

Multi-Atlas Segmentation of Biomedical Images: A Survey

Juan Eugenio Iglesias¹ and Mert R. Sabuncu²

¹ Basque Center on Cognition, Brain and Language (BCBL), Spain. email: e.iglesias@bcbl.eu

² A.A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, USA. email: msabuncu@nmr.mgh.harvard.edu

Abstract

Multi-atlas segmentation (MAS), first introduced and popularized by the pioneering work of Rohlfing, Brandt, Menzel and Maurer Jr (2004), Klein, Mensh, Ghosh, Tourville and Hirsch (2005), and Heckemann, Hajnal, Aljabar, Rueckert and Hammers (2006), is becoming one of the most widely-used and successful image segmentation techniques in biomedical applications. By manipulating and utilizing the entire dataset of “atlases” (training images that have been previously labeled, e.g., manually by an expert), rather than some model-based average representation, MAS has the flexibility to better capture anatomical variation, thus offering superior segmentation accuracy. This benefit, however, typically comes at a high computational cost. Recent advancements in computer hardware and image processing software have been instrumental in addressing this challenge and facilitated the wide adoption of MAS. Today, MAS has come a long way and the approach includes a wide array of sophisticated algorithms that employ ideas from machine learning, probabilistic modeling, optimization, and computer vision, among other fields. This paper presents a survey of published MAS algorithms and studies that have applied these methods to various biomedical problems. In writing this survey, we have three distinct aims. Our primary goal is to document how MAS was originally conceived, later evolved, and now relates to alternative methods. Second, this paper is intended to be a detailed reference of past research activity in MAS, which now spans over a decade (2003 - 2014) and entails novel methodological developments and application-specific solutions. Finally, our goal is to also present a perspective on the future of MAS, which, we believe, will be one of the dominant approaches in biomedical image segmentation.

Keywords: Multi-atlas segmentation, Label fusion, Survey

1. Historical Introduction and Background

Segmentation is one of the fundamental problems in biomedical image analysis and refers to the process of tagging image pixels or voxels with biologically meaningful labels, such as anatomical structures and tissue types. Depending on the application, these labels might constitute a handful of, possibly disjoint, regions of interest (ROIs) and a “background”, which would refer to the parts of the image one might ignore in subsequent analysis. Alternatively, the labels might densely cover a substantial portion or all of the image, which is sometimes referred to as “parcellation.”

The traditional approach to segment a given biomedical image involves the manual delineation (sometimes referred to as “annotation”) of the ROIs by a trained expert. This practice, however, can be painstakingly slow, prone to error, hard to reproduce, expensive, and unscalable. Furthermore, the quality of the results will depend on the performance of the expert. Thus, manual delineation is typically not suitable for deploying on large-scale datasets or in applications where time is critical, such as treatment planning. Automatic or semi-automatic segmentation al-

gorithms can address these challenges, by speeding up the process, reducing the cost, and offering reliability, repeatability, and scalability.

Some segmentation algorithms, such as those that assign voxels to tissue types (Kapur et al., 1996), might not require the availability of training data in the form of manually delineated images (commonly called “atlases”). However, the class of methods we consider for this survey will depend on such training data and thus can be viewed as supervised learning algorithms. The goal of atlas-guided segmentation is to use/encode the relationship between the segmentation labels and image intensities observed in the atlases, in order to assign segmentation labels to the pixels or voxels of an unlabeled (i.e., novel) image.

In the early days of atlas-guided segmentation, atlases were rare commodities. In fact, in many applications, there was only a single atlas¹, i.e., a single image that was delineated by an expert. In this context, the classi-

¹The word “atlas” is a legacy of this era, where, for a given problem, one exploited a single map of labels denoting the biological meaning of the observed anatomy, for example, as captured by an image.

cal atlas-guided approach treats segmentation as an image registration problem (Pham et al., 2000), where spatial correspondence is established between the atlas and novel image coordinates. Registration is typically a computationally expensive task that involves deforming (using some appropriate deformation model) one of the images until it is similar to the other one. The resulting mapping between the two coordinate systems can then be employed to transfer (or “propagate”) the segmentation labels from the atlas to the novel image voxels (Christensen et al., 1997; Collins et al., 1995; Davatzikos, 1996; Dawant et al., 1999; Lancaster et al., 1997; Sandor and Leahy, 1997). We refer to this technique as registration-based segmentation.

A single atlas coupled with a deformation model is usually insufficient to capture wide anatomical variation (Doan et al., 2010). Therefore, the use of several atlases is expected to yield improved segmentation results. Initial methods that utilized several atlases for segmentation took a two-step approach. In the first step, the most relevant atlas was identified, which was then used in a second registration-based segmentation step (Rohlfing et al., 2003a). As we will see below, this can be viewed as a special case of multi-atlas segmentation, since all atlases are consulted for segmentation. However, the approach that dominated early atlas-guided segmentation was probabilistic atlas-based segmentation (Ashburner and Friston, 2005; Fischl et al., 2002; Park et al., 2003; Pohl et al., 2006; Yeo et al., 2008), which had two distinctive properties. First, there was a single atlas coordinate frame, defined through the co-registration of the training images used to build the atlas. Second, statistics about the labels, such as the probability of observing a particular label at a given location, are precomputed in atlas space. The novel image was then segmented in the atlas coordinate frame with a probabilistic inference procedure that utilized *parametric* statistical models. The spatial normalization to the atlas could be computed via registration with a population template created at training, or estimated jointly with the segmentation within the probabilistic model; the latter alternative has the advantage that it is adaptive to variations in image intensity profiles, such as MRI contrast (Ashburner and Friston, 2005).

Probabilistic atlas-based segmentation offered two major advantages. First, by employing a single coordinate frame, to which all images were normalized, one automatically established spatial correspondence across all images. This facilitated the statistical analysis of biological variation across the population, as famously exemplified in voxel-based morphometry (Ashburner and Friston, 2000). The second advantage was computational. One needed to run the computationally expensive image registration step (spatial normalization) only once per novel image.

In 2003-2004, in a series of papers (Rohlfing et al., 2003b,c,d, 2004), Rohlfing and colleagues proposed an alternative segmentation strategy, which at the time might have not seemed radically different. Yet, as we elaborate below, this work inspired a rapidly growing class of meth-

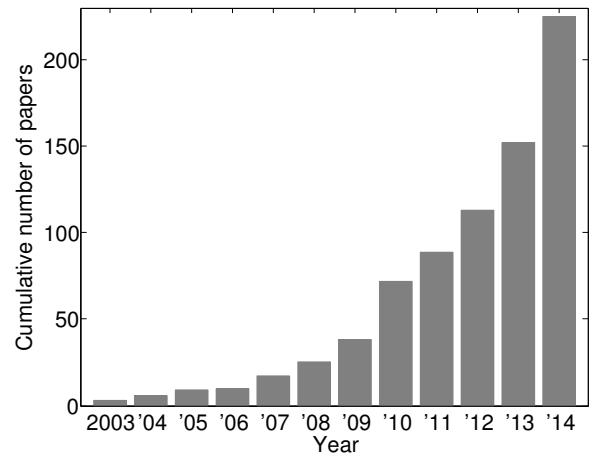


Figure 1: Cumulative number of papers, cited in this survey, that introduce a novel MAS method or present a novel application of MAS.

ods (see Figure 1), including the pioneering work of Klein et al. (2005), Heckemann et al. (2006) and others. We collectively refer to these methods as *multi-atlas segmentation* (MAS). In this approach, the atlases are *not* summarized in a (probabilistic) model. Instead, each atlas is available and potentially used for segmenting the novel image. A classical example involves applying a pairwise registration between the novel image and each atlas image. These registration results are then used to propagate the atlas labels to the novel image coordinates, where at each voxel, the most frequent label is selected. This is commonly referred to as “majority voting.”

We can subdivide a MAS algorithm into several components that we depict in Figure 2. These components might be implemented as independent, sequential steps, where earlier steps are placed above in the illustration. However, there are many exceptions to this structural organization. For example, in some algorithms, some blocks might be unified, form feedback loops, swap places, or even be omitted altogether. That said, we find this diagram useful for organizing methodological developments in MAS. Therefore, the part of our survey covering methods will adhere to this organization, with subsections corresponding to each one of these components.

The remainder of this survey is structured as follows. Section 2 presents an account of published MAS methods, organized into the aforementioned building blocks. Section 3 surveys published studies that apply a MAS algorithm to a novel biomedical problem. We conclude with a discussion and pointers to promising future directions of research in Section 4. Finally, we would like to note that we have made all effort to cover the relevant literature as comprehensively as possible (as of the end of 2014). Yet, we are bound to have missed some pertinent publications. Furthermore, we made the conscious choice to leave out some redundant papers. For example, earlier conference versions of more detailed journal publications were typi-

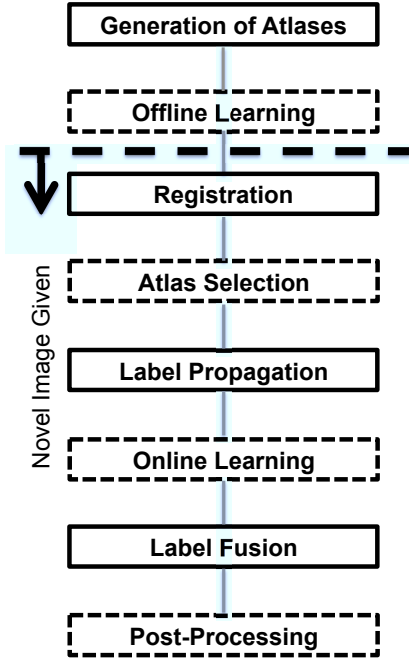


Figure 2: Building blocks of MAS. Dashed blocks can be considered optional.

cally omitted.

2. Survey of Methodological Developments

2.1. Generation of Atlases

Atlases, i.e., labeled training images, form the core foundation of MAS algorithms. They are typically obtained by the meticulous and costly effort of a domain-specific expert who relies on an interactive visualization software, such as (Criminisi et al., 2008; Heiberg et al., 2010; Pieper et al., 2004; Yushkevich et al., 2006), and might exploit multiple imaging modalities, while recruiting textbook anatomical knowledge. However, as we discuss below, there are exceptions to this rule.

Before seeing the to-be-segmented novel image, most methods treat each manually segmented image equally. Yet, to improve performance, one might identify high quality training cases, for example via visual inspection (Yang et al., 2010). Another option is to borrow ideas from the feature selection literature (Pudil et al., 1994) to automatically preselect the subset of atlases that is expected to yield the maximal performance when new data are segmented (Iglesias and Karssemeijer, 2009; Isgum et al., 2009). Prior knowledge on the problem at hand can also be used to preselect the atlases. For example, Tung et al. (2013) selected the atlases with narrow lumen in a neointima segmentation method in coronary optical coherence tomography (OCT), since neointima only exists in coronary arter-

ies with narrow lumen. Such approaches can increase the accuracy of segmentation by discarding low quality training data, but a decrease in segmentation quality is also possible due to the reduction of the atlas pool size, which also represents a waste of labeling effort. Another way of improving the performance of the segmentation is to apply population-level preprocessing (e.g., by co-registering the atlases) to increase their signal-to-noise (SNR) ratio (Zhuang et al., 2010). While higher SNR can yield a more accurate registration and segmentation of the novel image, this improvement will depend on the quality of the co-registration and noise properties of the data. A similar approach was recently used to compute multiple population templates via a clustering strategy, which were manually labeled (Gao et al., 2014).

In applications where the atlases might not be a representative sample of the population, one can synthesize atlases that offer a better representation of anatomical variability. Despite the higher computational cost in the subsequent analysis, such an approach can increase the accuracy of the segmentation by enriching the atlas pool. For example, Jia et al. (2012) used a statistical model based on principal component analysis (PCA) to synthesize deformations. One disadvantage of this technique is that the synthetic deformations might not always be anatomically plausible. In a related effort, Doshi et al. (2013) proposed to cluster all available training images using k-means on the L2 norm of intensity differences in order to identify a representative subset of cases that can be then manually labeled. Their algorithm, which is closely related to active learning, can greatly reduce the manual labeling effort, but also requires a large pool of unlabeled data. Recently, Awate and Whitaker (2014) presented a strategy that used a small number of labeled cases and a model of MAS based on non-parametric regression in the space of images, in order to predict the total number of atlases that need to be manually segmented to obtain a desired level of segmentation accuracy within a MAS framework. This technique can be useful for planning the manual segmentation phase.

In an alternative approach, one might exploit the wide availability of non-expert segmenters, instead of trying to obtain high quality expert manual segmentations. Trading off quality of annotations for number of atlases can be beneficial in some applications. For example, Bogovic et al. (2013) consider this scenario and propose to directly model the unknown “expertise” of each atlas, encoded in a confusion matrix between true and estimated labels. Bryan et al. (2014), on the other hand, consider relying on the self-declared “confidences” of the manual segmenters, which are used to weigh their contributions when merging their “opinions” in the segmentation of novel images. In other scenarios, atlases might have been segmented multiple times, as in (Weisenfeld and Warfield, 2011b; Wang and Yushkevich, 2012a), or only portions (e.g. certain slices) of the training data might have been manually traced, as in (Landman et al., 2012b). This information needs to be considered in the subsequent steps of

the MAS pipeline. Finally, there are several proposed approaches, e.g. (Chakravarty et al., 2013; Gass et al., 2013; Heckemann et al., 2006; Jia et al., 2012; Liao et al., 2013; Shen et al., 2010; Wang and Yushkevich, 2013a; Wolz et al., 2010a; Kotrotsou et al., 2014; Janes et al., 2014), that exploit the novel, unlabeled images to enrich the training data, for instance, by employing automatic segmentations as atlases, or by using them to generate different registration paths between the atlases and the target scan to segment. The former can be seen as a form of self-training (a semi-supervised learning technique in which unlabeled data are automatically classified and added to the pool of training samples). While it can take advantage of unlabeled data, it also inherits self-training’s shortcoming that segmentation mistakes reinforce themselves. The latter (using multiple registration paths) can increase the performance by generating more training segmentations for label fusion (i.e., the combination of propagated labels into a segmentation estimate), but at the same time, these might be poor candidates due to suboptimal registration and eventually worsen the final segmentation.

2.2. Offline Learning

Classical MAS algorithms applied no or very little processing to the atlas data *offline*, i.e., prior to observing the novel image. Atlases were manipulated and analyzed solely based on information from the image to be segmented. However, some of the more recent methods we review here perform what we call “offline learning,” where the atlases are analyzed offline and some sort of information is garnered to be used during the segmentation of the novel image. For example, one can learn a strategy to compute rough regions of interest in the novel image, in order to constrain or guide subsequent processing steps and/or reduce computational cost (Li et al., 2013; Ramus et al., 2010; van Rikxoort et al., 2010; Shi et al., 2010). In a very different approach, van der Lijn et al. (2008) proposed to construct a likelihood model on the training data, which quantifies the probability of observed image intensities conditioned on the underlying labels. While such a model can improve the segmentation by linking labels and image intensities, it can also degrade it if the intensities of the atlases and novel image are not well matched. Similarly, Zikic et al. (2014b) suggest to train a random forest classifier corresponding to each atlas, which learns to predict labels based on the image appearance. Instead of labels propagated via a registration step, atlas predictions computed using these classifiers are combined into a segmentation. Since classification with random forests is fast, this method can be computationally more efficient than conventional registration-based multi-atlas segmentation, but also less accurate, as it will fail to capitalize on the high accuracy of modern registration methods. In a related effort, Wang and Yushkevich (2013b) considered a tumor segmentation application, where the algorithm cannot rely on spatial correspondences between the images. Instead, they employed a data-driven clustering strategy

on atlas voxels to identify super-voxels (i.e., patches of irregular size), which were then used by a k-nearest neighbor classifier to segment the novel image. In a parallel effort, Wang et al. (2014a) proposed to use a local random forest classifier trained on the atlases to predict the segmentation label in the novel image.

Another direction involves analyzing the training data in order to learn how to assign weights to each atlas when conducting label fusion. One such strategy estimates measures of reliability associated with the atlases by co-registering them and computing the agreement between the propagated labels; atlases that can better predict the labels of others’ receive higher weights (Sdika, 2010; Wan et al., 2008). Alternatively, supervised learning approaches have been proposed to predict the weights from the novel image. For example, Sjöberg and Ahnesjö (2013) pre-registered the atlases and learned the distribution of Dice scores given the registration similarity measure; label fusion weights were derived from this distribution. In a similar fashion, Sanroma et al. (2014a) trained a support vector machine to predict a ranking of the atlases based on image features. The main disadvantage of these approaches is that they do not always generalize well beyond the training data. A related, yet different technique involves applying clustering (Langerak et al., 2013; Shi et al., 2010), manifold learning (Cao et al., 2011b,a, 2012; Duc et al., 2013; Wolz et al., 2010a; Gao et al., 2014), or computing a minimum spanning tree on the atlases (Jia et al., 2012). These learning algorithms are employed to construct a structure on the space of training images, which yields the means to efficiently compute distances between the atlases and novel image(s), run registrations, and propagate manual labels.

2.3. Registration

Registration is the task of establishing spatial correspondence between images and is considered one of the fundamental problems in biomedical image processing. Image registration involves deforming (or warping) one or more images to maximize an objective function that combines a metric of spatial alignment with a regularizer that quantifies the plausibility of the deformation. The three components of an image registration algorithm are thus the deformation model, the objective function, and the optimizer. The deformation model represents the class of spatial transformations that are plausible in the application at hand. This can be as simple as a rigid transform, or as complex as a non-parametric model in which each location is assigned a spatial transformation vector. Some deformation models incorporate constraints that exploit prior knowledge to make the spatial transforms more realistic. These constraints can be integrated in the deformation model (e.g., inverse consistency, diffeomorphism), or explicitly specified in the objective function through regularizers. The objective function is typically based on either the spatial distance between corresponding landmarks (manually placed or automatically detected) or on

image intensities. In the latter case, metrics such as sum of squares or cross-correlation have been widely used in intramodality scenarios, whereas statistical metrics such as mutual information have been popular when registering across modalities. Finally, the optimization method is often an iterative algorithm, e.g., gradient descent, conjugate gradient, Levenberg-Marquardt, or the BFGS algorithm. However, discrete, graph-based methods are also becoming popular, e.g., Glocker et al. (2008). An extensive review on deformation models, objective functions and optimizers can be found in Sotiras et al. (2013).

The optimal choice of algorithm specifics largely depends on the biomedical application, its goal (Yeo et al., 2010), and operational constraints, such as available computational resources, desired accuracy, and restrictions on time. Once registration is complete, the resulting spatial transform can then be used to map from the frame of one image to the coordinates of another.

In MAS, registration is the step that determines the spatial correspondence between each atlas and the novel image. Early MAS methods, such as (Heckemann et al., 2006; Rohlfing et al., 2004; Wan et al., 2008), relied on nonlinear deformation models, such as spline-based parameterized transformations (Rohde et al., 2003; Rueckert et al., 1999) or non-parametric diffeomorphisms (Beg et al., 2005; Vercauteren et al., 2009), which seek voxel-level alignment accuracy. Several studies (Bai et al., 2012; Lotjonen et al., 2009; Lötjönen et al., 2010; Sjöberg and Ahnesjö, 2013) have conducted empirical comparisons of the impact of different registration algorithms on MAS performance in different applications.

Typically, one independent registration is computed between each atlas and the novel image and generic intensity-based registration tools, such as (Avants et al., 2009; Klein et al., 2010; Rueckert et al., 1999; Ou et al., 2011), are used. Yet, Rohlfing and Maurer Jr (2005) experimented with running the registration step several times with different parameter settings and combining all resulting propagated labels. As explained in Section 2.1 above, generating more candidate segmentations can improve the subsequent fusion, but also might worsen it by introducing poor candidates generated by suboptimal registrations. A similar strategy, proposed by Wang et al. (2013a), employs pre-computed registrations between pairs of atlases to generate a multitude of propagated labels by concatenating the pairwise registration results. In a parallel effort, Datteri et al. (2014) relied on pre-registered atlases to estimate registration accuracy for the novel image based on registration circuits. Also, several authors proposed to employ the manual segmentations (Han et al., 2008; Nie and Shen, 2013; Tamez-Pena et al., 2012; Lee et al., 2014a), multiple imaging channels (Yushkevich et al., 2010), or automatically computed tissue maps (Heckemann et al., 2010; Ledig et al., 2014) to establish more accurate alignment, often at a higher computational cost. Instead of computing the pairwise atlas-to-novel image registrations independently, Lee et al. (2014a) recently proposed to solve them simul-

taneously in a group-wise registration framework. In another parallel effort, motivated by the observation that the registration step would benefit from the knowledge of the underlying segmentation labels, Hao et al. (2012), Iglesias et al. (2013c), Tang et al. (2013), and Stavros et al. (2014) developed MAS algorithms that integrated the registration and label fusion steps. Thus, instead of treating registration as an independent preprocessing step, these algorithms iterate between registration and segmentation, which yields a small increase in segmentation accuracy at the expense of reduced computational efficiency.

Typically, MAS treats the unknown deformation between the atlas and novel image as a nuisance, which once computed is only used to deform the atlas image intensities and/or propagate the labels. Yet, a growing number of methods recognize the value in the deformation fields themselves and propose to use information about the amount of deformation in the computations of the fusion weights. For example, Commowick and Malandain (2007) used the Euclidean norm of the deformation, Ramus et al. (2010) used its Jacobian determinant, and Wang et al. (2014b) used its harmonic energy.

The registration step is the computational bottleneck of the MAS algorithm and largely determines run time. One strategy to reduce the computational burden introduced by registration is via atlas selection (see next section), which can obviate expensive registrations with unselected atlases. An alternative, popular approach employs a common coordinate system, similar to conventional probabilistic atlas segmentation methods, either via a standard template (Aljabar et al., 2007, 2009), a population average (Artaechevarria et al., 2008; Commowick and Malandain, 2007; Commowick et al., 2009; Depa et al., 2011; Fonov et al., 2012; Ramus et al., 2010; Shi et al., 2010, 2013; Zhuang et al., 2010), or one of the atlases (van Rikxoort et al., 2010; Sjöberg et al., 2014). Here, all atlases are co-registered offline, and the novel image is registered with the template image. The template-to-novel image transformation can then be concatenated with the atlas-to-template transformations in order to propagate labels from the atlases to the novel image (Artaechevarria et al., 2008; Depa et al., 2011; Ramus et al., 2010; Rivest-Hénault et al., 2014). Such an approach can reduce the computational cost of registration, but also might negatively impact the performance due to the suboptimality of the registrations. The use of a common coordinate frame further enables the definition of regions of interests, which the segmentation algorithm can employ in subsequent steps, e.g., atlas selection (Shi et al., 2010) or label fusion (Commowick et al., 2009; Ramus et al., 2010). Yet another strategy to accelerate the registration step is to exploit the rapidly growing availability and capability of GPU processors, as proposed in (Duc et al., 2013; Han et al., 2009; Modat et al., 2010; Cardoso et al., 2013).

A recent technique is inspired by the non-local means method (Buades et al., 2005) and utilizes a patch-based search strategy to identify correspondences with the at-

lases. This technique was introduced to biomedical MAS by Coupé et al. (2011) and recently has been gaining popularity (Asman and Landman, 2013; Bai et al., 2013; Fonov et al., 2012; Konukoglu et al., 2013; Rousseau et al., 2011; Wang et al., 2011, 2013b; Wolz et al., 2013; Wang et al., 2014e; Zhang et al., 2011, 2012; Wang et al., 2014c,a; Ta et al., 2014; Wang et al., 2014f,d; Sanroma et al., 2014b). These papers have demonstrated that a patch-based search strategy can be used in a wide range of MAS methods to improve performance, for example, by relaxing the one-to-one correspondence assumption or eliminating the need for highly accurate registration results. In contrast with common implementations of non-local means in computer vision, its biomedical applications can be computationally efficient, cf. (Ta et al., 2014), for instance, by assuming a rough alignment (e.g., achieved via a linear transformation model), which allows one to restrict the patch search to a local neighborhood of each voxel. Furthermore, the anatomical context can be used to improve the quality of the patch matches, as demonstrated in (Wang et al., 2014f,d).

2.4. Atlas Selection

There are two main motivations not to use all available atlases in MAS. First, by reducing the number of atlases, one can improve computational efficiency. This might be particularly important for applications where time is a significant constraint. A typical MAS algorithm’s computational demand is at least linear with respect to the number of utilized atlases. So, selecting only half of all available atlases would be expected to about double the speed of the algorithm and reduce the memory requirements by up to a half. Second, by excluding irrelevant atlases that might misguide the segmentation procedure, one might expect to improve final segmentation accuracy. The specifics of the problem and utilized algorithm determine how applicable and significant these two points are. For example, it has been observed that atlas selection can improve the accuracy of majority voting (Aljabar et al., 2007), but is less critical for weighted fusion (Sabuncu et al., 2010).

The effectiveness and efficiency of atlas selection are closely related to the registration step. If all atlases are nonlinearly registered to the novel images, the atlas selection step will be well informed by the similarity of the images, but the computational savings will be nonexistent or minimal. If no registration is performed, the computational efficiency of the algorithm is much higher, but it is more difficult to select the atlases that are most relevant to the novel images to segment. A compromise can be achieved by linearly registering all the atlases to the novel scan, performing the selection, and continuing with the nonlinear component of the registration only for the chosen atlases. While the selection is typically conducted prior to the segmentation of the novel image, several authors (Langerak et al., 2013, 2010; van Rikxoort et al., 2010; Weisenfeld and Warfield, 2011a) have proposed methods that iterate between segmentation and at-

las selection, pruning or adding to the selected atlas set based on the current estimate of the segmentation.

Early atlas selection methods employed a metric to rank the relevance of the atlases. These metrics included similarity measures based on image intensities, e.g., sum of squared differences, correlation or mutual information (Aljabar et al., 2007, 2009; Aribisala et al., 2013; Tung et al., 2013; Xie and Ruan, 2014; Wu et al., 2007); non-image meta-data such as age (Aljabar et al., 2007; Aribisala et al., 2013); registration consistency (Heckemann et al., 2009); amount of deformation (Commowick and Malandain, 2007; Commowick et al., 2009); and anatomical geometry (Teng et al., 2010). Several studies have conducted empirical comparisons of these different atlas selection strategies in various MAS applications (Aljabar et al., 2007; Acosta et al., 2011; Avants et al., 2010; Lötjönen et al., 2010; Ramus and Malandain, 2010), concluding that sum of squared differences and cross-correlation (after histogram matching) of intensity values, along with age difference, are reliable metrics to rank the relevance of atlases.

More recently, other works have proposed to define similarity measures based on an image manifold structure. For example, Duc et al. (2013) used Isomap, locally linear embedding and Laplacian eigenmaps to learn the manifold. Cao et al. (2011a) used locality preserving projections, and Asman et al. (2014) used principal components analysis. These approaches introduce additional complexity to the system, but can outperform standard similarity measures in the atlas selection task. Another approach to increase the efficiency and accuracy of atlas selection utilizes clustering, where the atlases, possibly together with the novel image(s), are analyzed to identify clusters of similar cases using methods such as k-means (Nouranian et al., 2014), affinity propagation (Langerak et al., 2013) and Floyd’s algorithm (Wang et al., 2014b). Then, cluster representatives (or exemplars) are used for the initial search of the most relevant atlases. These methods can yield a performance similar to approaches that do not preselect atlases, but at a much lower computational cost. Alternatively, atlas selection can be treated as a learning problem, where the optimal strategy to choose the relevant atlases can be learned on the atlases themselves, utilizing the manual segmentations, as demonstrated by Cao et al. (2012) (manifold learning), Konukoglu et al. (2013) (random forests) and Sanroma et al. (2014a) (support vector machines). These strategies are not straightforward to implement but have been shown to improve segmentation performance.

While the atlas selection method has a significant impact on segmentation performance, with notable exceptions (Awate and Whitaker, 2014; Heckemann et al., 2006), the optimal number of atlases to be selected seems to be an overlooked topic of research. Some algorithms simply choose the most suitable single atlas, and apply registration-based segmentation (Commowick and Malandain, 2007; Teng et al., 2010; Wu et al., 2007). Yet most MAS methods end up using more than one atlas (Rohlfing et al., 2004; Klein et al., 2005; Heckemann et al., 2006). Typically, the

number of atlases to be selected is either estimated, e.g., based on heuristics such as computational expectations, or determined empirically via cross-validation, bootstrapping or a similar sampling strategy.

2.5. Label Propagation

Once the relevant atlases are selected, and spatial correspondence is established with the novel image, the classical multi-atlas segmentation strategy proceeds by propagating the atlas labels to the novel image coordinates. Since early MAS methods (Heckemann et al., 2006), one of the most popular strategies has been to utilize “nearest neighbor interpolation,” where each atlas transfers a single label to each novel image voxel, e.g., (Artaechevarria et al., 2009; Langerak et al., 2010, 2013; Sabuncu et al., 2010). Although this label corresponds to that of the closest voxel in atlas space, Sdika (2010) showed that higher performance can be achieved by augmenting the information with a tissue consistency step. That is, the nearest neighbor search is conducted among those atlas voxels with a tissue segmentation (obtained automatically, from a separate step) consistent with the target voxel. However, this approach depends on the performance of the separate tissue segmentation step, which can be sensitive to outliers, as in the case of very old or diseased subjects not represented in the atlas pool. The nearest neighbor strategy can further be refined using, for example, linear interpolation (Rohlfing et al., 2004; Sabuncu et al., 2010), where each atlas’s vote is spread over multiple labels, with associated weights that reflect the ratio of partial volumes.

An alternative approach involves using the signed distance maps of the original atlas label images (Gholipour et al., 2012; Gorthi et al., 2013; Sabuncu et al., 2010; Sjöberg and Ahnesjö, 2013; Weisenfeld and Warfield, 2011a; Xu et al., 2014b). Each label has an associated signed distance map, which takes positive values within the corresponding structure, negative values outside, and the magnitude is proportional to the closest distance to the label boundary. The signed distance map encodes the uncertainty close to label boundaries and the relative confidence deep within a region. While signed distance maps are not naturally normalized (i.e., the scale depends on the size and shape of the anatomical structure), one strategy is to use them to compute label probabilities, e.g., via the logistic mapping (Sabuncu et al., 2010). A complementary technique transforms the atlas label boundaries directly, rather than applying a volumetric warp to the images (Chou et al., 2008; Klein et al., 2005; Nie and Shen, 2013; Tamez-Pena et al., 2012). Finally, rather than transferring over atlas labels via a geometric deformation model, one can employ learning algorithms trained on each atlas to generate voxel-level candidate label estimates for each atlas, as recently proposed in (Zikic et al., 2014b). This technique does not seem to increase segmentation accuracy, but can considerably reduce the run time of the algorithm by requiring only one nonlinear registration (to align a probabilistic atlas to the novel scan).

2.6. Online Learning

The labels of the atlases that have been propagated to novel image coordinates are often merged directly into a single estimate of the segmentation with a label fusion algorithm. However, several MAS methods perform an “online learning” step, which aims to boost the performance of the algorithm by exploiting the relationships between the registered atlases and the novel images.

Some methods use the estimated segmentation of the novel image to iteratively perform atlas selection and/or registration. For example, the selected atlas set can be determined based on the similarity between the deformed atlas labels and the current estimate of the segmentation, which can increase segmentation accuracy by excluding outlier atlases from fusion (Langerak et al., 2010). Alternatively, van Rikxoort et al. (2010) divide the novel image into blocks, which are used to update the local registrations and selection of atlases. This way, they are able to automatically stop the local registration of atlases when no improvement is expected, reducing the computational cost without a negative impact on performance.

Other approaches exploit the relationship between the labels and intensities of the novel image in order to assist the fusion step. This can be achieved via conditional Gaussian models (Lotjonen et al., 2009), or non-parametric density estimators (Weisenfeld and Warfield, 2011a), which can be employed to refine the propagated labels. However, this strategy can be counterproductive when the intensities of the atlases and the novel scan are not well matched. In a related effort, Hao et al. (2014) use a discriminative technique to model the posterior label probability directly. More specifically, they use the intensities and labels of the registered atlases in a window around each spatial location to build a set of local classifiers – one per voxel of the novel image. Each classifier is an L1-regularized support vector machine that predicts the label of the voxel at hand from hundreds of image features computed with different filters. In this case, label fusion is implicitly carried out in the classification.

Finally, we have semi-supervised approaches that utilize the collection of novel images along with the atlases. For example, the LEAP algorithm (Wolz et al., 2010a) first learns a manifold structure on all (novel plus training) images. Next, a small number of novel images closest to the atlases are automatically segmented via a multi-atlas procedure. These automatically segmented novel images are then added to the atlas list and the whole procedure is repeated. By using stepping stones, this strategy boosts the performance in cases where some novel images are considerably different from the atlases. There are other algorithms that also rely on self-training, i.e., using the automatically segmented images as new training data (Chakravarty et al., 2013; Liao et al., 2013; Shen et al., 2010; Wang and Yushkevich, 2013a). These approaches have the disadvantage that segmentation mistakes reinforce themselves. Another way of exploiting un-

labeled data is to use unlabeled scans to generate multiple deformations of a single labeled atlas to a novel scan, again using the unlabeled volumes as stepping stones (Gass et al., 2013).

2.7. Label Fusion

Label fusion, i.e., the step of combining propagated atlas labels, is one of the core components of MAS. The earliest and simplest fusion methods are best atlas selection (Rohlfing et al., 2004) and majority voting (Heckemann et al., 2006; Klein et al., 2005; Rohlfing et al., 2004). In best atlas selection, a single atlas is utilized, which is usually chosen based on examining the match between the registered atlas and novel image intensities, for example, as captured by the registration cost function (e.g., sum of squared differences, normalized cross-correlation, or mutual information). Relying on a single atlas disregards potentially useful information in all other atlases. Majority voting chooses the most frequent label at each location, therefore using information from all atlases at all locations; however, it has the drawback that it ignores image intensity information.

An extension of majority voting is weighted voting, where each atlas is associated with a weight (global or local) that reflects the similarity between the atlas and novel image. The first method using global weights was proposed by Artaechevarria et al. (2008), who used weights proportional to the normalized mutual information between the registered atlas image and novel image intensities. A related approach is to estimate the weights by posing it as a least squares problem, where the novel image intensities are assumed to be equal to the weighted combination of atlas intensities (Cao et al., 2011b). An alternative strategy involves defining the weights based on the similarity of the labels, which can be computed iteratively with respect to the current segmentation, either globally (Langerak et al., 2010) or within predefined ROIs (Langerak et al., 2011); or estimated by examining the pairwise similarities between the atlases (Datteri et al., 2011).

Global weights cannot model the spatially-varying nature of registration accuracy. For this reason, the use of global weights was later replaced by local and semi-local weighting schemes. The earliest examples of this strategy used weights inversely proportional to the absolute difference between local intensities of the novel image and deformed atlas (Isgum et al., 2009; Iglesias and Karssemeijer, 2009), and standard local intensity-based registration metrics such as local cross-correlation (Artaechevarria et al., 2009). Alternative local weighting strategies were further explored, including the use of a precomputed local reliability measure (Wan et al., 2008), the Jacobian determinant of the deformation fields (Ramus et al., 2010), a Gaussian intensity difference function (Depa et al., 2010; Jia et al., 2012), the inverse of the squared standard score (Tamez-Pena et al., 2012), a measure of the saliency of each atlas (Ou et al., 2012), local mutual information (Nie and

Shen, 2013), estimates of local registration accuracy (Datteri et al., 2014), and structural relationships between locally extracted wavelet features (Kasiri et al., 2014). Other studies have used weights defined as a function of ranks of local similarity, computed with correlations (Yushkevich et al., 2010) or Jacobian determinants (Doshi et al., 2013). Bridging global and local weighting, Wolz et al. (2013) used weights that combined three different terms, reflecting global, organ-level and local (intensity-based) similarities. In a related effort, a combination of region-wise and voxel-wise similarities (all based on sum of squares) were used in (Xie and Ruan, 2014). Even though these strategies improve the segmentation accuracy obtained with global weights, the optimality of the chosen local weight metrics remains unclear.

In a series of papers, Wang et al. (2011, 2013b) computed fusion weights that exploited the correlation structure between the atlases. The weights were optimized to minimize the expectation of segmentation error, which in turn led to relaxing the common independence assumption on the atlases. Moreover, the registration-determined correspondence was refined via a local patch search. In a later paper, the same authors improved their algorithm to make the segmentations of the novel images consistent, such that the automatic segmentations are recruited as atlases, but with a lower weight than the manually labeled ones (Wang and Yushkevich, 2013a).

Other works have used more complicated schemes to define local weights, for example via offline learning. One such method assumed that the weights were a linear combination of the dissimilarities of the voxels at each location, and learned them with Tikhonov-regularized least squares (ridge regression) (Khan et al., 2011). Another related approach pre-registered all the atlases with each other to compute a reliability metric as the average agreement of the propagated labels; the reliabilities were then used as weights in the fusion (Sdika, 2010). Along a similar direction, Zhang et al. (2011) used a forward-backward, patch-based search to compute a measure of correspondence specificity with respect to each atlas. Label fusion is then conducted in a sequential manner, starting at voxels that the algorithm is confident about segmenting and employing already segmented voxels within the neighborhood for guiding the segmentation of yet-to-be-labeled voxels. A related method that was recently proposed by Koch et al. (2014), uses a graph that connects similar regions across images to allow label information to iteratively flow from high confidence to low confidence voxels. In a different approach, Wachinger and Golland (2012) used spectral clustering to identify homogeneous regions, and then performed semi-local label fusion within each region to finally compute a single label per region by pooling the votes within its boundaries.

An alternative label fusion strategy involves the use of patches to compute weights at each voxel, which can be used with a conventional label fusion method (Coupé et al., 2011; Fonov et al., 2012). This technique has recently be-

come more sophisticated. For example, Xiao et al. (2014a) and Wang et al. (2014d) used this approach to compute local label fusion weights using multi-channel MRI data. Wang et al. (2014f) proposed to use the anatomical context to improve the quality of the patch matches. Instead of labeling the central voxel, one can segment the whole patch, and overlapping segmentations can then be fused (e.g. via majority voting) (Rousseau et al., 2011; Sanroma et al., 2014b). These methods have produced state-of-the-art segmentation accuracy, often at a high computational cost.

Rather than directly using the similarity between patches, one can also compute the label fusion weights by seeking sparse linear combinations from a patch dictionary to reconstruct each patch of the novel image (Liao et al., 2013; Zhang et al., 2012; Wang et al., 2014d). Along a similar direction, Sanroma et al. (2014b) recently formulated label fusion as a matrix completion problem, which can be viewed as unifying the weight estimation framework with a learning-based approach. Cao et al. (2011a) also used weights that best reconstructed the intensities of the novel images from the k nearest atlases, computed on an image manifold. Instead of focusing on reconstruction error, a different method (Wu et al., 2014) involves modifying the framework to reflect the consistency in the segmentations, such that atlases that propagate similar labels to the segmentation have a similar contribution.

A different view of label fusion formulates segmentation as an optimization problem, where the agreement with the propagated atlases makes up a data fit term. In this framework, one can incorporate prior expectations such as spatial and temporal smoothness in longitudinal data, as in (Li et al., 2014).

Some of the label fusion techniques discussed above can be derived from probabilistic models of the data. Casting a segmentation method as a Bayesian inference problem in a probabilistic model has several advantages. First, it can easily deal with missing data, e.g., lack of labels in a given region of an atlas. Second, the estimated parameters of the model might have a direct interpretation that can provide us with some insight about the data and the fit of the model. Third, the modeling assumptions have to be clearly stated and their effect on final accuracy can be empirically examined. Fourth, the impact of the inference or estimation strategy can also be assessed by investigating alternative methods. Finally, Bayesian methods are based on a principled and flexible framework, which can be adapted to the specifications of the problem at hand.

A generative probabilistic model of label fusion was first proposed by Sabuncu et al. (2010, 2009). The model comprises of an *unknown* discrete membership field that indexes the atlas that “generated” each voxel of the novel image and an additive Gaussian noise component. This generative model framework unifies some of the most popular label fusion algorithms, generalizing local, semi-local and global weighted fusion methods, including majority voting and best atlas selection. The generative model

has been extended to intermodality fusion (Iglesias et al., 2013b), replacing the Gaussian noise by a joint histogram; and to patch-based fusion (Bai et al., 2013), by augmenting the membership field with a spatial shift and defining the intensity likelihood term as a function of patches.

There is a family of generative models for label fusion that can be viewed as a modification of Sabuncu’s model (Sabuncu et al., 2009, 2010), where the latent membership field is only used to define a prior on the labels and the novel image intensities are generated directly from the underlying segmentation, e.g., via a parametric Gaussian. This model does not utilize the relationship between the image intensities and labels observed in the atlases and thus can be used to segment images of a modality different from the atlases (Iglesias et al., 2012a), or multi-channel images (Iglesias et al., 2012b). This strategy will be suboptimal for scenarios where the intensity profiles of the atlases and the novel scans are matched. Iglesias, Tang and colleagues later proposed to integrate registration into this generative model (Iglesias et al., 2013c; Tang et al., 2013), which offers a small but significant improvement in segmentation accuracy at an increased computational cost. Finally, many methods that use label fusion to construct a prior in a probabilistic segmentation algorithm (Lotjonen et al., 2009; van der Lijn et al., 2008; Van Der Lijn et al., 2012; Wang et al., 2014c; Wachinger and Golland, 2014; Wolz et al., 2009, 2010b; Platero and Tobar, 2014) can be viewed to be (approximate and/or modified) instantiations of the probabilistic generative label fusion framework.

Another family of probabilistic fusion methods builds on the STAPLE algorithm (Warfield et al., 2004). STAPLE was originally developed to model manual segmentations as noisy observations of the hidden (ground truth) segmentation and the noise was modeled with a stationary confusion matrix $\{\theta_n\}$. The original STAPLE algorithm only supported binary segmentations (Warfield et al., 2004), but was soon after extended to the multi-class setting (Rohlfing et al., 2003b,c,d). Many extensions of STAPLE correspond (or can be shown to correspond) to modifications of the original probabilistic model, for example placing a Beta prior on the parameters of the confusion matrix (Commowick and Warfield, 2010), replacing the hard atlas segmentations with probabilistic maps (Weisenfeld and Warfield, 2011b), dealing with missing atlas label data (Landman et al., 2012b), altering the confusion matrix to account for self-assessed uncertainty (Asman and Landman, 2011; Bryan et al., 2014), employing a hierarchical noise model (Asman and Landman, 2014), introducing and estimating unknown reliability weight maps (Akhondi-Asl et al., 2014), and learning and exploiting the relationship between performance parameters and intensity similarities (Gorthi et al., 2014).

Rather than making explicit changes to the original framework and solving the corresponding model, some extensions of STAPLE are based on *ad-hoc* modifications. For instance, some researchers have introduced spatially varying performance parameters to the model by estimat-

ing local confusion matrices from windows around each voxel (Asman and Landman, 2012; Commowick et al., 2012). One can view these methods as approximate solvers to a version of the STAPLE model, in which the noise parameters vary smoothly over space. In a different approach, Langerak et al. (2010), Cardoso et al. (2013), Nouranian et al. (2014), and Weisenfeld and Warfield (2011a) proposed using only a subset of atlases in label fusion. Langerak *et al.*’s SIMPLE algorithm integrates the atlas selection step into STAPLE and solves for that iteratively. The SIMPLE method was recently integrated with context learning to exploit exogenous information, e.g., about tissue likelihood (Xu et al., 2014a). Cardoso *et al.*’s method obtains the subset by ranking the atlases in terms of local similarity to the novel image; Nouranian *et al.*’s algorithm iteratively computes the segmentation with STAPLE and removes the atlases that do not agree with the current estimate of the labels; and Weisenfeld *et al.* use a probabilistic formulation to disregard atlases that do not agree with the current segmentation estimate. Again, these algorithms can be seen as an approximate solution to a model, in which the atlases to be explained are indicated by a latent field. Finally, Asman and Landman (2013) incorporated information from intensity image patches in STAPLE. From a probabilistic modeling perspective, this approach would require modifying STAPLE’s model to connect the novel image intensities to the training images.

2.8. Post-processing

The label fusion result does not necessarily represent the final segmentation; sometimes it is fed to another algorithm to estimate the output labels. The extent to which this post-processing changes the segmentation varies across methods.

Some methods use the output of label fusion to simply initialize a subsequent algorithm, for instance, to determine the bounding box where a segmentation method is applied (van Rikxoort et al., 2007a), to start the evolution of an active contour (Fritscher et al., 2014; Hollensen et al., 2010), or to fit a smooth contour to the object boundary (Nouranian et al., 2014). Other MAS algorithms rely on applying heavy post-processing to the label fusion output, for example by employing an error detection and correction classifier (Yushkevich et al. 2010, who use AdaBoost), deriving features to drive a subsequent voxel-wise segmentation method, based for example on level sets (Gholipour et al., 2012; Schreibmann et al., 2014), random forests (Han, 2013), support vector machines (Hao et al., 2014), patch-based techniques (Wang et al., 2014e), or a graph-cut-based method (Candemir et al., 2014; Lee et al., 2014b). Along a similar direction, one can apply a refinement to the MAS results, for example, by comparing the observed intensities in the novel image to tissue-based expected intensity profiles (Ledig et al., 2014). Alternatively, label fusion results have been used to compute priors in probabilistic segmentation algorithms (Fortunati et al., 2013; Shi et al., 2010; van der Lijn

et al., 2008; Van Der Lijn et al., 2012; Wang et al., 2014c; Wolz et al., 2009, 2010b; Xu et al., 2014b; Platero et al., 2014; Makropoulos et al., 2014; Yan et al., 2014). These methods can be robust to changes in image appearance, for example, in applications where there is significant variation in imaging parameters or the novel subject’s anatomy is not represented in the atlases. However, they will be less accurate than standard MAS methods when the intensity profiles and appearance distribution are well matched between the novel image and atlases. In a related effort, Liu et al. (2014) used MAS to define a prior for the detection of lymph nodes in thoracic CT scans. A different strategy is to examine summary measurements (e.g., volume of an ROI) computed from the MAS to statistically determine whether the segmentation result is an outlier and thus might have failed – in which case one can resort to manual delineation (van Rikxoort et al., 2009).

There are also methods that operate on the posterior probability map obtained from label fusion, rather than applying a hard threshold to obtain a segmentation. For example, applying a deconvolution to the probability map has been shown to reduce the spatial bias in the segmentation of convex structures (Wang and Yushkevich, 2012b). In the context of neointima (scar tissue) segmentation in coronary optical coherence tomography, Tung et al. (2013) augment the posterior probability with an anatomically-informed probability, defined upon the distance to the vessel wall. While this prior knowledge enhances the performance of the method, it is highly domain specific and not applicable to other problems. In a parallel approach, Asman et al. (2013) propose to analyze the posterior probabilities to detect outliers that are not well represented in the atlas set. This is shown to be beneficial in the presence of anomalous regions (e.g., tumors).

3. Survey of Applications

Since its original application to confocal microscopy of bee brains (Rohlfing et al., 2004; Rohlfing and Maurer Jr, 2005), MAS has been successfully used in a large variety of biomedical segmentation problems. The most prevalent field of application has been brain MRI analysis, for two different reasons; first, segmentation’s crucial role in a wide range of widely studied neuroimaging problems; and second, the success of image registration techniques in this field.

Most of the MAS work applied to brain MRI data has focused on the segmentation of cortical and subcortical regions in structural images, typically acquired with T1-weighted MRI sequences. Many methods have been developed to parcellate the whole brain, segmenting it into a large number of regions (Aljabar et al., 2008; Babalola et al., 2009; Fonov et al., 2012; Han et al., 2009; Hecke-mann et al., 2010, 2011; Keihaninejad et al., 2010; Kotrotsou et al., 2014; Svarer et al., 2005; Wang et al., 2012; Ledig et al., 2014), while other studies have focused on

small sets or individual ROIs, such as the caudate nucleus (van Rikxoort et al., 2007b); the cerebellum (Park et al., 2014; Van Der Lijn et al., 2012; Weier et al., 2014); the amygdala (Hanson et al., 2012; Klein-Koerkamp et al., 2014); the corpus callosum (Ardekani et al., 2014; Gao et al., 2014; Meyer, 2014); the striatum (Janes et al., 2014); the subthalamic nucleus, red nucleus and substantia nigra (Xiao et al., 2014b,a); the ventricles (Chou et al., 2008; Raamana et al., 2014); and, most notably, the hippocampus, which has attracted much attention due to its association with dementia and Alzheimer’s disease (Akhondi-Asl et al., 2010; Bishop et al., 2010; Clerx et al., 2013; Hammers et al., 2007; Iglesias et al., 2010; Kim et al., 2012; Leung et al., 2010; Pipitone et al., 2014; Pluta et al., 2012; Raamana et al., 2014; van der Lijn et al., 2008; Van Der Lijn et al., 2012; Winston et al., 2013; Wolz et al., 2010b; Yushkevich et al., 2010; Platero et al., 2014; Ta et al., 2014).

In the context of segmentation of structural human brain MRI, multi-atlas techniques have also been applied to preprocessing tasks such as skull stripping (Leung et al., 2011; Weisenfeld and Warfield, 2011b) and tissue classification (Bouix et al., 2007; Crum, 2009), the segmentation of tumors (Zikic et al., 2014a; Wang and Yushkevich, 2013b; Warfield et al., 2004), eyes and optic nerves (Datteri et al., 2014; Harrigan et al., 2014). MAS has also been employed for the segmentation of cortical and sub-cortical structures in MRI data from fetuses, neonates, and infants too (Gholipour et al., 2012; Gousias et al., 2008, 2010, 2013; Shi et al., 2010; Wang et al., 2014e; Li et al., 2014; Koch et al., 2014; Makropoulos et al., 2014; Wang et al., 2014d), in which the contrast inversion due to ongoing myelination complicates the segmentation. Another area of application of MAS has been the segmentation of brain MRI in animal studies, e.g., mice (Da et al., 2012; Ma et al., 2014; Nie and Shen, 2013; Lee et al., 2014a; Khan et al., 2014), rats (Lancelot et al., 2014) and non-human primates (Ballanger et al., 2013). Finally, there are also studies that have applied MAS to the analysis of diffusion brain MRI data of humans (Jin et al., 2012; Tang et al., 2014; Traynor et al., 2010), which requires specific strategies for the registration, atlas selection and label fusion steps, due to the nature of the data, which are typically described by directional functions defined on the sphere at each voxel.

Outside brain imaging, the prevalence of prostate cancer in men has sparked interest in applications within prostate imaging, using modalities such as MRI (Langerak et al., 2010; Litjens et al., 2014; Rivest-Hénault et al., 2014), CT (Acosta et al., 2011; Sjöberg et al., 2013; Acosta et al., 2014) and ultrasound (Nouranian et al., 2014). Likewise, interest in radiotherapy treatment planning has been the main driver of applications in head, neck, and thoracic CT segmentation (Han et al., 2008; Wang et al., 2014c), which have mainly focused on segmenting tumors (Ramus and Malandain, 2010), organs at risk (e.g. the parotid glands, Fritscher et al. 2014; Gorthi et al. 2010; Han et al.

2010; Hollensen et al. 2010; Yang et al. 2010 or mediastinal lymph nodes, Liu et al. 2014) and lymph node metastases (Sjöberg et al., 2013; Teng et al., 2010). MAS has also been used in abdominal imaging, despite the relatively poor performance of image registration in this domain (e.g., compared with brain MRI) due to the shifting of organs within the abdominal cavity. Nonetheless, MAS has been successful in liver (van Rikxoort et al., 2007a; Platero and Tobar, 2014), spleen (Li et al., 2013; Xu et al., 2014b) and multi-organ segmentation (Wolz et al., 2013; Schreibmann et al., 2014) in CT scans.

Finally, there are many other applications that have benefited from MAS within human medical imaging, including: segmentation of pelvic bones in MRI (Weisenfeld and Warfield, 2011b; Akhondi-Asl et al., 2014); lungs in CT scans (van Rikxoort et al., 2009) and chest X-rays (Candemir et al., 2014); heart and its ventricles in CT (van Rikxoort et al., 2010; Dey et al., 2010), MRI (Zhuang et al., 2010; Zuluaga et al., 2014), MR angiography (Wachinger and Golland, 2012), ultrasound (Wang et al., 2014a), and CT angiography (Kirişli et al., 2010; Yang et al., 2014a); breast tissues and lesions in X-ray mammography (Iglesias and Karssemeijer, 2009) and MRI (Gubern-Mérida et al., 2012; Lee et al., 2013); cartilage and bone in knee MRI (Tamez-Pena et al., 2012; Lee et al., 2014b; Shan et al., 2014); the vertebrae in spinal MRI (Asman et al., 2014); scar tissue in intravascular coronary optical coherence tomography (OCT) (Tung et al., 2013); the mitral valve in transesophageal echocardiography (Wang et al., 2013a; Pouch et al., 2014); skeletal muscle in whole-body MRI (Karlsson et al., 2014); kidneys in CT images (Yang et al., 2014b); and bone in dental cone-beam CT images (Wang et al., 2014c).

4. Discussion and Future Directions

By taking full advantage of the entire training data, rather than a model-based summary, MAS delivers highly accurate segmentation algorithms. This approach has come a long way since the early days of “majority voting”, which basically consisted of three independent steps: registration, label propagation, and fusion. Today, most MAS algorithms have many more steps, some of which form feedback loops. Furthermore, each one of these steps is becoming increasingly more sophisticated, employing ideas from optimization, computer vision, machine learning, probabilistic modeling, and other fields.

The biggest shortcoming of MAS is its ravenous appetite for computational resources. Analyzing, manipulating, and processing all atlases typically demands a substantial amount of memory and time. We believe this is one of the main reasons why MAS has not been widely adopted in clinical applications yet, even though, research suggests that it can produce state of the art segmentation tools in many domains. However we expect that several recent developments alleviate the computational challenges

of MAS. Firstly, the continued exponential growth in computer hardware technologies is to our advantage. We note, however, this technological benefit is to some extent countered by the rapidly increasing resolution of biomedical images, which multiplies the computational burden. Secondly, we observe that many of the subcomponents of MAS are parallelizable and thus can take advantage of multi-core architectures and GPUs. At the coarsest level, the registrations that need to be computed with each atlas can be solved in parallel. Furthermore, each registration can be implemented such that the bulk of the voxel- or region-level computations can be distributed over multiple processors. This approach has already been used for the GPU-acceleration of the registration step (Cardoso et al., 2013; Duc et al., 2013; Han et al., 2009; Modat et al., 2010). A similar strategy can be adopted in the label fusion step, particularly by algorithms that conduct numerical optimization in label fusion and not just simple counting. Finally, some of the online computational burden can be shifted offline, via learning structure on the training images, which can then be utilized to optimize the processing of the novel image, as proposed in (Jia et al., 2012).

The manually delineated training data form the main foundation of atlas-based segmentation. Empirical evidence suggests that the number and quality of training cases can critically impact segmentation accuracy. For this reason, the careful definition and standardization of annotation guidelines is paramount to obtaining accurate automatic segmentations, especially when the atlases are manually delineated by multiple experts. Yet, obtaining high quality segmentations annotated by experts is both time consuming and expensive. Most past research has dealt with scenarios where the development of the segmentation algorithm is independent of the manual segmentation process. We believe a better strategy is to integrate the two pipelines. For example, as recently demonstrated (Awate and Whitaker, 2014), given a segmentation method, one can estimate the number of cases that need to be manually delineated to achieve a desired level of accuracy.

Furthermore, one can imagine an algorithm that indicates the cases, which, if manually segmented, assist the segmentation algorithm the most. Active learning can provide the framework to derive such an algorithm. An alternative approach is to use automatic segmentations as atlases, after applying a quality control step. Yet a different strategy is to harness the potential of non-expert segmenters (Bogovic et al., 2013; Bryan et al., 2014), for example, via a crowd-sourcing framework (Landman et al., 2012a; Maier-Hein et al., 2014). Although many biomedical segmentation problems rely on anatomical expertise, it is not clear whether this expertise has to be deployed in the delineation of every single atlas. One can imagine certain scenarios, where the expert(s) provides a handful of example annotations, which can be used to train or guide non-experts. Finally, we believe that the idea to combine heterogeneous sets of atlases, delineated with different pro-

ocols, is a promising future direction. This strategy can both yield better accuracy by enriching the training data and offer the ability to identify ROIs that were technically not part of any single manual delineation protocol but can be defined by intersections (Iglesias et al., 2015). Moreover, such an approach could also potentially minimize the impact of the variability of the manual delineations (within or across experts) on the final segmentation, automatically learning the biases of the annotations.

Crowdsourcing offers another attractive solution to the atlas generation problem of MAS. Instead of high quality manual delineations from trained experts, one might consider using lower quality data from non-experts (Maier-Hein et al., 2014). Alternatively, the non-expert crowd can be used to correct or filter the segmentations. We expect that outsourcing certain aspects of MAS, particularly those related to the offline stages of the pipeline, to non-expert and/or expert masses in an online community will be investigated in the near future.

While speeding up the registration step might be considered top priority for some applications, many biomedical problems seek very high accuracy, even at high computational cost. For such applications, one strategy is to improve registration accuracy and the quality of propagated labels. The probabilistic modeling perspective offers a complementary approach. From this viewpoint, registration is a nuisance parameter and thus should be marginalized out, e.g., via variational techniques (Simpson et al., 2011) or a sampling procedure such as Markov Chain Monte Carlo (Iglesias et al., 2013a). In other words, one should integrate over all possible registration results, rather than attempting to find the most likely one and using that for the fusion step. Note that this approach would be different from the latest algorithms that combine the registration and label fusion steps, as done in (Hao et al., 2012; Iglesias et al., 2013c; Tang et al., 2013). Currently, the marginalization strategy might seem computationally prohibitive for MAS. However, the recent successful applications of this idea in other biomedical image analysis scenarios suggest that in the near future we can expect to see label fusion algorithms that integrate out the unknown registrations.

Another direction of future work in MAS is to develop algorithms that are robust against changes in image intensity profiles, e.g., MRI contrast, due to variation in acquisition protocols, hardware, and other imaging parameters. Such robust methods will be invaluable for large-scale multi-site studies and clinical applications, where the standardization of the imaging protocol might be unrealistic. Although some existing label fusion algorithms (e.g., Iglesias et al. 2012a, 2013b) have been developed to handle different modalities, they are application-specific and make strong assumptions about the data (e.g., locally Gaussian intensity distributions).

We believe that the fields of machine learning and computer vision have also a lot to contribute to MAS. Recent years have witnessed dramatic technical advances in

both of these fields, such as unsupervised feature learning in vision (Erhan et al., 2010) and efficient learning methods on deep architectures (Hinton and Salakhutdinov, 2006), which have facilitated tremendous gains in performance. Recent developments suggest that researchers are currently working on translating such ideas to biomedical image analysis problems, including MAS.

The probabilistic perspective, with its flexibility and principled inference machinery, offers another promising direction for future research. In particular, this approach enables the derivation of methods that can handle missing labels, heterogeneous labels, variable imaging modalities, estimate and utilize model uncertainty, and integrate domain knowledge, for example about the anatomy or imaging physics. Furthermore, probabilistic algorithms offer the capability to quantify the uncertainty in the final segmentation estimate, which can further be utilized for obtaining more accurate measurements, for example of the volume of structures (Iglesias et al., 2013a).

Rather than segmenting each novel image independently, empirical evidence suggests that solving the segmentations of multiple novel images simultaneously might yield improved results (Wang et al., 2012). This can be a particularly promising approach for segmenting serial scans. Longitudinal image analysis is an area of growing importance and the detection of subtle longitudinal changes can call for highly accurate segmentation (Reuter et al., 2012). Encouraged by some recent applications (Wolz et al. 2010b and Li et al. 2014), we believe MAS will be a critical tool for longitudinal biomedical image analysis.

So far, most of the applications of MAS have been in the domain of human brain MRI, in which modern registration algorithms achieve good alignment and even the simplest fusion algorithms (e.g., majority voting) yield good performance. Registration is however less effective in other modalities and body parts, such as in abdominal imaging, in which the sliding between organ walls (e.g., due to respiratory motion) is problematic for current algorithms. We believe, though, that the development of registration methods that can cope with these difficulties, along with the improvements in label fusion techniques (which will make them more robust against misregistration), will make the use of the multi-atlas approach ubiquitous in a growing number of novel biomedical image segmentation problems.

Finally, it is important to note that there is no *universally* optimal segmentation algorithm. Each application brings with it a unique set of constraints and objectives, making certain types of methods more suitable than others. Yet, we believe that the large class of MAS methods, with its rich set of instantiations that enable compromising between different application tradeoffs and considering various objectives while exploiting different sources of information, offer a framework that promises to yield effective and useful solutions for a wide range of biomedical applications. That said, we can identify general trends that have emerged. There seems to be a global tradeoff

between computational efficiency and segmentation accuracy. Incorporating domain knowledge and adopting realistic models that are based on the physical and biological context, can yield significant improvements. Complex, more advanced methods can pay off and should be something we continue to work on. However, this endeavor critically depends on a proper evaluation of the methods, as demonstrated in some recent efforts (Rueda et al., 2014; Menze et al., 2014; Panda et al., 2014; Goksel et al., 2014). Going forward, a grand challenge of biomedical image segmentation will be to establish standardized datasets and performance evaluation metrics to be used to objectively compare various segmentation algorithms, including MAS-based techniques.

Acknowledgment

Juan Eugenio Iglesias is supported by the Gipuzkoako Foru Aldundia (Fellows Gipuzkoa Program). Mert R. Sabuncu is supported by NIH NIBIB 1K25EB013649-01 and a Bright-Focus Alzheimer’s disease pilot research grant (AHAF-A2012333).

References

- Acosta, O., Dowling, J., Drean, G., Simon, A., De Crevoisier, R., Haigron, P., 2014. Multi-atlas-based segmentation of pelvic structures from CT scans for planning in prostate cancer radiotherapy, in: *Abdomen and Thoracic Imaging*. Springer, pp. 623–656.
- Acosta, O., Simon, A., Monge, F., Commandeur, F., Bassirou, C., Cazoulat, G., De Crevoisier, R., Haigron, P., 2011. Evaluation of multi-atlas-based segmentation of CT scans in prostate cancer radiotherapy, in: *Biomedical Imaging: From Nano to Macro*, 2011 IEEE International Symposium on, IEEE. pp. 1966–1969.
- Akhondi-Asl, A., Hoyte, L., Lockhart, M., Warfield, S., 2014. A logarithmic opinion pool based staple algorithm for the fusion of segmentations with associated reliability weights. *IEEE Transactions on Medical Imaging*.
- Akhondi-Asl, A., Jafari-Khouzani, K., Soltanian-Zadeh, H., 2010. Multiple-atlas-based automatic segmentation of hippocampus for lateralization in temporal lobe epilepsy, in: *Biomedical Imaging: From Nano to Macro*, 2010 IEEE International Symposium on, IEEE. pp. 836–839.
- Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J.V., Rueckert, D., 2007. Classifier selection strategies for label fusion using large atlas databases, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*. Springer, pp. 523–531.
- Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 46, 726–738.
- Aljabar, P., Rueckert, D., Crum, W.R., 2008. Automated morphological analysis of magnetic resonance brain imaging using spectral analysis. *Neuroimage* 43, 225–235.
- Ardekani, B.A., Bachman, A.H., Figarsky, K., Sidtis, J.J., 2014. Corpus callosum shape changes in early Alzheimer’s disease: an MRI study using the OASIS brain database. *Brain Structure and Function* 219, 343–352.
- Aribisala, B.S., Cox, S.R., Ferguson, K.J., MacPherson, S.E., MacLulich, A.M., Royle, N.A., Hernández, M.V., Bastin, M.E., Deary, I.J., Wardlaw, J.M., 2013. Assessing the performance of atlas-based prefrontal brain parcellation in an ageing cohort. *Journal of computer assisted tomography* 37.
- Artachevarria, X., Muñoz-Barrutia, A., Ortiz-de Solorzano, C., 2008. Efficient classifier generation and weighted voting for atlas-based segmentation: Two small steps faster and closer to the

- combination oracle, in: Medical Imaging, International Society for Optics and Photonics. pp. 69141W–69141W.
- Arteachevarria, X., Munoz-Barrutia, A., Ortiz-de Solorzano, C., 2009. Combination strategies in multi-atlas image segmentation: Application to brain MR data. Medical Imaging, IEEE Transactions on 28, 1266–1277.
- Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry: the methods. Neuroimage 11, 805–821.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851.
- Asman, A.J., Bryan, F.W., Smith, S.A., Reich, D.S., Landman, B.A., 2014. Groupwise multi-atlas segmentation of the spinal cord's internal structure. Medical image analysis 18, 460–471.
- Asman, A.J., Chambless, L.B., Thompson, R.C., Landman, B.A., 2013. Out-of-atlas likelihood estimation using multi-atlas segmentation. Medical physics 40, 043702.
- Asman, A.J., Landman, B.A., 2011. Robust statistical label fusion through consensus level, labeler accuracy, and truth estimation (COLLATE). Medical Imaging, IEEE Transactions on 30, 1779–1794.
- Asman, A.J., Landman, B.A., 2012. Formulating spatially varying performance in the statistical fusion framework. Medical Imaging, IEEE Transactions on 31, 1326–1336.
- Asman, A.J., Landman, B.A., 2013. Non-local statistical label fusion for multi-atlas segmentation. Medical image analysis 17, 194–208.
- Asman, A.J., Landman, B.A., 2014. Hierarchical performance estimation in the statistical label fusion framework. Medical image analysis 18, 1070–1081.
- Avants, B.B., Tustison, N., Song, G., 2009. Advanced normalization tools (ANTS). Insight J.
- Avants, B.B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., Gee, J.C., 2010. The optimal template effect in hippocampus studies of diseased populations. Neuroimage 49, 2457–2466.
- Awate, S.P., Whitaker, R.T., 2014. Multiatlas segmentation as non-parametric regression. IEEE Transactions on Medical Imaging.
- Babalola, K.O., Patenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., Smith, S., Cootes, T., Jenkinson, M., Rueckert, D., 2009. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. Neuroimage 47, 1435–1447.
- Bai, J., Trinh, T.L.H., Chuang, K.H., Qiu, A., 2012. Atlas-based automatic mouse brain image segmentation revisited: model complexity vs. image registration. Magnetic resonance imaging 30, 789–798.
- Bai, W., Shi, W., O'Regan, D.P., Tong, T., Wang, H., Jamil-Copley, S., Peters, N.S., Rueckert, D., 2013. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: application to cardiac MR images. Medical Imaging, IEEE Transactions on 32, 1302–1315.
- Ballanger, B., Tremblay, L., Sgambato-Faure, V., Beaudoin-Gobert, M., Lavenne, F., Le Bars, D., Costes, N., 2013. A multi-atlas based method for automated anatomical macaca fascicularis brain MRI segmentation and PET kinetic extraction. Neuroimage 77, 26–43.
- Beg, M.F., Miller, M.I., Trounev, A., Younes, L., 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International journal of computer vision 61, 139–157.
- Bishop, C.A., Jenkinson, M., Declerck, J., Merhof, D., 2010. Evaluation of hippocampal segmentation methods for healthy and pathological subjects, in: Proceedings of the 2nd Eurographics conference on Visual Computing for Biology and Medicine, Eurographics Association. pp. 17–24.
- Bogovic, J.A., Jedynek, B., Rigg, R., Du, A., Landman, B.A., Prince, J.L., Ying, S.H., 2013. Approaching expert results using a hierarchical cerebellum parcellation protocol for multiple inexperienced human raters. Neuroimage 64, 616–629.
- Bouix, S., Martin-Fernandez, M., Ungar, L., Nakamura, M., Koo, M.S., McCarley, R.W., Shenton, M.E., 2007. On evaluating brain tissue classifiers without a ground truth. Neuroimage 36, 1207–1224.
- Bryan, F.W., Xu, Z., Asman, A.J., Allen, W.M., Reich, D.S., Landman, B.A., 2014. Self-assessed performance improves statistical fusion of image labels. Medical physics 41, 031903.
- Buades, A., Coll, B., Morel, J.M., 2005. A non-local algorithm for image denoising, in: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, IEEE. pp. 60–65.
- Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Xue, Z., Karargyris, A., Antani, S., Thoma, G., McDonald, C.J., 2014. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. Medical Imaging, IEEE Transactions on 33, 577–590.
- Cao, Y., Li, X., Yan, P., 2012. Multi-atlas based image selection with label image constraint, in: Machine Learning and Applications (ICMLA), 2012 11th International Conference on, IEEE. pp. 311–316.
- Cao, Y., Yuan, Y., Li, X., Turkbey, B., Choyke, P.L., Yan, P., 2011a. Segmenting images by combining selected atlases on manifold, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011. Springer, pp. 272–279.
- Cao, Y., Yuan, Y., Li, X., Yan, P., 2011b. Putting images on a manifold for atlas-based image segmentation, in: Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE. pp. 289–292.
- Cardoso, M., Leung, K., Modat, M., Keihaninejad, S., Cash, D., Barnes, J., Fox, N.C., Ourselin, S., 2013. STEPS: similarity and truth estimation for propagated segmentations and its application to hippocampal segmentation and brain parcellation. Medical image analysis 17, 671–684.
- Chakravarty, M.M., Steadman, P., Eede, M.C., Calcott, R.D., Gu, V., Shaw, P., Raznahan, A., Collins, D.L., Lerch, J.P., 2013. Performing label-fusion-based segmentation using multiple automatically generated templates. Human brain mapping 34, 2635–2654.
- Chou, Y.Y., Lepore, N., de Zubicaray, G.I., Carmichael, O.T., Becker, J.T., Toga, A.W., Thompson, P.M., 2008. Automated ventricular mapping with multi-atlas fluid image alignment reveals genetic effects in Alzheimer's disease. Neuroimage 40, 615–630.
- Christensen, G.E., Joshi, S.C., Miller, M.I., 1997. Volumetric transformation of brain anatomy. Medical Imaging, IEEE Transactions on 16, 864–877.
- Clerx, L., van Rossum, I.A., Burns, L., Knol, D.L., Scheltens, P., Verhey, F., Aalten, P., Lapuerta, P., van de Pol, L., van Schijndel, R., et al., 2013. Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. Neurobiology of aging 34, 2003–2013.
- Collins, D.L., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. Human brain mapping 3, 190–208.
- Commowick, O., Akhondi-Asl, A., Warfield, S.K., 2012. Estimating a reference standard segmentation with spatially varying performance parameters: Local MAP STAPLE. Medical Imaging, IEEE Transactions on 31, 1593–1606.
- Commowick, O., Malandain, G., 2007. Efficient selection of the most similar image in a database for critical structures segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. Springer, pp. 203–210.
- Commowick, O., Warfield, S.K., 2010. Incorporating priors on expert performance parameters for segmentation validation and label fusion: a maximum a posteriori STAPLE, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010. Springer, pp. 25–32.
- Commowick, O., Warfield, S.K., Malandain, G., 2009. Using Frankenstein's creature paradigm to build a patient specific atlas, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009. Springer, pp. 993–1000.
- Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L., 2011. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. Neuroimage 54, 940–954.
- Criminisi, A., Sharp, T., Blake, A., 2008. Geos: Geodesic image segmentation, in: Computer Vision–ECCV 2008. Springer, pp. 99–112.
- Crum, W.R., 2009. Spectral clustering and label fusion for 3D tissue

- classification: Sensitivity and consistency analysis. *Annals of the BMVA* 2009, 1–12.
- Da, M., Cardoso, M.J., Modat, M., Powerll, N., Holmes, H., Lyhgoe, M., Ourselin, S., 2012. Multi atlas segmentation applied to in vivo mouse brain MRI, in: *Workshop on Multi-atlas Labeling*, pp. 363–374.
- Datteri, R., Asman, A., Landman, B., Dawant, B., 2011. Estimation of registration accuracy applied to multiatlas segmentation, in: *Proceedings of the 2011 MICCAI Workshop on Multi-Atlas Labeling and Statistical Fusion*, pp. 78–87.
- Datteri, R., Asman, A.J., Landman, B.A., Dawant, B.M., 2014. Applying the algorithm assessing quality using image registration circuits (aquirc) to multi-atlas segmentation, in: *SPIE Medical Imaging, International Society for Optics and Photonics*. pp. 90341F–90341F.
- Davatzikos, C., 1996. Spatial normalization of 3D brain images using deformable models. *Journal of computer assisted tomography* 20, 656–665.
- Dawant, B.M., Hartmann, S.L., Thirion, J.P., Maes, F., Vandermeulen, D., Demaerel, P., 1999. Automatic 3-D segmentation of internal structures of the head in 3D MR images using a combination of similarity and free-form transformations i: Methodology and validation on normal subjects. *Medical Imaging, IEEE Transactions on* 18, 909–916.
- Depa, M., Holmvang, G., Schmidt, E.J., Golland, P., Sabuncu, M.R., 2011. Towards efficient label fusion by pre-alignment of training data, in: *Proceedings of the MICCAI Workshop on Multi-Atlas Labeling and Statistical Fusion*, pp. 38–46.
- Depa, M., Sabuncu, M.R., Holmvang, G., Nezafat, R., Schmidt, E.J., Golland, P., 2010. Robust atlas-based segmentation of highly variable anatomy: Left atrium segmentation, in: *Statistical Atlases and Computational Models of the Heart*. Springer, pp. 85–94.
- Dey, D., Ramesh, A., Slomka, P.J., Nakazato, R., Cheng, V.Y., Germano, G., Berman, D.S., 2010. Automated algorithm for atlas-based segmentation of the heart and pericardium from non-contrast CT, in: *SPIE Medical Imaging, International Society for Optics and Photonics*. pp. 762337–762337.
- Doan, N.T., de Xivry, J.O., Macq, B., 2010. Effect of inter-subject variation on the accuracy of atlas-based segmentation applied to human brain structures, in: *SPIE Medical Imaging, International Society for Optics and Photonics*. pp. 76231S–76231S.
- Doshi, J., Erus, G., Ou, Y., Gaonkar, B., Davatzikos, C., 2013. Multi-atlas skull-stripping. *Academic radiology* 20, 1566–1576.
- Duc, A.K.H., Modat, M., Leung, K.K., Cardoso, M.J., Barnes, J., Kadir, T., Ourselin, S., Initiative, A.D.N., et al., 2013. Using manifold learning for atlas selection in multi-atlas segmentation. *PloS one* 8, e70059.
- Erhan, D., Bengio, Y., Courville, A., Manzagol, P.A., Vincent, P., Bengio, S., 2010. Why does unsupervised pre-training help deep learning? *The Journal of Machine Learning Research* 11, 625–660.
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klavenness, S., et al., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.
- Fonov, V., Coupé, P., Eskildsen, S.F., Manjon, J.V., Collins, L., et al., 2012. Multi-atlas labeling with population-specific template and non-local patch-based label fusion, in: *MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling*, pp. 63–66.
- Fortunati, V., Verhaart, R.F., van der Lijn, F., Niessen, W.J., Veenland, J.F., Paulides, M.M., van Walsum, T., 2013. Tissue segmentation of head and neck CT images for treatment planning: A multiatlas approach combined with intensity modeling. *Medical physics* 40, 071905.
- Fritscher, K.D., Peroni, M., Zaffino, P., Spadea, M.F., Schubert, R., Sharp, G., 2014. Automatic segmentation of head and neck CT images for radiotherapy treatment planning using multiple atlases, statistical appearance models, and geodesic active contours. *Medical Physics* 41, 051910.
- Gao, Q., Tong, T., Rueckert, D., et al., 2014. Multi-atlas propagation via a manifold graph on a database of both labeled and unlabeled images, in: *SPIE Medical Imaging, International Society for Optics and Photonics*. pp. 90350A–90350A.
- Gass, T., Székely, G., Goksel, O., 2013. Semi-supervised segmentation using multiple segmentation hypotheses from a single atlas, in: *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. Springer, pp. 29–37.
- Gholipour, A., Akhondi-Asl, A., Estroff, J.A., Warfield, S.K., 2012. Multi-atlas multi-shape segmentation of fetal brain MRI for volumetric and morphometric analysis of ventriculomegaly. *Neuroimage* 60, 1819–1831.
- Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N., 2008. Dense image registration through MRFs and efficient linear programming. *Medical image analysis* 12, 731–741.
- Goksel, O., Gass, T., Székely, G., 2014. Segmentation and landmark localization based on multiple atlases, in: *Proceedings of the VISCERAL Challenge at ISBI. CEUR Workshop Proceedings*, Beijing, China, pp. 37–43.
- Gorthi, S., Akhondi-Asl, A., Thiran, J.P., Warfield, S.K., 2014. Optimal map parameters estimation in staple-learning from performance parameters versus image similarity information, in: *Machine Learning in Medical Imaging*. Springer, pp. 174–181.
- Gorthi, S., Bach Cuadra, M., Schick, U., Tercier, P.A., Allal, A.S., Thiran, J.P., 2010. Multi-atlas based segmentation of head and neck CT images using active contour framework, in: *MICCAI workshop on 3D Segmentation Challenge for Clinical Applications*, pp. 313–321.
- Gorthi, S., Bach Cuadra, M., Tercier, P.A., Allal, A.S., Thiran, J.P., 2013. Weighted shape-based averaging with neighborhood prior model for multiple atlas fusion-based medical image segmentation. *IEEE Signal Processing Letters* 20.
- Gousias, I., Hammers, A., Heckemann, R., Counsell, S., Dyet, L., Boardman, J., Edwards, A., Rueckert, D., 2010. Atlas selection strategy for automatic segmentation of pediatric brain MRIs into 83 rois, in: *Imaging Systems and Techniques (IST), 2010 IEEE International Conference on, IEEE*. pp. 290–293.
- Gousias, I.S., Hammers, A., Counsell, S.J., Srinivasan, L., Rutherford, M.A., Heckemann, R.A., Hajnal, J.V., Rueckert, D., Edwards, A.D., 2013. Magnetic resonance imaging of the newborn brain: automatic segmentation of brain images into 50 anatomical regions. *PloS one* 8, e59990.
- Gousias, I.S., Rueckert, D., Heckemann, R.A., Dyet, L.E., Boardman, J.P., Edwards, A.D., Hammers, A., 2008. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. *Neuroimage* 40, 672–684.
- Gubern-Mérida, A., Kallenberg, M., Martí, R., Karssemeijer, N., 2012. Segmentation of the pectoral muscle in breast MRI using atlas-based approaches, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*. Springer, pp. 371–378.
- Hammers, A., Heckemann, R., Koepp, M.J., Duncan, J.S., Hajnal, J.V., Rueckert, D., Aljabar, P., 2007. Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: a proof-of-principle study. *Neuroimage* 36, 38–47.
- Han, X., 2013. Learning-boosted label fusion for multi-atlas auto-segmentation, in: *Machine Learning in Medical Imaging*. Springer, pp. 17–24.
- Han, X., Hibbard, L.S., O’connell, N.P., Willcut, V., 2010. Automatic segmentation of parotids in head and neck CT images using multi-atlas fusion. *Medical Image Analysis for the Clinic: A Grand Challenge*, 297–304.
- Han, X., Hibbard, L.S., Willcut, V., 2009. GPU-accelerated, gradient-free mi deformable registration for atlas-based MR brain image segmentation, in: *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on, IEEE*. pp. 141–148.
- Han, X., Hoogeman, M.S., Levendag, P.C., Hibbard, L.S., Teguh, D.N., Voet, P., Cowen, A.C., Wolf, T.K., 2008. Atlas-based auto-segmentation of head and neck CT images, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*. Springer, pp. 434–441.
- Hanson, J.L., Suh, J.W., Nacewicz, B.M., Sutterer, M.J., Cayo,

- A.A., Stodola, D.E., Burghy, C.A., Wang, H., Avants, B.B., Yushkevich, P.A., et al., 2012. Robust automated amygdala segmentation via multi-atlas diffeomorphic registration. *Frontiers in neuroscience* 6.
- Hao, Y., Jiang, T., Fan, Y., 2012. Iterative multi-atlas based segmentation with multi-channel image registration and jackknife context model, in: *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE*. pp. 900–903.
- Hao, Y., Wang, T., Zhang, X., Duan, Y., Yu, C., Jiang, T., Fan, Y., 2014. Local label learning (LLL) for subcortical structure segmentation: Application to hippocampus segmentation. *Human brain mapping* 35, 2674–2697.
- Harrigan, R.L., Panda, S., Asman, A.J., Nelson, K.M., Chaganti, S., DeLisi, M.P., Yvernault, B.C., Smith, S.A., Galloway, R.L., Mawn, L.A., et al., 2014. Robust optic nerve segmentation on clinically acquired computed tomography. *Journal of Medical Imaging* 1, 034006–034006.
- Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A., 2006. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *Neuroimage* 33, 115–126.
- Heckemann, R.A., Hammers, A., Aljabar, P., Rueckert, D., Hajnal, J.V., 2009. The mirror method of assessing segmentation quality in atlas label propagation, in: *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on, IEEE*. pp. 1194–1197.
- Heckemann, R.A., Keihaninejad, S., Aljabar, P., Gray, K.R., Nielsen, C., Rueckert, D., Hajnal, J.V., Hammers, A., 2011. Automatic morphometry in alzheimer's disease and mild cognitive impairment. *Neuroimage* 56, 2024–2037.
- Heckemann, R.A., Keihaninejad, S., Aljabar, P., Rueckert, D., Hajnal, J.V., Hammers, A., 2010. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *Neuroimage* 51, 221–227.
- Heiberg, E., Sjögren, J., Ugander, M., Carlsson, M., Engblom, H., Arheden, H., 2010. Design and validation of segment-freely available software for cardiovascular image analysis. *BMC medical imaging* 10, 1.
- Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. *Science* 313, 504–507.
- Hollensen, C., Hansen, M.F., Højgaard, L., Specht, L., Larsen, R., 2010. Segmenting the parotid gland using registration and level set methods. *MICCAI: Grand Challenges in Medical Image Analysis: Head & Neck Autosegmentation Challenge*, Beijing.
- Iglesias, J.E., Dinov, I., Singh, J., Tong, G., Tu, Z., 2010. Synthetic MRI signal standardization: application to multi-atlas analysis, in: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010. Springer*, pp. 81–88.
- Iglesias, J.E., Karssemeijer, N., 2009. Robust initial detection of landmarks in film-screen mammograms using multiple FFDM atlases. *Medical Imaging, IEEE Transactions on* 28, 1815–1824.
- Iglesias, J.E., Sabuncu, M.R., Aganj, I., Bhatt, P., Casillas, C., Salat, D., Boxer, A., Fischl, B., Van Leemput, K., 2015. An algorithm for optimal fusion of atlases with different labeling protocols. *NeuroImage* 106, 451–463.
- Iglesias, J.E., Sabuncu, M.R., Van Leemput, K., 2012a. A generative model for multi-atlas segmentation across modalities, in: *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE*. pp. 888–891.
- Iglesias, J.E., Sabuncu, M.R., Van Leemput, K., 2012b. A generative model for probabilistic label fusion of multimodal data, in: *Multimodal Brain Image Analysis. Springer*, pp. 115–133.
- Iglesias, J.E., Sabuncu, M.R., Van Leemput, K., 2013a. Improved inference in bayesian segmentation using monte carlo sampling: Application to hippocampal subfield volumetry. *Medical image analysis* 17, 766–778.
- Iglesias, J.E., Sabuncu, M.R., Van Leemput, K., 2013b. A probabilistic, non-parametric framework for inter-modality label fusion, in: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013. Springer*, pp. 576–583.
- Iglesias, J.E., Sabuncu, M.R., Van Leemput, K., 2013c. A unified framework for cross-modality multi-atlas segmentation of brain MRI. *Medical image analysis* 17, 1181–1191.
- Isgum, I., Staring, M., Rutten, A., Prokop, M., Viergever, M.A., van Ginneken, B., 2009. Multi-atlas-based segmentation with local decision fusion: Application to cardiac and aortic segmentation in CT scans. *Medical Imaging, IEEE Transactions on* 28, 1000–1010.
- Janes, A.C., Park, M.T.M., Farmer, S., Chakravarty, M.M., 2014. Striatal morphology is associated with tobacco cigarette craving. *Neuropsychopharmacology*, 406–411.
- Jia, H., Yap, P.T., Shen, D., 2012. Iterative multi-atlas-based multi-image segmentation with tree-based registration. *Neuroimage* 59, 422–430.
- Jin, Y., Shi, Y., Zhan, L., Li, J., de Zubicaray, G.I., McMahon, K.L., Martin, N.G., Wright, M.J., Thompson, P.M., 2012. Automatic population HARDI white matter tract clustering by label fusion of multiple tract atlases, in: *Multimodal Brain Image Analysis. Springer*, pp. 147–156.
- Kapur, T., Grimson, W.E.L., Wells III, W.M., Kikinis, R., 1996. Segmentation of brain tissue from magnetic resonance images. *Medical image analysis* 1, 109–127.
- Karlsson, A., Rosander, J., Romu, T., Tallberg, J., Grönqvist, A., Borga, M., Dahlqvist Leinhard, O., 2014. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. *Journal of Magnetic Resonance Imaging*.
- Kasiri, K., Fieguth, P., Clausi, D., et al., 2014. Cross modality label fusion in multi-atlas segmentation, in: *Image Processing (ICIP), 2014 IEEE International Conference on, IEEE*. pp. 16–20.
- Keihaninejad, S., Heckemann, R., Gousias, I.S., Aljabar, P., Hajnal, J.V., Rueckert, D., Hammers, A., 2010. Automatic volumetry can reveal visually undetected disease features on brain MR images in temporal lobe epilepsy, in: *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on, IEEE*. pp. 105–108.
- Khan, A.R., Cherbuin, N., Wen, W., Anstey, K.J., Sachdev, P., Beg, M.F., 2011. Optimal weights for local multi-atlas fusion using supervised learning and dynamic information (SuperDyn): Validation on hippocampus segmentation. *Neuroimage* 56, 126–139.
- Khan, U.A., Liu, L., Provenzano, F.A., Berman, D.E., Profaci, C.P., Sloan, R., Mayeux, R., Duff, K.E., Small, S.A., 2014. Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical alzheimer's disease. *Nature neuroscience* 17, 304–311.
- Kim, H., Chupin, M., Colliot, O., Bernhardt, B.C., Bernasconi, N., Bernasconi, A., 2012. Automatic hippocampal segmentation in temporal lobe epilepsy: impact of developmental abnormalities. *Neuroimage* 59, 3178–3186.
- Kirish, H., Schaap, M., Klein, S., Papadopoulou, S., Bonardi, M., Chen, C., Weustink, A., Mollnet, N., Vonken, E., van der Geest, R., et al., 2010. Evaluation of a multi-atlas based method for segmentation of cardiac CTA data: a large-scale, multicenter, and multivendor study. *Medical physics* 37, 6279–6291.
- Klein, A., Mensh, B., Ghosh, S., Tourville, J., Hirsch, J., 2005. Mind-boggle: automated brain labeling with multiple atlases. *BMC medical imaging* 5, 7.
- Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P., 2010. Elastix: a toolbox for intensity-based medical image registration. *Medical Imaging, IEEE Transactions on* 29, 196–205.
- Klein-Koerkamp, Y., A Heckemann, R., T Ramdeen, K., Moreaud, O., Keignart, S., Krainik, A., Hammers, A., Baciuc, M., Hot, P., Initiative, A.D.N., et al., 2014. Amygdalar atrophy in early Alzheimer's disease. *Current Alzheimer Research* 11, 239–252.
- Koch, L.M., Wright, R., Vatansever, D., Kyriakopoulou, V., Malamateniou, C., Patkee, P.A., Rutherford, M., Hajnal, J.V., Aljabar, P., Rueckert, D., 2014. Graph-based label propagation in fetal brain MR images, in: *Machine Learning in Medical Imaging. Springer*, pp. 9–16.
- Konukoglu, E., Glocker, B., Zikic, D., Criminisi, A., 2013. Neighbourhood approximation using randomized forests. *Medical image*

- analysis 17, 790–804.
- Kotrotsou, A., Bennett, D.A., Schneider, J.A., Dawe, R.J., Golak, T., Leurgans, S.E., Yu, L., Arfanakis, K., 2014. Ex vivo MR volumetry of human brain hemispheres. *Magnetic Resonance in Medicine* 71, 364–374.
- Lancaster, J., Rainey, L., Summerlin, J., Freitas, C., Fox, P., Evans, A., Toga, A., Mazziotta, J., 1997. Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. *Human brain mapping* 5, 238.
- Lancelot, S., Roche, R., Slimen, A., Bouillot, C., Levigoureux, E., Langlois, J.B., Zimmer, L., Costes, N., 2014. A multi-atlas based method for automated anatomical rat brain MRI segmentation and extraction of PET activity. *PloS one* 9, e109113.
- Landman, B.A., Asman, A.J., Scoggins, A.G., Bogovic, J.A., Stein, J.A., Prince, J.L., 2012a. Foibles, follies, and fusion: Web-based collaboration for medical image labeling. *NeuroImage* 59, 530–539.
- Landman, B.A., Asman, A.J., Scoggins, A.G., Bogovic, J.A., Xing, F., Prince, J.L., 2012b. Robust statistical fusion of image labels. *Medical Imaging, IEEE Transactions on* 31, 512–522.
- Langerak, T., van der Heide, U.A., Kotte, A.N., Berendsen, F.F., Pluim, J.P., 2011. Local atlas selection and performance estimation in multi-atlas based segmentation, in: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, IEEE. pp. 669–672.
- Langerak, T.R., Berendsen, F.F., Van der Heide, U.A., Kotte, A.N., Pluim, J.P., 2013. Multiatlas-based segmentation with preregistration atlas selection. *Medical physics* 40, 091701.
- Langerak, T.R., van der Heide, U.A., Kotte, A.N., Viergever, M.A., van Vulpen, M., Pluim, J.P., 2010. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). *Medical Imaging, IEEE Transactions on* 29, 2000–2008.
- Ledig, C., Heckemann, R.A., Hammers, A., Lopez, J.C., Newcombe, V.F., Makropoulos, A., Lötjönen, J., Menon, D.K., Rueckert, D., 2014. Robust whole-brain segmentation: application to traumatic brain injury. *Medical image analysis* .
- Lee, A.W., Rajagopal, V., Babarenda Gamage, T.P., Doyle, A.J., Nielsen, P.M., Nash, M.P., 2013. Breast lesion co-localisation between X-ray and MR images using finite element modelling. *Medical image analysis* 17, 1256–1264.
- Lee, J., Lyu, I., Styner, M., 2014a. Multi-atlas segmentation with particle-based group-wise image registration, in: *SPIE Medical Imaging, International Society for Optics and Photonics*. pp. 903447–903447.
- Lee, J.G., Gumus, S., Moon, C.H., Kwoh, C.K., Bae, K.T., 2014b. Fully automated segmentation of cartilage from the MR images of knee using a multi-atlas and local structural analysis method. *Medical physics* 41, 092303.
- Leung, K.K., Barnes, J., Modat, M., Ridgway, G.R., Bartlett, J.W., Fox, N.C., Ourselin, S., 2011. Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. *Neuroimage* 55, 1091–1108.
- Leung, K.K., Barnes, J., Ridgway, G.R., Bartlett, J.W., Clarkson, M.J., Macdonald, K., Schuff, N., Fox, N.C., Ourselin, S., 2010. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *Neuroimage* 51, 1345–1359.
- Li, B., Panda, S., Xu, Z., Asman, A.J., Shanahan, P.L., Abramson, R.G., Landman, B.A., 2013. Regression forest region recognition enhances multi-atlas spleen labeling, in: *MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (SATA)*, pp. 82–92.
- Li, G., Wang, L., Shi, F., Lin, W., Shen, D., 2014. Simultaneous and consistent labeling of longitudinal dynamic developing cortical surfaces in infants. *Medical image analysis* 18, 1274–1289.
- Liao, S., Gao, Y., Lian, J., Shen, D., 2013. Sparse patch-based label propagation for accurate prostate localization in CT images. *Medical Imaging, IEEE Transactions on* 32, 419–434.
- van der Lijn, F., den Heijer, T., Breteler, M., Niessen, W.J., 2008. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *Neuroimage* 43, 708–720.
- Litjens, G., Debats, O., Barentsz, J., Karssemeijer, N., Huisman, H., 2014. Computer-aided detection of prostate cancer in MRI. *IEEE transactions on medical imaging* 33, 1083–1092.
- Liu, J., Zhao, J., Hoffman, J., Yao, J., Lu, L., Turkbey, E.B., Kim, C., Summers, R.M., 2014. Detection and station mapping of mediastinal lymph nodes on thoracic computed tomography using spatial prior from multi-atlas label fusion, in: *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, IEEE. pp. 1107–1110.
- Lotjonen, J., Koikkalainen, J., Thurfjell, L., Rueckert, D., 2009. Atlas-based registration parameters in segmenting sub-cortical regions from brain MRI-images, in: *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, IEEE. pp. 21–24.
- Lötjönen, J.M., Wolz, R., Koikkalainen, J.R., Thurfjell, L., Walde-mar, G., Soininen, H., Rueckert, D., 2010. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *Neuroimage* 49, 2352–2365.
- Ma, D., Cardoso, M.J., Modat, M., Powell, N., Wells, J., Holmes, H., Wiseman, F., Tybulewicz, V., Fisher, E., Lythgoe, M.F., et al., 2014. Automatic structural parcellation of mouse brain MRI using multi-atlas label fusion. *PloS one* 9, e86576.
- Maier-Hein, L., Mersmann, S., Kondermann, D., Stock, C., Keng-gott, H.G., Sanchez, A., Wagner, M., Preukschas, A., Wekerle, A.L., Helfert, S., et al., 2014. Crowdsourcing for reference correspondence generation in endoscopic images, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, pp. 349–356.
- Makropoulos, A., Gousias, I.S., Ledig, C., Aljabar, P., Serag, A., Hajnal, J.V., Edwards, A.D., Counsell, S.J., Rueckert, D., 2014. Automatic whole brain MRI segmentation of the developing neonatal brain. *IEEE Transactions on Medical Imaging* .
- Menze, B., Reyes, M., Van Leemput, K., 2014. The multimodal brain tumorimage segmentation benchmark (BRATS). *Medical Imaging, IEEE Transactions on* .
- Meyer, A., 2014. Multi-atlas based segmentation of corpus callosum on MRIs of multiple sclerosis patients, in: *Pattern Recognition*. Springer, pp. 729–735.
- Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine* 98, 278–284.
- Nie, J., Shen, D., 2013. Automated segmentation of mouse brain images using multi-atlas multi-ROI deformation and label fusion. *Neuroinformatics* 11, 35–45.
- Nouranian, S., Mahdavi, S., Spadinger, I., Morris, W., Salcudean, S., Abolmaesumi, P., 2014. A multi-atlas-based segmentation framework for prostate brachytherapy. *IEEE Transactions on Medical Imaging* .
- Ou, Y., Doshi, J., Erus, G., Davatzikos, C., 2012. Attribute similarity and mutual-saliency weighting for registration and label fusion, in: *MICCAI Workshop on Multi-Atlas Segmentation*. pp. 95–98, pp. 95–98.
- Ou, Y., Sotiras, A., Paragios, N., Davatzikos, C., 2011. Dramms: Deformable registration via attribute matching and mutual-saliency weighting. *Medical image analysis* 15, 622–639.
- Panda, S., Asman, A.J., Khare, S.P., Thompson, L., Mawn, L.A., Smith, S.A., Landman, B.A., 2014. Evaluation of multiatlas label fusion for in vivo magnetic resonance imaging orbital segmentation. *Journal of Medical Imaging* 1, 024002–024002.
- Park, H., Bland, P.H., Meyer, C.R., 2003. Construction of an abdominal probabilistic atlas and its application in segmentation. *Medical Imaging, IEEE Transactions on* 22, 483–492.
- Park, M.T.M., Pipitone, J., Baer, L.H., Winterburn, J.L., Shah, Y., Chavez, S., Schira, M.M., Lobaugh, N.J., Lerch, J.P., Voineskos, A.N., et al., 2014. Derivation of high-resolution MRI atlases of the human cerebellum at 3T and segmentation using multiple automatically generated templates. *Neuroimage* 95, 217–231.
- Pham, D.L., Xu, C., Prince, J.L., 2000. Current methods in medical image segmentation 1. *Annual review of biomedical engineering*

- 2, 315–337.
- Pieler, S., Halle, M., Kikinis, R., 2004. 3D slicer, in: Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on, IEEE. pp. 632–635.
- Pipitone, J., Park, M.T.M., Winterburn, J., Lett, T.A., Lerch, J.P., Pruessner, J.C., Lepage, M., Voineskos, A.N., Mallar Chakravarty, M., 2014. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. *Neuroimage* .
- Platero, C., Tobar, M., Sanguino, J., Velasco, O., 2014. A new label fusion method using graph cuts: application to hippocampus segmentation, in: XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013, Springer. pp. 174–177.
- Platero, C., Tobar, M.C., 2014. A multiatlas segmentation using graph cuts with applications to liver segmentation in CT scans. *Computational and mathematical methods in medicine* 2014.
- Pluta, J., Yushkevich, P., Das, S., Wolk, D., 2012. In vivo analysis of hippocampal subfield atrophy in mild cognitive impairment via semi-automatic segmentation of T2-weighted MRI. *Journal of Alzheimer's Disease* 31, 85–99.
- Pohl, K.M., Fisher, J., Grimson, W.E.L., Kikinis, R., Wells, W.M., 2006. A Bayesian model for joint segmentation and registration. *Neuroimage* 31, 228–239.
- Pouch, A.M., Wang, H., Takabe, M., Jackson, B.M., Gorman, J., Gorman, R.C., Yushkevich, P.A., Sehgal, C.M., 2014. Fully automatic segmentation of the mitral leaflets in 3D transesophageal echocardiographic images using multi-atlas joint label fusion and deformable medial modeling. *Medical image analysis* 18, 118–129.
- Pudil, P., Novovičová, J., Kittler, J., 1994. Floating search methods in feature selection. *Pattern recognition letters* 15, 1119–1125.
- Raamana, P.R., Rosen, H., Miller, B., Weiner, M.W., Wang, L., Beg, M.F., 2014. Three-class differential diagnosis among Alzheimer disease, frontotemporal dementia, and controls. *Frontiers in neurology* 5.
- Ramus, L., Commowick, O., Malandain, G., 2010. Construction of patient specific atlases from locally most similar anatomical pieces, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010. Springer, pp. 155–162.
- Ramus, L., Malandain, G., 2010. Assessing selection methods in the context of multi-atlas based segmentation, in: Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on, IEEE. pp. 1321–1324.
- Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61, 1402–1418.
- van Rikxoort, E., Arzhaeva, Y., van Ginneken, B., 2007a. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching, in: Proceeding MICCAI Workshop, Citeseer. pp. 101–108.
- van Rikxoort, E., Arzhaeva, Y., van Ginneken, B., 2007b. A multi-atlas approach to automatic segmentation of the caudate nucleus in MR brain images, in: Proceedings of MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge, pp. 29–36.
- van Rikxoort, E.M., de Hoop, B., Viergever, M.A., Prokop, M., van Ginneken, B., 2009. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical physics* 36, 2934–2947.
- van Rikxoort, E.M., Isgum, I., Arzhaeva, Y., Staring, M., Klein, S., Viergever, M.A., Pluim, J.P., van Ginneken, B., 2010. Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus. *Medical image analysis* 14, 39–49.
- Rivest-Hénault, D., Ghose, S., Pluim, J.P., Greer, P.B., Frapp, J., Dowling, J.A., 2014. Fast multiatlas selection using composition of transformations for radiation therapy planning, in: Medical Computer Vision: Algorithms for Big Data. Springer, pp. 105–115.
- Rohde, G.K., Aldroubi, A., Dawant, B.M., 2003. The adaptive bases algorithm for intensity-based nonrigid image registration. *Medical Imaging, IEEE Transactions on* 22, 1470–1479.
- Rohlfing, T., Brandt, R., Menzel, R., Maurer Jr, C.R., 2003a. Segmentation of three-dimensional images using non-rigid registration: Methods and validation with application to confocal microscopy images of bee brains, in: Medical Imaging 2003, International Society for Optics and Photonics. pp. 363–374.
- Rohlfing, T., Brandt, R., Menzel, R., Maurer Jr, C.R., 2004. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *Neuroimage* 21, 1428–1442.
- Rohlfing, T., Maurer Jr, C.R., 2005. Multi-classifier framework for atlas-based image segmentation. *Pattern Recognition Letters* 26, 2070–2079.
- Rohlfing, T., Russakoff, D.B., Maurer, C.R., 2003b. Expectation maximization strategies for multi-atlas multi-label segmentation, in: Information Processing in Medical Imaging, Springer. pp. 210–221.
- Rohlfing, T., Russakoff, D.B., Maurer Jr, C.R., 2003c. An expectation maximization-like algorithm for multi-atlas multi-label segmentation, in: Bildverarbeitung für die Medizin 2003. Springer, pp. 348–352.
- Rohlfing, T., Russakoff, D.B., Maurer Jr, C.R., 2003d. Extraction and application of expert priors to combine multiple segmentations of human brain tissue, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2003. Springer, pp. 578–585.
- Rousseau, F., Habas, P.A., Studholme, C., 2011. A supervised patch-based approach for human brain labeling. *Medical Imaging, IEEE Transactions on* 30, 1852–1862.
- Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *Medical Imaging, IEEE Transactions on* 18, 712–721.
- Rueda, S., Fathima, S., Knight, C.L., Yaqub, M., Papageorgiou, A.T., Rahmatullah, B., Foi, A., Maggioni, M., Pepe, A., Tohka, J., et al., 2014. Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge. *Medical Imaging, IEEE Transactions on* 33, 797–813.
- Sabuncu, M.R., Yeo, B.T., Van Leemput, K., Fischl, B., Golland, P., 2009. Supervised nonparametric image parcellation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009. Springer, pp. 1075–1083.
- Sabuncu, M.R., Yeo, B.T., Van Leemput, K., Fischl, B., Golland, P., 2010. A generative model for image segmentation based on label fusion. *Medical Imaging, IEEE Transactions on* 29, 1714–1729.
- Sandor, S., Leahy, R., 1997. Surface-based labeling of cortical anatomy using a deformable atlas. *Medical Imaging, IEEE Transactions on* 16, 41–54.
- Sanroma, G., Wu, G., Gao, Y., Shen, D., 2014a. Learning to rank atlases for multiple-atlas segmentation. *IEEE transactions on medical imaging* .
- Sanroma, G., Wu, G., Thung, K., Guo, Y., Shen, D., 2014b. Novel multi-atlas segmentation by matrix completion, in: Machine Learning in Medical Imaging. Springer, pp. 207–214.
- Schreibmann, E., Marcus, D.M., Fox, T., 2014. Multiatlas segmentation of thoracic and abdominal anatomy with level set-based local search. *Journal of Applied Clinical Medical Physics* 15, 22–38.
- Sdika, M., 2010. Combining atlas based segmentation and intensity classification with nearest neighbor transform and accuracy weighted vote. *Medical Image Analysis* 14, 219–226.
- Shan, L., Zach, C., Charles, C., Niethammer, M., 2014. Automatic atlas-based three-label cartilage segmentation from MR knee images. *Medical image analysis* 18, 1233–1246.
- Shen, K., Bourgeat, P., Frapp, J., Meriaudeau, F., Ames, D., Ellis, K.A., Masters, C.L., Villemagne, V.L., Rowe, C.C., Salvado, O., 2010. Supervised method to build an atlas database for multi-atlas segmentation-propagation, in: SPIE Medical Imaging, International Society for Optics and Photonics. pp. 76241N–76241N.
- Shi, F., Yap, P.T., Fan, Y., Gilmore, J.H., Lin, W., Shen, D., 2010. Construction of multi-region-multi-reference atlases for neonatal brain MRI segmentation. *Neuroimage* 51, 684–693.
- Shi, Y., Lai, R., Toga, A.W., 2013. Conformal mapping via metric optimization with application for cortical label fusion, in: Infor-

- mation Processing in Medical Imaging, Springer. pp. 244–255.
- Simpson, I.J., Woolrich, M.W., Schnabel, J.A., 2011. Probabilistic segmentation propagation from uncertainty in registration, in: *Proceedings Medical Image Analysis and Understanding (MIUA)*.
- Sjöberg, C., Ahnesjö, A., 2013. Multi-atlas based segmentation using probabilistic label fusion with adaptive weighting of image similarity measures. *Computer methods and programs in biomedicine* 110, 308–319.
- Sjöberg, C., Johansson, S., Ahnesjö, A., 2014. How much will linked deformable registrations decrease the quality of multi-atlas segmentation fusions? *Radiation Oncology* 9, 251.
- Sjöberg, C., Lundmark, M., Granberg, C., Johansson, S., Ahnesjö, A., Montelius, A., 2013. Clinical evaluation of multi-atlas based segmentation of lymph node regions in head and neck and prostate cancer patients. *Radiation Oncology* 8, 229.
- Sotiras, A., Davatzikos, C., Paragios, N., 2013. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on* 32, 1153–1190.
- Stavros, A., Aristeidis, S., Nikos, P., 2014. Discrete multi atlas segmentation using agreement constraints, in: *British Machine Vision Conference*.
- Svarer, C., Madsen, K., Hasselbalch, S.G., Pinborg, L.H., Haugbøl, S., Frøkjær, V.G., Holm, S., Paulson, O.B., Knudsen, G.M., 2005. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *Neuroimage* 24, 969–979.
- Ta, V.T., Giraud, R., Collins, D.L., Coupé, P., 2014. Optimized patchmatch for near real time and accurate label fusion, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, pp. 105–112.
- Tamez-Pena, J.G., Farber, J., Gonzalez, P.C., Schreyer, E., Schneider, E., Totterman, S., 2012. Unsupervised segmentation and quantification of anatomical knee features: data from the Osteoarthritis Initiative. *Biomedical Engineering, IEEE Transactions on* 59, 1177–1186.
- Tang, X., Oishi, K., Faria, A.V., Hillis, A.E., Albert, M.S., Mori, S., Miller, M.I., 2013. Bayesian parameter estimation and segmentation in the multi-atlas random orbit model. *PloS one* 8, e65591.
- Tang, X., Yoshida, S., Hsu, J., Huisman, T.A., Faria, A.V., Oishi, K., Kuttan, K., Poretti, A., Li, Y., Miller, M.I., et al., 2014. Multi-contrast multi-atlas parcellation of diffusion tensor imaging of the human brain. *PloS one* 9, e96985.
- Teng, C.C., Shapiro, L.G., Kalet, I.J., 2010. Head and neck lymph node region delineation with image registration. *Biomedical engineering online* 9, 30.
- Traynor, C., Heckemann, R.A., Hammers, A., O’Muircheartaigh, J., Crum, W.R., Barker, G.J., Richardson, M.P., 2010. Reproducibility of thalamic segmentation based on probabilistic tractography. *Neuroimage* 52, 69–85.
- Tung, K.P., Bei, W.J., Shi, W.Z., Wang, H.Y., Tong, T., De Silva, R., Edwards, E., Rueckert, D., 2013. Multi-atlas based neointima segmentation in intravascular coronary OCT, in: *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on, IEEE*. pp. 1280–1283.
- Van Der Lijn, F., De Bruijne, M., Klein, S., Den Heijer, T., Hoogendam, Y.Y., Van Der Lugt, A., Breteler, M.M., Niessen, W.J., 2012. Automated brain structure segmentation based on atlas registration and appearance models. *Medical Imaging, IEEE Transactions on* 31, 276–286.
- Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: Efficient non-parametric image registration. *Neuroimage* 45, S61–S72.
- Wachinger, C., Golland, P., 2012. Spectral label fusion, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*. Springer, pp. 410–417.
- Wachinger, C., Golland, P., 2014. Atlas-based under-segmentation, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, pp. 315–322.
- Wan, J., Carass, A., Resnick, S.M., Prince, J.L., 2008. Automated reliable labeling of the cortical surface, in: *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, IEEE*. pp. 440–443.
- Wang, H., Cao, Y., Syeda-Mahmood, T., 2014a. Multi-atlas segmentation with learning-based label fusion, in: *Machine Learning in Medical Imaging*. Springer, pp. 256–263.
- Wang, H., Pouch, A., Takabe, M., Jackson, B., Gorman, J., Gorman, R., Yushkevich, P.A., 2013a. Multi-atlas segmentation with robust label transfer and label fusion, in: *Information Processing in Medical Imaging*. Springer Berlin Heidelberg, pp. 548–559.
- Wang, H., Suh, J.W., Das, S., Pluta, J., Altinay, M., Yushkevich, P., 2011. Regression-based label fusion for multi-atlas segmentation, in: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE*. pp. 1113–1120.
- Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A., 2013b. Multi-atlas segmentation with joint label fusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35, 611–623.
- Wang, H., Yushkevich, P., 2012a. Guiding automatic segmentation with multiple manual segmentations. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, 429–436.
- Wang, H., Yushkevich, P.A., 2012b. Spatial bias in multi-atlas based segmentation, in: *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE*. pp. 909–916.
- Wang, H., Yushkevich, P.A., 2013a. Groupwise segmentation with multi-atlas joint label fusion, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, pp. 711–718.
- Wang, H., Yushkevich, P.A., 2013b. Multi-atlas segmentation without registration: A supervoxel-based approach, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, pp. 535–542.
- Wang, J., Vachet, C., Rumple, A., Gouttard, S., Ouziel, C., Perrot, E., Du, G., Huang, X., Gerig, G., Styner, M.A., 2014b. Multi-atlas segmentation of subcortical brain structures via the AutoSeg software pipeline. *Frontiers in Neuroinformatics* 8.
- Wang, L., Chen, K.C., Gao, Y., Shi, F., Liao, S., Li, G., Shen, S.G., Yan, J., Lee, P.K., Chow, B., et al., 2014c. Automated bone segmentation from dental CBCT images using patch-based sparse representation and convex optimization. *Medical physics* 41, 043503. [2014WangMedPhys.pdf](#), replaces wang2013automated.
- Wang, L., Shi, F., Gao, Y., Li, G., Gilmore, J.H., Lin, W., Shen, D., 2014d. Integration of sparse multi-modality representation and anatomical constraint for iso-intense infant brain MR image segmentation. *NeuroImage* 89, 152–164.
- Wang, L., Shi, F., Li, G., Gao, Y., Lin, W., Gilmore, J.H., Shen, D., 2014e. Segmentation of neonatal brain MR images using patch-driven level sets. *Neuroimage* 84, 141–158.
- Wang, Y., Jia, H., Yap, P.T., Cheng, B., Wee, C.Y., Guo, L., Shen, D., 2012. Groupwise segmentation improves neuroimaging classification accuracy, in: *Multimodal Brain Image Analysis*. Springer, pp. 185–193.
- Wang, Z., Bhatia, K.K., Glocker, B., Marvao, A., Dawes, T., Misawa, K., Mori, K., Rueckert, D., 2014f. Geodesic patch-based segmentation, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, pp. 666–673.
- Warfield, S.K., Zou, K.H., Wells, W.M., 2004. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *Medical Imaging, IEEE Transactions on* 23, 903–921.
- Weier, K., Fonov, V., Lavoie, K., Doyon, J., Collins, D.L., 2014. Rapid automatic segmentation of the human cerebellum and its lobules (RASCAL) implementation and application of the patch-based label-fusion technique with a template library to segment the human cerebellum. *Human brain mapping*.
- Weisenfeld, N.I., Warfield, S.K., 2011a. Learning likelihoods for labeling (L3): a general multi-classifier segmentation algorithm, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*. Springer, pp. 322–329.
- Weisenfeld, N.I., Warfield, S.K., 2011b. SoftSTAPLE: truth and performance-level estimation from probabilistic segmentations, in: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE*. pp. 440–443.

- tional Symposium on, IEEE. pp. 441–446.
- Winston, G.P., Cardoso, M.J., Williams, E.J., Burdett, J.L., Bartlett, P.A., Espak, M., Behr, C., Duncan, J.S., Ourselin, S., 2013. Automated hippocampal segmentation in patients with epilepsy: Available free online. *Epilepsia* 54, 2166–2173.
- Wolz, R., Aljabar, P., Hajnal, J.V., Hammers, A., Rueckert, D., 2010a. LEAP: learning embeddings for atlas propagation. *Neuroimage* 49, 1316–1325.
- Wolz, R., Aljabar, P., Rueckert, D., Heckemann, R.A., Hammers, A., 2009. Segmentation of subcortical structures and the hippocampus in brain MRI using graph-cuts and subject-specific a-priori information, in: *Biomedical Imaging: From Nano to Macro*, 2009. ISBI'09. IEEE International Symposium on, IEEE. pp. 470–473.
- Wolz, R., Chu, C., Misawa, K., Fujiwara, M., Mori, K., Rueckert, D., 2013. Automated abdominal multi-organ segmentation with subject-specific atlas generation. *IEEE transactions on medical imaging* 32, 1723.
- Wolz, R., Heckemann, R.A., Aljabar, P., Hajnal, J.V., Hammers, A., Lotjonen, J., Rueckert, D., 2010b. Measuring atrophy by simultaneous segmentation of serial MR images using 4-D graph-cuts, in: *Biomedical Imaging: From Nano to Macro*, 2010 IEEE International Symposium on, IEEE. pp. 960–963.
- Wu, G., Wang, Q., Zhang, D., Nie, F., Huang, H., Shen, D., 2014. A generative probability model of joint label fusion for multi-atlas based brain segmentation. *Medical image analysis* 18, 881–890.
- Wu, M., Rosano, C., Lopez-Garcia, P., Carter, C.S., Aizenstein, H.J., 2007. Optimum template selection for atlas-based segmentation. *Neuroimage* 34, 1612–1618.
- Xiao, Y., Fonov, V.S., Beriault, S., Gerard, I., Sadikot, A.F., Pike, G.B., Collins, D.L., 2014a. Patch-based label fusion segmentation of brainstem structures with dual-contrast MRI for parkinson's disease. *International journal of computer assisted radiology and surgery* , 1–13.
- Xiao, Y., Jannin, P., D'Albis, T., Guizard, N., Haegelen, C., Lalys, F., Verin, M., Collins, D.L., 2014b. Investigation of morphometric variability of subthalamic nucleus, red nucleus, and substantia nigra in advanced parkinson's disease patients using automatic segmentation and pca-based analysis. *Human brain mapping* .
- Xie, Q., Ruan, D., 2014. Low-complexity atlas-based prostate segmentation by combining global, regional, and local metrics. *Medical physics* 41, 041909.
- Xu, Z., Asman, A.J., Shanahan, P.L., Abramson, R.G., Landman, B.A., 2014a. Simple is a good idea (and better with context learning), in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2014*. Springer, pp. 364–371.
- Xu, Z., Li, B., Panda, S., Asman, A.J., Merkle, K.L., Shanahan, P.L., Abramson, R.G., Landman, B.A., 2014b. Shape-constrained multi-atlas segmentation of spleen in CT, in: *SPIE Medical Imaging*, International Society for Optics and Photonics. pp. 903446–903446.
- Yan, Z., Zhang, S., Liu, X., Metaxas, D.N., Montillo, A., et al., 2014. Accurate whole-brain segmentation for alzheimers disease combining an adaptive statistical atlas and multi-atlas, in: *Medical Computer Vision. Large Data in Medical Imaging*. Springer, pp. 65–73.
- Yang, G., Chen, Y., Tang, L., Shu, H., Toumoulin, C., 2014a. Automatic left ventricle segmentation based on multiatlas registration in 4d CT images, in: *Biomedical Imaging (ISBI)*, 2014 IEEE 11th International Symposium on, IEEE. pp. 413–416.
- Yang, G., Gu, J., Chen, Y., Liu, W., Tang, L., Shu, H., Toumoulin, C., 2014b. Automatic kidney segmentation in CT images based on multi-atlas image registration, in: *Engineering in Medicine and Biology Society (EMBC)*, 2014 36th Annual International Conference of the IEEE, IEEE. pp. 5538–5541.
- Yang, J., Zhang, Y., Zhang, L., Dong, L., 2010. Automatic segmentation of parotids from CT scans using multiple atlases. *Medical Image Analysis for the Clinic: A Grand Challenge* , 323–330.
- Yeo, B., Sabuncu, M.R., Desikan, R., Fischl, B., Golland, P., 2008. Effects of registration regularization and atlas sharpness on segmentation accuracy. *Medical image analysis* 12, 603–615.
- Yeo, B.T., Sabuncu, M.R., Vercauteren, T., Holt, D.J., Amunts, K., Zilles, K., Golland, P., Fischl, B., 2010. Learning task-optimal registration cost functions for localizing cytoarchitecture and function in the cerebral cortex. *Medical Imaging, IEEE Transactions on* 29, 1424–1441.
- Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31, 1116–1128.
- Yushkevich, P.A., Wang, H., Pluta, J., Das, S.R., Craige, C., Avants, B.B., Weiner, M.W., Mueller, S., 2010. Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. *Neuroimage* 53, 1208–1224.
- Zhang, D., Guo, Q., Wu, G., Shen, D., 2012. Sparse patch-based label fusion for multi-atlas segmentation, in: *Multimodal Brain Image Analysis*. Springer, pp. 94–102.
- Zhang, D., Wu, G., Jia, H., Shen, D., 2011. Confidence-guided sequential label fusion for multi-atlas based segmentation, in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011*. Springer, pp. 643–650.
- Zhuang, X., Leung, K., Rhode, K., Razavi, R., Hawkes, D., Ourselin, S., 2010. Whole heart segmentation of cardiac MRI using multiple path propagation strategy, in: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2010*. Springer, pp. 435–443.
- Zikic, D., Glocker, B., Criminisi, A., 2014a. Classifier-based multi-atlas label propagation with test-specific atlas weighting for correspondence-free scenarios, in: *Medical Computer Vision: Algorithms for Big Data*. Springer, pp. 116–124.
- Zikic, D., Glocker, B., Criminisi, A., 2014b. Encoding atlases by randomized classification forests for efficient multi-atlas label propagation. *Medical image analysis* 18, 1262–1273.
- Zuluaga, M., Mendelson, A., Cardoso, M., Taylor, A., Ourselin, S., 2014. Multi-atlas based pathological stratification of d-tga congenital heart disease, in: *Biomedical Imaging (ISBI)*, 2014 IEEE 11th International Symposium on, IEEE. pp. 109–112.