in $\mathcal{N}(x)$ with respect to the similarity between patches $P_I(x)$ and $P_{I_i}(\hat{y}_i)$. Then, we have:

$$\hat{y}_i = \underset{y \in \mathcal{N}(x)}{\arg\min} \, \|P_I(x) - P_{I_i}(y)\| \tag{9}$$

Only one weight $w_i$ is then to be computed for each image $I_i$. Nevertheless, it is defined as spatially variant. In particular, at $x$, the set $\hat{\mathbf{w}}(x) = \{\hat{w}_i(x), i = 1 \dots n\}$ is defined as the following minimizer:

$$\hat{\mathbf{w}}(x) = \underset{\mathbf{w}(x) \in W}{\arg\min} \, \mathbf{w}(x)^t M_x \mathbf{w}(x) \tag{10}$$

where $W$ is the set of vectors $(w_i)_{i=1}^n \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$ and $M_x$ is the correlation matrix between the segmentation error probabilities induced by the $n$ images $I_i$ of $\mathcal{E}$ (see [30, Equations (6–12)] for more details).

### *3.3. STAPLE*

Considering a collection of segmentation maps, the segmentation fusion step aims at computing a probabilistic estimate of the true segmentation. It appears then that the STAPLE method proposed in [33] can be used for multi-atlas segmentation. The goal of STAPLE is to estimate both the segmentation $S$, and performance parameters $\theta$ describing the agreement over the whole image between the experts (i.e. the set of registered images of $\mathcal{E}$) and $S$. An iterative optimization approach (i.e. the Expectation-Maximization algorithm) is used to estimate $S$ and $\theta$ in a probabilistic framework. $S$ is obtained by maximizing its posterior probability:

$$P(S = s | D, \theta^{(k)}) = \prod_{x \in \Omega} P(S(x) = s \mid D, \theta^{(k)}) \tag{11}$$

$$= \prod_{x \in \Omega} \frac{P(S(x) = s) \prod_{i=1}^{n} \theta^{(k)}_{id_{xi}s}}{\sum_{s'} P(S(x) = s') \prod_{i=1}^{n} \theta^{(k)}_{id_{xi}s'}} \tag{12}$$

where $d_{xi}$ is the segmentation decision of expert $i$ for the voxel $x$ ($x \in \Omega$, where $\Omega$ is the support of the image $I$ to segment) and $\theta^{(k)}_{id_{xi}s}$ is the probability that expert $i$ gives the label $s'$ to a voxel $x$ when the reference standard label is $s$ ($\theta_{id_{xi}s} = P(d_{xi} = s' \mid S = s)$) at the iteration $k$. The variable $\theta$ corresponds then to the weights $w$ used for the segmentation fusion. In [34], a non-local version of STAPLE is proposed to add search and patch neighborhoods to compute the weights for segmentation fusion.

## 4. Optimization-based iterative proposed approach

In this section, we describe an optimization-based iterative algorithm based on two key principles. On the one hand, the proposed optimization scheme is iterative in the sense that the computed weights $w_i(x, y)$ can evolve in order to converge progressively towards a satisfactory solution. For each voxel $x$, the set of weights, as a latent representation, is not fixed: both values of weights and the set of neighbours may change during the segmentation process. On the other hand, the optimization-based weight estimation does not longer rely only on data-fidelity, but also on the current segmentation. In other words, the evolution of the weights $w_i(x, y)$ are guided by two kinds of information: proximity between patches in the image space, and proximity in the segmentation space.

### *4.1. Optimization-based approach for weights computation*

Multi-atlas segmentation methods ([28, 29, 30]) implicitly rely on the isometry assumption, *i.e.* the latent representation (provided by the set of weights) of the current voxel $x$ with respect to its neighbours in the learning dataset $\mathcal{E}$ is the same in the intensity space and in the segmentation space. Thanks to this assumption, the segmentation $S$ of $I$ at $x$ is defined as a (normalized) linear combination of the segmentations, the weights being estimated from the intensity images (Eq. 4, 6). Bearing this assumption in mind, the computation of weights $w_i(x, y)$ can also be seen as an optimization problem expressed in the image space. Indeed, one can set the energy function:

$$\phi_I(x) = \left\| P_I(x) - \sum_{i=1}^{n} \sum_{y \in \mathcal{N}(x)} w_i(x, y) P_{I_i}(y) \right\| \tag{13}$$

that defines the distance between the patch $P_I(x)$ and the linear combination of the patches $P_{I_i}(y)$ in $I_i$ for all the $y$ of $\mathcal{N}(x)$. The weights $\hat{\mathbf{w}}(x) = \{w_i(x, y) \mid$

$y \in \mathcal{N}(x), i = 1 \dots n\}$ can then be computed by minimizing $\phi_I(x)$ as follows:

$$\hat{\mathbf{w}}(x) = \underset{\mathbf{w} \in W}{\arg\min}\, \phi_I(x) \tag{14}$$

where $W$ is the set of normalized vectors taking their values in $[0, 1]$. The weights are then estimated such that the reconstruction error of the current image patch $P_I$ using a linear model is minimized. This approach is inspired from the LLE (Locally Linear Embedding) algorithm [35], initially designed for dimension reduction purposes.

Since the computed weights are then used for segmentation purposes based on Eq. 4 or 6, the above strategy relies also on the isometry assumption. The isometry assumption can be relaxed by keeping the way the segmentation is performed from the weights (Eq. 4, 6). The principle is to directly compute the weights in the space of segmentations. To this end, it is then necessary to replace, in the definition of $\phi_I(x)$ (Equation (13)), the patches $P_I$ and $P_{I_i}$ on images $I$ and $I_i$ by the patches $P_S$ and $P_{S_i}$ on the segmentations $S$ and $S_i$ associated with these images. The same optimization scheme (Equation (14)) can then be used for determining the weights $\hat{\mathbf{w}}(x) = \{w_i(x, y) \mid y \in \mathcal{N}(x), i = 1 \dots n\}$. However, this formulation requires to know beforehand the segmentation $S$ of the target image $I$. Since $S$ is the segmentation to estimate, we propose to derive an iterative scheme to relax the isometry assumption.

### *4.2. Mixed patches*

To this end, we propose to define mixed patches $P_{E_\star}$ from image–segmentation couples $E_\star = (I_\star, S_\star)$. All the considered images $I_\star$ (namely, $I$ and the $I_i$) are defined on a same support $\Omega$ (indeed, the image registration step is performed beforehand), and they take their values within an interval $V \subset \mathbb{R}$. The associated segmentations $S_\star$ are defined on the same support $\Omega$, but they take their values in $[0, 1]$. Without loss of generality, we consider that $V$ has been normalized. Under these assumptions, both the $I_\star$ and the $S_\star$ can be expressed as functions $\Omega \to [0, 1]$ (with, however, distinct semantics). This normalization is crucial for defining a non-biased inter-patch distance.

Practically, for an image–segmentation couple $E_\star = (I_\star, S_\star)$, the space of associated patches $E_\star$ is a function:

$$\left| \begin{array}{rcccl} P_{E_\star} & : & \Omega & \to & [0,1]^p \times [0,1]^p \\ & & x & \mapsto & (P_{I_\star}(x), P_{S_\star}(x)) \end{array} \right. \tag{15}$$

where $P_{I_\star}(x)$ and $P_{S_\star}(x)$ are the usual patches of image and segmentation, respectively, that are vectors of $[0,1]^p$, that indicate the values of $I_\star$ and $S_\star$ inside a window $\mathcal{N}(x)$ of $\Omega$ of size $p$, locally centred on $x$. (For the sake of concision, we will equivalently consider $P_{E_\star}(x)$ as a vector $(p^k_{E_\star}(x))_{k=1}^{2p}$ of $[0,1]^{2p}$ instead of a couple of vectors of size $p$.)

### *4.3. Inter-patch distance and energy function*

In order to define the inter-patch distance, the $L_k$ norms (and especially $L_1$ and $L_2$) can be considered:

$$\|P_{E_\alpha}(x) - P_{E_\beta}(x)\|_1 = \sum_{k=1}^{2p} \left|p^k_{E_\alpha}(x) - p^k_{E_\beta}(x)\right| \tag{16}$$

$$\|P_{E_\alpha}(x) - P_{E_\beta}(x)\|_2 = \Big(\sum_{k=1}^{2p} \left(p^k_{E_\alpha}(x) - p^k_{E_\beta}(x)\right)^2\Big)^{\frac{1}{2}} \tag{17}$$

In particular, it is possible to define, in a way similar to Equation (13), an energy function $\phi_E$ which, given a image–segmentation couple $E = (I, S)$ and a multi-atlas set $\mathcal{E} = \{E_i = (I_i, S_i), i = 1 \dots n\}$, expresses the distance between $E$ and a linear combination on $\mathcal{E}$ at a point $x$ of $\Omega$:

$$\phi_E(x) = \Big\|P_E(x) - \sum_{i=1}^{n} \sum_{y \in \mathcal{N}(x)} w_i(x,y) P_{E_i}(y)\Big\| \tag{18}$$

As in Equation (14), our purpose is to minimize the function $\phi_E(x)$ in order to obtain a vector of optimal weights for computing the segmentation $S$ of $I$ at every voxel $x$:

$$\hat{\mathbf{w}}(x) = \arg\min_{\mathbf{w} \in W} \phi_E(x) \tag{19}$$

with the same notations as in Section 4.1. This criterion can be optimized efficiently in the same way than the one of Eq. 14.

### *4.4. Iterative Multi-Atlas Patch-based approach*

In order to relax the isometry assumption, the computation of the weights $w_i(x,y)$ is based on proximity between patches in the image space, and proximity in the segmentation space. Since the segmentation of the target image is not known, an iterative method is used (Algorithm 1).

Its successive steps are discussed hereafter.

**Algorithm 1** Iterative Multi-Atlas Patch-Based approach

**Require:** $I$: input image ; $\mathcal{E} = \{E_i = (I_i, S_i), i = 1 \dots n\}$: learning dataset, N: number of iterations, $\{\alpha_0 = 0, \alpha_1, \dots, \alpha_{N-1}\}$: intensity *vs* segmentation trade-off ; $S^{(0)}$: arbitrarily chosen initial segmentation

**Ensure:** $S$: segmentation of $I$

```
for j = 1 ... N do
    for all x ∈ Ω do
        Compute ŵ^(j)(x) (Equation (20))
        Compute S^(j)(x) from ŵ^(j)(x) (Equation. (21))
    end for
end for
S ← S^(N)
```

#### *4.4.1. Initialization*

The first iteration of the process requires to define beforehand an initial image–segmentation couple $E^{(0)} = (I, S^{(0)})$ associated to $I$. The initial segmentation $S^{(0)}$ can be chosen arbitrarily. This initialization is justified hereafter, in the weighting policy description.

#### *4.4.2. Iterations*

During the iterative process, the current energy function $\phi_E^{(j+1)}(x)$ is optimized by considering $P_E^{(j)}(x) = (P_I(x), P_{S^{(j)}}(x))$, in order to define the weights $w_i^{(j+1)}(x, y)$ (Equations (18–19)):

$$\hat{\mathbf{w}}^{(j+1)}(x) = \arg\min_{\mathbf{w} \in W} \left\| P_E^{(j)}(x) - \sum_{i=1}^{n} \sum_{y \in \mathcal{N}(x)} w_i(x, y) P_{E_i}(y) \right\| \tag{20}$$

and then to compute $S^{(j+1)}$ from these weights $w_i^{(j+1)}(x, y)$ of $\hat{\mathbf{w}}^{(j+1)}(x)$:

$$S^{(j+1)}(x) = \sum_{i=1}^{n} \sum_{y \in \mathcal{N}(x)} w_i^{(j+1)}(x, y) S_i(y) \tag{21}$$

#### *4.4.3. Weighting*

We propose to balance the influence of the terms related to intensity and estimated segmentation, respectively. Here, the underlying idea is to relax progressively the hypothesis of isometry between intensity space and segmentation space.

In particular, at each iteration $j$, we consider a parameter $\alpha_j \in [0,1]$ such that the terms of $P_{E_\star}$ linked to the patch $P_{S_\star}$ (resp. $P_{I_\star}$) are weighted by $\alpha_j$ (resp. $1-\alpha_j$).

Practically, this weighting can be performed without altering the formulation of Equations (19–21). Indeed, it is sufficient that the $\alpha_j$ weights be involved in the definition of $P_{E_\star}$ by applying a scale function onto the value space. In other words, we define a function:

$$\left| \begin{array}{rcccl} \Pi_\alpha & : & [0,1]^p \times [0,1]^p & \to & \mathbb{R}^p \times \mathbb{R}^p \\ & & (P_{I_\star}, P_{S_\star}) & \mapsto & ((1-\alpha)P_{I_\star}, \alpha P_{S_\star}) \end{array} \right. \tag{22}$$

It is then sufficient to substitute $\Pi_{\alpha_j} \circ P_{E_\star}$ to $P_{E_\star}$ in the above optimization scheme. In order to guarantee a continuous evolution of the process, we choose an increasing sequence of weights $(\alpha_j)$ depending on $j$. At the first iteration, $\alpha_0$ is set to $0$. Then any initialization of the segmentation $S^{(0)}$ can be considered.

## 5. Experiments

In this section, we apply the proposed method for processing 3D neonatal brain MR images. In this work, we focus on the segmentation of cerebral cortex, by taking advantage of a base of similar images, endowed with associated segmentation maps.

### *5.1. Data*

The considered images are T2-weighted MRI data, made available [36] in the context of the Developing Human Connectome Project (dHCP)[1] [37]. Infants were recruited and imaged at the Evelina Neonatal Imaging Centre, London. Informed parental consent was obtained for imaging and data release, and the study was approved by the UK Health Research Authority. All infants were born and imaged at term age (37-44 weeks of age). Imaging was carried out on 3T Philips Achieva using a dedicated neonatal imaging system which included a neonatal 32 channel phased array head coil. The infants were imaged in natural sleep. T2w images were acquired in sagittal and axial slice stacks with in-plane resolution $0.8 \times 0.8$ mm$^2$ and $1.6$ mm slices overlapped by $0.8$ mm. Other parameters were: $12000/156$ms TR/TE, SENSE factor $2.11$ (axial) and $2.58$ (sagittal). In this work, a set of 40 images with a voxel size of $0.5 \times 0.5 \times 0.5$ mm$^3$ is considered. Among these data, 30 are used for building the learning dataset $\mathcal{E}$, while the other 10

[1] http://www.developingconnectome.org

are used as testing dataset. For each of these 40 images, the algorithmic pipeline dHCP[2] has been applied. The obtained segmentations, visually validated, are used as reference.

### *5.2. Data preprocessing*

The example images were registered in two steps (affine, then non-rigid registration) by using ANTs[3] [38]. To improve the robustness with respect to intensity variability, the intensities of example images ($I_i$) were normalized by histogram matching with the subjects to be segmented ($I$). This is done in order to avoid additive bias effects on inter-patch distances (Equations (16–17)).

### *5.3. Compared methods*

The proposed method, described in Section 4, and denoted by IMAPA (IMAPA stands for "Iterative Multi-Atlas Patch-based Approach") in the sequel, was compared to the following methods:

- Non-Local Means – NLM (Section 3.1);
- Joint Label Fusion – JLF [30] (Section 3.2);
- MANTIS [18].

It is worth mentioning that the first two are patch-based multi-atlas methods, such as IMAPA. By contrast, MANTIS is not multi-atlas; it is based on classification techniques. Indeed, this latter method relies on an adaptation of the unified segmentation algorithm [39], enriched with morphological processing and topological filtering steps.

### *5.4. Parameters*

The MANTIS method presents few parameters. These are regularization parameters for the initial classification of the unified segmentation [39]. The used values are those defined by default.

The three other methods based on patches all present similar characteristics. In particular, they share various parameters: the number $K$ of nearest patches considered for the calculus of energy functions $\phi(.)$ (Equation (13) and next); the size $p$ of the patches $P(.)$; and the size of research areas $\mathcal{N}(.)$.

[2] https://github.com/DevelopingHCP/structural-pipeline
[3] http://stnava.github.io/ANTs

The use of the $K$ nearest patches (among $n.|\mathcal{N}(x)|$) for the calculus of $\phi(x)$ is aimed to reduce the influence of an excessive amount of patches with a low similarity with the considered patch in each point $x \in \Omega$. The parameter $K$ is set to $15$, except for JLF where, by definition, we have $K = n$ (here, $30$). The value of $p$ is set to $27$; it corresponds to patches of size $3 \times 3 \times 3$ voxels. The search areas $\mathcal{N}(.)$ are cubes of $7 \times 7 \times 7$ voxels. These parameters were set based on preliminary experiments.

Some parameters are specific to some of these methods. In JLF, the supplementary parameters were set at a fixed default value. For the IMAPA method, the inter-patch distance relies on the $L_2$ norm (Equation (17)). The number of iterations $N$ was empirically set to $2$. The values of $\alpha_j$ (trade-off coefficient) are set to $0$ and $0.25$ for the first and second iteration, respectively. The $0$ value allows us at perform a first iteration based on data-fidelity only. The $0.25$ value for the second iteration was determined from preliminary experiments, within a range of values from $10^{-2}$ to $0.5$. The convergence of the algorithm has been assessed by computing the fuzzy DICE index for each value of $\alpha \in \{0; 0.25; 0.5; 0.75; 1\}$. Although the fuzzy DICE index increases over iterations, the first two iterations (from $\alpha = 0$ to $0.25$) led to the main improvement. Figure 1 shows estimated segmentation maps through the iterative process for one subject. It can be seen that small isolated regions are removed over the iterations. To shorten the computation time, we have used in this work only two iterations (i.e. $\alpha \in \{0; 0.25\}$).

The $K$ nearest neighbours are updated at each iteration. This allows us to combine examples that are increasingly adapted to estimate the segmentation, according to the evolution of $\alpha_j$.

(a) $\alpha = 0$ (b) $\alpha = 0.25$ (c) $\alpha = 0.5$ (d) $\alpha = 0.75$ (e) $\alpha = 1$

Figure 1: Slices of estimation segmentation maps of the same subject through the iterative process (i.e. $\alpha \in \{0; 0.25; 0.5; 0.75; 1\}$.

### *5.5. Results*

Quantitative evaluations of the results provided by the four methods are gathered in Table 1. In particular, two evaluation measures were considered: the Dice index and the Peak Signal to Noise Ratio (PSNR). The Dice index (that varies between 0 and 1) is an overlapping measure between the estimated segmentation and the ground-truth; the closer to 1, the better the adequacy between the result and the ground-truth both in terms of false positives and false negatives. The PSNR provides another complimentary measure of quality (mean quadratic error) of the estimated image; the higher its value (in dB), the better the quality of the obtained result. The PSNR is computed as follows:

$$PSNR = 10 \times \log_{10} \left( \frac{MAX^2}{MSE} \right) \tag{23}$$

where $MAX$ is the maximum value of the reference image and $MSE$ is the mean squared error (between the estimated image and the reference image).

Table 1: Performances of IMAPA, compared to three other segmentation methods (NLM, JLF, MANTIS, IMAPA$^{(0)}$) applied on 10 subjects, with a learning dataset of 30 subjects. Best results are highlighted in bold.

| | NLM | JLF [30] | MANTIS [18] | IMAPA$^{(0)}$ ($\alpha = 0$) | IMAPA |
|---|---|---|---|---|---|
| Dice (mean) | 0.876 | 0.845 | 0.793 | 0.873 | **0.887** |
| Dice (std. dev.) | **0.011** | 0.018 | 0.028 | 0.011 | **0.011** |
| PSNR (mean, dB) | 20.758 | 18.966 | 16.661 | 20.671 | **21.086** |
| PSNR (std. dev., dB) | 0.396 | 0.487 | 0.533 | 0.411 | **0.392** |

The IMAPA and NLM methods generate results defined as fuzzy maps. In order to compare these results with those of JLF and MANTIS (that generate binary maps), a thresholding of the fuzzy maps is set at value $0.5$. IMAPA$^{(0)}$ stands for the non-iterative IMAPA method, *i.e.* only the intensity patches are used for the weight computation ($\alpha = 0$).

A more qualitative assessment of the results obtained by the four methods is also available via illustrative samples of segmentations, visualized on 2D slices, in Figures 2–3 (axial slices), and as 3D visualizations of meshed obtained from binary volumes (sagittal view) in Figure 4 and using meshing tools from BrainVISA software[4] [40].

## 6. Discussion

From a quantitative point of view (Table 1), the results obtained with the proposed IMAPA method are better than those obtained with the other three tested methods, for both Dice and PSNR measures. This improvement is significant compared to the JLF and MANTIS methods; it is lower compared to the NLM method (approximately $0.01$ gap for Dice and $0.3$ dB for PSNR). These results on the dHCP dataset tend to show that the proposed strategy of coupling patch-based and iterative optimization is indeed of interest, as the results are at the level of the state of the art. In addition to this algorithmic approach, the simultaneous use of a data-fidelity term and a regularization term from both the example images and their segmentations in same patches, also seems relevant. Indeed, compared to

[4] http://brainvisa.info

IMAPA, the NLM weight computation method, although being also patch-based, is neither iterative nor image/segmentation mixing.

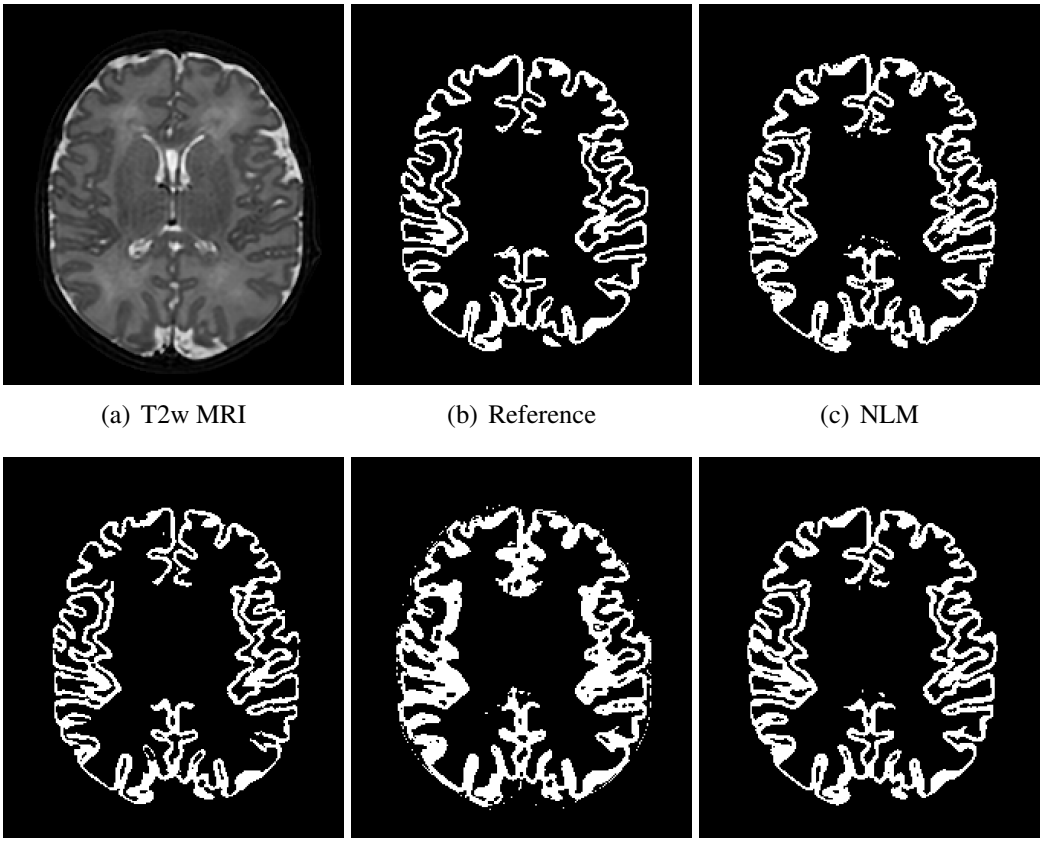
Despite slight quantitative differences between IMAPA and the other methods, observation of the visual results confirms the satisfactory behaviour of the IMAPA method from a more qualitative point of view. First, this can be seen on 2D slices in Figures 2–3. In Figure 2, one can observe that MANTIS clearly over-segments the cerebral cortex, compared to the other three methods, while JLF tends to slightly under-segment it. IMAPA and NLM provide visually comparable results, with slightly more noisy results for NLM. These trends are confirmed by Figure 3, where we can focus on a zoomed area of the slice, and observe the zones of false positives and false negatives. This illustration emphasizes the over- and under-segmentation behaviours of MANTIS and JLF, respectively. Once again, NLM and IMAPA provide close results, but IMAPA seems to present lower false negatives.

Secondly, the behaviour of the four methods can be qualitatively observed from 3D cortical surfaces computed based on segmentation results. Indeed, in Figure 4, one can observe that JLF leads to topologically incorrect surfaces (holes), due to under-segmentation, while the over-segmentation of MANTIS leads to noisy patterns on the surface, and disconnected elements. Once again, NLM and IMAPA provide globally correct surfaces, with a slightly more regular appearance with IMAPA.

Overall, these experiments argue in favour of considering mixed patches and iterative optimization schemes for patch-based segmentation approaches. This is, in particular, strengthened by the fact that these results were obtained in a complicated applicative context, namely the analysis of cortical surface.

However, these results are yet preliminary. At this stage, they cover a small data set of 10 + 30 images. In addition, these data present good contrast quality and signal-to-noise ratio. Consequently, further work will focus on the assessment of the robustness of IMAPA in more challenging contexts, with data of lower quality. In addition, since the method is based on a multi-atlas paradigm, it may be relevant to investigate the impact of example quality, in order to understand the side effects of imperfect segmentation examples, but also heterogeneous data collected in multi-centric studies.

The space of parameters of the IMAPA method has not been fully explored. In particular, the type of distance, the number of closest neighbours, the weighting policies will deserve a wider study. We will also consider larger ranges of iterations, in order to observe if smooth evolutions of the trade-off parameter $\alpha$ could enable to improve the overall quality of the segmentation results.

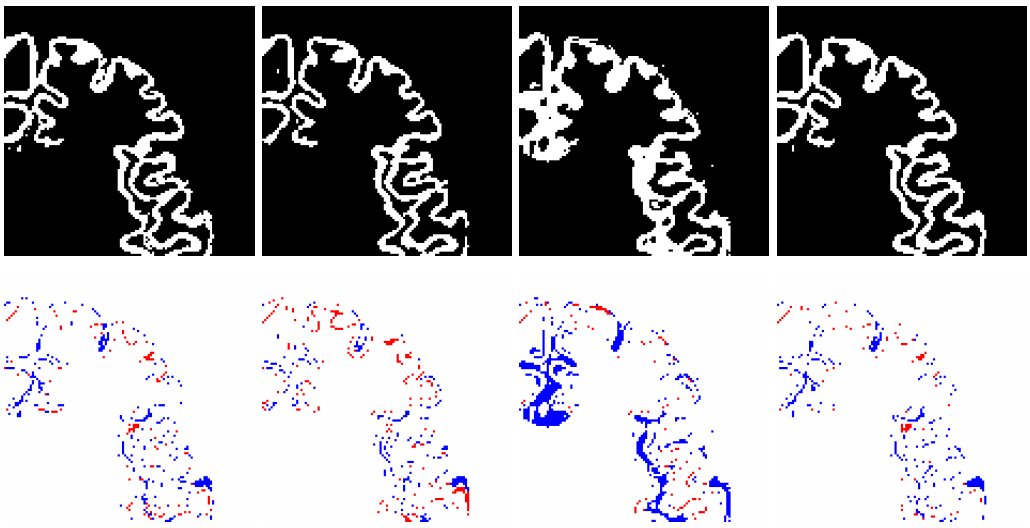


(a) T2w MRI (b) Reference (c) NLM

(d) JLF (e) MANTIS (f) IMAPA

Figure 2: Segmentation results on a T2-weighted brain MRI (axial slice). (a) Input image. (b) Reference segmentation obtained from the dHCP pipeline [17]. Segmentation results obtained with: (c) NLM, (d) JLF [30], (e) MANTIS [18], (f) IMAPA.

It has also to be mentioned that IMAPA shares similarities with sparse coding based methods, such as the one described in [41, 42]. More specifically, the representation of the input patch as a linear combination of a set of atoms (called a dictionary in sparse coding) is close to non-local patch-based linear modeling. Main differences between sparse coding-based methods and the framework used here are: 1) data representation (dictionaries vs raw patches), 2) sparsity (via a penalty term vs nearest neighbors in patch space). The estimation of the weights through a reconstruction term brings sparsity-based and non-local approaches closer. The iterative nature of the proposed method is also related to cascading approaches [43] (also called auto-context [44]) that make use of previous estimations to refine the



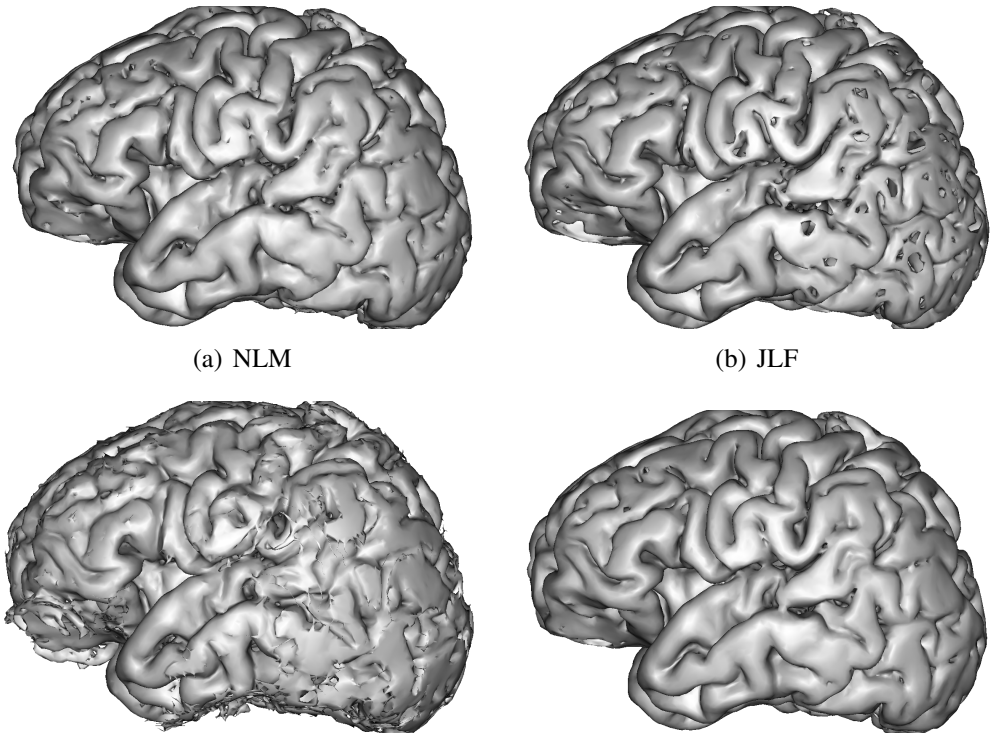
(a) NLM (b) JLF (c) MANTIS (d) IMAPA

Figure 3: First row: zoom on a region of Figure 2. Second row: false positives (in blue) and false negatives (in red) provided by the different methods, compared to the reference image (Figure 2(b)). (a) NLM, (b) JLF [30], (c) MANTIS [18], (d) IMAPA.

segmentation maps. It has to be noticed that the probabilistic framework used in STAPLE is also related to this cascading approach by the alternation between the estimation of weights and the segmentation map during the EM optimization. Future research directions could focus on these methodological aspects to further study the link between IMAPA, sparse coding based methods, cascading frameworks and the STAPLE probabilistic modeling, and to propose a unified framework highlighting the key components of these approaches.


## Acknowledgements

The research leading to these results has been supported by the ANR MAIA project, grant ANR-15-CE23-0009 of the French National Research Agency (project MAIA, "Multiphysics image-based analysis for premature brain development understanding", http://recherche.imt-atlantique.fr/maia), INSERM and Institut Mines Télécom Atlantique (Chaire "Imagerie médicale en thérapie interventionnelle"), Fondation pour la Recherche Médicale (FRM grant DIC20161236453) and the American Memorial Hospital Foundation.



(a) NLM (b) JLF

(c) MANTIS (d) IMAPA

Figure 4: 3D visualization of mesh surfaces computed from the segmentation results obtained with (a) NLM, (b) JLF [30], (c) MANTIS [18], (d) IMAPA.

## References


[1] A. W. Toga, Brain Mapping: An Encyclopedic Reference, Academic Press, 2015.

[2] F. Rousseau, C. Studholme, R. Jardri, M. Thomason, In vivo human fetal brain analysis using MR imaging, in: Fetal Development: Research on Brain and Behavior, Environmental Influences, and Emerging Technologies, Springer International Publishing, pp. 407–427.

[3] O. M. Benkarim, G. Sanroma, V. A. Zimmer, E. Muñoz-Moreno, N. Hahner, E. Eixarch, O. Camara, M. A. González Ballester, G. Piella, Toward the automatic quantification of in utero brain development in 3D structural MRI: A review, Human Brain Mapping 38 (2017) 2772–2787.

[4] C. N. Devi, A. Chandrasekharan, V. K. Sundararaman, Z. C. Alex, Neonatal brain MRI segmentation: A review, Computers in Biology and Medicine 64 (2015) 163–178.

[5] A. Makropoulos, S. J. Counsell, D. Rueckert, A review on automatic fetal and neonatal brain MRI segmentation, NeuroImage 170 (2018) 231–248.

[6] L. R. Ment, D. Hirtz, P. S. Hüppi, Imaging biomarkers of outcome in the developing preterm brain, The Lancet Neurology 8 (2009) 1042–1055.

[7] P. S. Hüppi, Cortical development in the fetus and the newborn: Advanced MR techniques, Topics in Magnetic Resonance Imaging: TMRI 22 (2011) 33–38.

[8] J. Dubois, M. Benders, A. Cachia, F. Lazeyras, R. Ha-Vinh Leuchter, S. V. Sizonenko, C. Borradori-Tolsa, J. F. Mangin, P. S. Hüppi, Mapping the early cortical folding process in the preterm newborn brain, Cerebral Cortex 18 (2008) 1444–1454.

[9] J. Lefèvre, D. Germanaud, J. Dubois, F. Rousseau, I. de Macedo Santos, H. Angleys, J.-F. Mangin, P. S. Hüppi, N. Girard, F. De Guio, Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns?, Cerebral Cortex 26 (2016) 3023–3035.

[10] E. Orasanu, A. Melbourne, M. J. Cardoso, H. Lomabert, G. S. Kendall, N. J. Robertson, N. Marlow, S. Ourselin, Cortical folding of the preterm brain: A longitudinal analysis of extremely preterm born neonates using spectral matching, Brain and Behavior 6 (2016).

[11] G. Ball, J. P. Boardman, P. Aljabar, A. Pandit, T. Arichi, N. Merchant, D. Rueckert, A. D. Edwards, S. J. Counsell, The influence of preterm birth on the developing thalamocortical connectome, Cortex 49 (2013) 1711–1721.

[12] G. Ball, L. Srinivasan, P. Aljabar, S. J. Counsell, G. Durighel, J. V. Hajnal, M. A. Rutherford, A. D. Edwards, Development of cortical microstructure in the preterm human brain, Proceedings of the National Academy of Sciences of the United States of America 110 (2013) 9541–9546.

[13] Q. Yu, A. Ouyang, L. Chalak, T. Jeon, J. Chia, V. Mishra, M. Sivarajan, G. Jackson, N. Rollins, S. Liu, H. Huang, Structural development of Human fetal and preterm brain cortical plate based on population-averaged templates, Cerebral Cortex 26 (2016) 4381–4391.

[14] L. Gui, R. Lisowski, T. Faundez, P. S. Hüppi, F. Lazeyras, M. Kocher, Morphology-driven automatic segmentation of MR images of the neonatal brain, Medical Image Analysis 16 (2012) 1565–1579.

[15] B. Morel, Y. Xu, A. Virzi, T. Géraud, C. Adamsbaum, I. Bloch, A challenging issue: Detection of white matter hyperintensities in neonatal brain MRI, in: EMBC, Procs., 2016, pp. 93–96.

[16] F. Leroy, J.-F. Mangin, F. Rousseau, H. Glasel, L. Hertz-Pannier, J. Dubois, G. Dehaene-Lambertz, Atlas-free surface reconstruction of the cortical grey-white interface in infants, PloS One 6 (2011) e27128.

[17] A. Schuh, A. Makropoulos, R. Wright, E. C. Robinson, N. Tusor, J. Steinweg, E. Hughes, L. C. Grande, A. Price, J. Hutter, J. V. Hajnal, D. Rueckert, A deformable model for the reconstruction of the neonatal cortex, in: ISBI, Procs., 2017, pp. 800–803.

[18] R. J. Beare, J. Chen, C. E. Kelly, D. Alexopoulos, C. D. Smyser, C. E. Rogers, W. Y. Loh, L. G. Matthews, J. L. Y. Cheong, A. J. Spittle, P. J. Anderson, L. W. Doyle, T. E. Inder, M. L. Seal, D. K. Thompson, Neonatal brain tissue classification with morphological adaptation and unified segmentation, Frontiers in Neuroinformatics 10 (2016) 12.

[19] P. Moeskops, M. J. Benders, S. M. Chiţă, K. J. Kersbergen, F. Groenendaal, L. S. de Vries, M. A. Viergever, I. Išgum, Automatic segmentation of MR brain images of preterm infants using supervised classification, NeuroImage 118 (2015) 628–641.

[20] P. Anbeek, I. Išgum, B. J. M. van Kooij, C. P. Mol, K. J. Kersbergen, F. Groenendaal, M. A. Viergever, L. S. d. Vries, M. J. N. L. Benders, Automatic segmentation of eight tissue classes in neonatal brain MRI, PLOS ONE 8 (2013) e81895.

[21] L. Wang, F. Shi, G. Li, Y. Gao, W. Lin, J. H. Gilmore, D. Shen, Segmentation of neonatal brain MR images using patch-driven level sets, NeuroImage 84 (2014) 141–158.

[22] M. J. Cardoso, A. Melbourne, G. S. Kendall, M. Modat, N. J. Robertson, N. Marlow, S. Ourselin, AdaPT: An adaptive preterm segmentation algorithm for neonatal brain MRI, NeuroImage 65 (2013) 97–108.

[23] J. E. Iglesias, M. R. Sabuncu, Multi-atlas segmentation of biomedical images: A survey, Medical Image Analysis 24 (2015) 205–219.

[24] M. I. Miller, G. E. Christensen, Y. Amit, U. Grenander, Mathematical textbook of deformable neuroanatomies, Proceedings of the National Academy of Sciences of the United States of America 90 (1993) 11944–11948.

[25] A. Makropoulos, I. S. Gousias, C. Ledig, P. Aljabar, A. Serag, J. V. Hajnal, A. D. Edwards, S. J. Counsell, D. Rueckert, Automatic whole brain MRI segmentation of the developing neonatal brain, IEEE Transactions on Medical Imaging 33 (2014) 1818–1831.

[26] N. I. Weisenfeld, S. K. Warfield, Automatic segmentation of newborn brain MRI, NeuroImage 47 (2009) 564–572.

[27] A. Buades, B. Coll, J. Morel, A review of image denoising algorithms, with a new one, Multiscale Modeling & Simulation 4 (2005) 490–530.

[28] F. Rousseau, P. A. Habas, C. Studholme, A supervised patch-based approach for human brain labeling, IEEE Transactions on Medical Imaging 30 (2011) 1852–1862.

[29] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, D. L. Collins, Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation, NeuroImage 54 (2011) 940–954.

[30] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, P. A. Yushkevich, Multi-atlas segmentation with joint label fusion, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (2013) 611–623.

[31] M. Liu, A. Kitsch, S. Miller, V. Chau, K. Poskitt, F. Rousseau, D. Shaw, C. Studholme, Patch-based augmentation of Expectation-Maximization for brain MRI tissue segmentation at arbitrary age after premature birth, NeuroImage 127 (2016) 387–408.

[32] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, C. Barillot, An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images, IEEE Transactions on Medical Imaging 27 (2008) 425–441.

[33] S. K. Warfield, K. H. Zou, W. M. Wells, Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation, IEEE transactions on medical imaging 23 (2004) 903–921.

[34] A. J. Asman, B. A. Landman, Non-Local STAPLE: An Intensity-Driven Multi-Atlas Rater Model, Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 15 (2012) 426–434.

[35] S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (2000) 2323–2326.

[36] E. Hughes, L. Cordero-Grande, M. Murgasova, J. Hutter, A. Price, A. Santos Gomes, J. Allsop, J. Steinweg, N. Tusor, J. Wurie, J. Bueno-Conde, J.-D. Tournier, M. Abaei, S. Counsell, M. Rutherford, M. Pietsch, D. Edwards, J. Hajnal, S. Fitzgibbon, E. Duff, M. Bastiani, J. Andersson, S. Jbabdi, S. Sotiropoulos, M. Jenkinson, S. Smith, S. Harrison, L. Griffanti, R. Wright, J. Bozek, C. Beckmann, A. Makropoulos, E. Robinson, A. Schuh, J. Passerat Palmbach, G. Lenz, F. Mortari, T. Tenev, D. Rueckert, The Developing Human Connectome: Announcing the first release of open access neonatal brain imaging, in: OHBM, Procs., 2017.

[37] A. Makropoulos, E. C. Robinson, A. Schuh, R. Wright, S. Fitzgibbon, J. Bozek, S. J. Counsell, J. Steinweg, J. Passerat-Palmbach, G. Lenz, F. Mortari, T. Tenev, E. P. Duff, M. Bastiani, L. Cordero-Grande, E. Hughes, N. Tusor, J.-D. Tournier, J. Hutter, A. N. Price, M. Murgasova, C. Kelly, M. A. Rutherford, S. M. Smith, A. D. Edwards, J. V. Hajnal, M. Jenkinson, D. Rueckert, The Developing Human Connectome Project: A minimal processing pipeline for neonatal cortical surface reconstruction, bioRxiv (2017) 125526.

[38] B. B. Avants, C. L. Epstein, M. Grossman, J. C. Gee, Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain, Medical Image Analysis 12 (2008) 26–41.

[39] J. Ashburner, K. J. Friston, Unified segmentation, NeuroImage 26 (2005) 839–851.

[40] D. Geffroy, D. Rivière, I. Denghien, N. Souedet, S. Laguitton, Y. Cointepas, BrainVISA: A complete sotor-ftware platform for neuroimaging, in: Python in Neuroscience Workshop, Procs., 2011.

[41] S. Liao, Y. Gao, J. Lian, D. Shen, Sparse patch-based label propagation for accurate prostate localization in CT images, IEEE Transations on Medical Imaging 32 (2013) 419–434.

[42] T. Tong, R. Wolz, Z. Wang, Q. Gao, K. Misawa, M. Fujiwara, K. Mori, J. V. Hajnal, D. Rueckert, Discriminative dictionary learning for abdominal multi-organ segmentation, Medical Image Analysis 23 (2015) 92–104.

[43] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Computer Vision and Pattern Recognition, 2001.

[44] Z. Tu, X. Bai, Auto-Context and Its Application to High-Level Vision Tasks and 3d Brain Image Segmentation, Ieee Transactions on Pattern Analysis and Machine Intelligence 32 (2010) 1744–1757.