Image Segmentation with Adaptive Spatial Priors from Joint Registration[*]

Haifeng Li[†], Weihong Guo[‡], Jun Liu[†], Li Cui[†], and Dongxing Xie[§]

Abstract. Image segmentation is a crucial but challenging task that has many applications. In medical imaging, for instance, intensity inhomogeneity and noise are common. In thigh muscle images, different muscles are closely packed together and there are often no clear boundaries between them. Intensity based segmentation models cannot separate one muscle from another. To solve such problems, in this work we present a segmentation model with adaptive spatial priors from joint registration. This model combines segmentation and registration in a unified framework to leverage their positive mutual influence. The segmentation is based on a modified Gaussian mixture model, which integrates intensity inhomogeneity and spatial smoothness. The registration plays the role of providing a shape prior. We adopt a modified sum of squared difference fidelity term and Tikhonov regularity term for registration and also utilize a Gaussian pyramid and parametric method for robustness. The connection between segmentation and registration is guaranteed by the cross entropy metric that aims to make the segmentation map (from segmentation) and deformed atlas (from registration) as similar as possible. This joint framework is implemented within a constraint optimization framework, which leads to an efficient algorithm. We evaluate our proposed model on synthetic and thigh muscle MR images. Numerical results show the improvement as compared to segmentation and registration performed separately and other joint models.

Key words. image segmentation, shape priors, image registration, intensity inhomogeneity, joint model, Gaussian mixture model, variational method, thigh muscle segmentation

MSC codes. 68U10, 65K10, 35R30

DOI. 10.1137/21M1444874

1. Introduction. Image segmentation is a classical problem in image processing. It becomes very challenging when the boundaries between objects of interest are of similar intensity and texture. In medical analysis, for example, the segmentation of different thigh muscles is an essential task for the evaluation of musculoskeletal diseases such as osteoarthritis. In addition, noise and intensity inhomogeneity make it even harder. Figure 1 shows an example of thigh muscle MR image.

In this case, segmentation methods only depending on intensity can hardly distinguish different muscles, so it is necessary to introduce auxiliary information as priors. So far, several approaches have been devoted to solve such problems. One kind of approach are semiautomatic

[†]Laboratory of Mathematics and Complex Systems (Ministry of Education of China), School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, People's Republic of China (hfli@mail.bnu.edu.cn, jliu@bnu.edu.cn, licui@bnu.edu.cn).

[‡]Department of Mathematics, Applied Mathematics and Statistics, Case Western Reserve University, Cleveland, OH 44106 USA (wxg49@case.edu).

[§]Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan, 410008, People's Republic of China (xdx1024@csu.edu.cn).
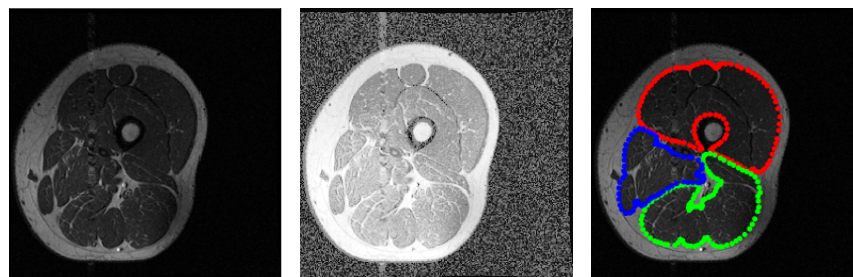
Figure 1. *A sample axial slice of thigh muscle in T1-weight MR image. Left: original MR image; middle: histogram equalized image; right: manually segmented objects. The muscle groups of quadriceps, hamstrings, and others are in red, green, and blue, respectively.*

methods (e.g., [1, 16, 30, 29, 27]) that obtain priors by the manual delineation of lines, polygons, or manual segmentation of beginning slices, and so on. These semiautomatic methods have proven to have some efficiency but need manual intervention that can be time consuming and tedious. Another kind of approach are fully automatic methods. Most of them are shape based methods (e.g., [5, 2, 18], and [8, 21, 36] for more general images) or atlas-based methods (e.g., [39, 46, 26]). However, these automatic methods are either single segmentation/registration techniques or sequential methods (like "pipeline" procedures), which ignore the mutual promotion between segmentation and registration. Recently, approaches based on deep learning have been emerging, such as using the AlexNet network [12], U-Net architecture [17], bounding boxes with 3D U-Net [28], and the edge-aware network based on U-Net [14]. Although deep learning has great potential in muscle segmentation, we should note that this data driven method needs a large amount of annotated data (manual segmentation of thigh muscles has been recognized as time consuming) and a specific network tuning for each dataset, which make it not very suitable and competitive for this task. Considering these facts, this current work is devoted to proposing a fully automatic model-based segmentation model, which integrates adaptive spatial priors from joint registration.

A joint segmentation and registration model can exploit the strong correlation between them, thus achieving more accurate results. For the segmentation, the registration can be viewed as a prior (e.g., shape or topology prior) to guide the segmentation process. In turn, relevant segmented structures can promote registration by providing a reliable estimation of the deformation between the encoded structures, making the registration not only based on intensity matching (local criterion), but also on geometrical and shape pairing (nonlocal character) [9]. There have been some related works discussing joint segmentation and registration. Most of them adopt active contour or level set—based variational approaches. For example, in the pioneering work, Yezzi, Zollei, and Kapur [45] introduced a variational framework that integrates an active contour segmentation model with a rigid registration technique to simultaneously segment and register features from multiple images. In [15], the authors combined a matching criterion based on the active contour without edges for segmentation and a nonlinear elasticity-based smoother on the displacement vector field to perform joint segmentation and registration. In [4], the authors formulated the joint problem as a minimization of a functional that integrates a nonlinear elastic registration with geodesic active contours, which is

introduced together with a weighted total variation term to segment the deforming template image. Pawar et al. [31] presented a joint approach using bidirectional composition-based level set formulation, in which the implicit level set function defining the segmentation contour and the displacement field for registration are both defined using B-splines. Swierczynski et al. [38] proposed an algorithm based on a level set formulation, which merges a classic Chan–Vese segmentation with an active dense displacement field estimation. Debroux and Guyader [10] established a joint model which is based on a nonlocal characterization of weighted total variation and nonlocal shape descriptors inspired by the piecewise constant Mumford–Shah model. However, these variational methods have some drawbacks. First, all the above methods [45, 15, 4, 31, 38, 10] are designed for segmentation of two classes, and the extension to multiclasses is not straightforward. Second, some of the joint models [45, 4, 10] are still intensity-based methods, as they only consider the pixel value information from two images, without using any atlas as a shape prior. Third, in [15, 38], the registration is a matching of segmentation maps, without considering intensity matching of the two images, thus making a rough registration. Therefore, these joint models cannot accomplish the segmentation of thigh muscle images with regions of interest that are closely packed together without clear boundaries in between.

In addition to the variational methods, there are also statistical methods for joint segmentation and registration, and most of them are based on the expectation maximization (EM) algorithm. For example, Wyatt and Noble [44] applied a Markov random field framework, a mathematical technique for embedding local spatial information, within which they seek to obtain a maximum a posteriori estimate of the segmentation and registration. Ashburner and Friston [3] presented a probabilistic framework based on GMM that enables image registration, tissue classification, and intensity bias correction to be combined within the same generative model. In [32], the authors developed a statistical model that combines the registration of an atlas with the segmentation of MR images. The model is solved by an EM algorithm which simultaneously estimates image artifacts, anatomical label maps, and a structure-dependent hierarchical mapping from the atlas to the image space. Gooya et al. [13] presented a generative method for simultaneously registering a probabilistic atlas of a healthy population to brain MR scans showing glioma and segmenting the scans into tumor as well as healthy tissue labels, by integrating the EM algorithm with a glioma growth model, such that EM iteratively refines the estimates of the posterior probabilities of tissue labels, the deformation field, and the tumor growth model parameters. These statistical methods are more natural and good at dealing with big data, as well as segmentation of multiclasses. However, in [44, 3, 32, 13], registrations only consider the matching between the segmentation map and the deformed atlas, thus making a rough registration. Besides, statistical methods are susceptible to noise and are not easy to combine with excellent variational properties such as spatial regularization and geometric structure. In contrast, variational methods can incorporate these properties more flexibly. In [20], the authors proposed a variational framework to solve GMM-based methods for image segmentation, thus making it flexible to combine with spatial regularization and bias correction, but it didn't consider a registration. Inspired by the works [3] and [20], we combine the GMM-based segmentation with a nonrigid registration in a variational framework, which integrates not only the merits of both variational and statistical methods, but also the advantages of a joint model. To the best of our knowledge, this

is the first time such an approach has been used in the joint segmentation and registration framework.

The main work of the current paper can be summarized as follows:

- We present a fully automatic model for joint segmentation and registration, which is designed for multiclasses segmentation. By interpreting the GMM-based model in a variational sense, this joint model absorbs the merits of both variational and statistical methods. The segmentation component combines GMM with bias correction and spatial regularization in a variational framework. The registration component consists of a modified sum of squared difference (SSD) and the Tikhonov regularization. The coupling between segmentation and registration is achieved by a cross entropy metric, making the segmentation map as close as possible to the deformed atlas.

- In this joint model, the registration can provide an adaptive spatial prior for the segmentation process. In turn, the segmentation can promote registration by providing geometrical structures encoded in the segmentation map. Therefore, the registration process considers not only the matching of image intensity, but also geometrical pairing through the matching of the segmentation map and the deformed atlas. This two levels matching makes a more accurate registration, and further a more reliable spatial prior for segmentation.

- This proposed joint model shows improved performance on synthetic images and thigh muscle MR images compared to sequential methods (segmentation and registration done separately) and other joint models.

The remainder of this paper is organized as follows. In section 2, we introduce the related work and the motivation of this work. Then we introduce the proposed joint segmentation and registration model and its optimization algorithm in section 3. We show some numerical experiments to demonstrate the effectiveness of the proposed model in section 4. Finally, a brief conclusion is drawn in section 5.

2. Related work and motivation. In this section, we will review some related work based on GMM, introduce the motivation of our work, and make a description of image registration.

2.1. Gaussian mixture model. Gaussian mixture models have been widely used in many classification problems (e.g., [6] and [23]), such as image segmentation. For a gray image $I : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ (e.g., $d = 2$ for two-dimensional (2D) images and $d = 3$ for 3D images), the probability of pixel intensity z can be modeled as

$$(2.1) \qquad p(z) = \sum_{k=1}^{K} \gamma_k p_k(z; c_k, \sigma_k^2) \quad \text{with} \quad \sum_{k=1}^{K} \gamma_k = 1, \gamma_k \geq 0,$$

where K is the total number of mixtures, γ_k is the weight of the kth Gaussian component, and p_k is the Gaussian probability density function parameterized by mean c_k and variance σ_k^2. Denote parameters in GMM as

$$\Theta = \{\gamma_1, \ldots, \gamma_K, c_1, \ldots, c_K, \sigma_1^2, \ldots, \sigma_K^2\},$$

and assume that all pixels are independent, then the probability density function of the entire image I can be written as

$$p(I|\Theta) = \prod_{x \in \Omega} \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{[I(x)-c_k]^2}{2\sigma_k^2}\right\}.$$

This probability can be maximized with respect to the unknown parameter Θ, by minimizing the following negative log-likelihood function:

$$(2.2) \qquad \mathcal{L}(\Theta) = -\int_{\Omega} \log \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{[I(x)-c_k]^2}{2\sigma_k^2}\right\} dx.$$

The parameters in the GMM can be efficiently estimated using the EM algorithm [22], and then the clusters can be obtained through the estimated parameters.

2.2. Unified segmentation. A probabilistic framework is presented in [3] that combines image segmentation, image registration, and bias correction within a GMM-based model.

In this work, bias correction is included in GMM by extra parameters that account for smooth intensity variations. The bias field at pixel x is denoted as $\rho(x;\boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is a vector of unknown parameters. Assume that the bias is multiplicative, and the kth cluster is normally distributed with mean $c_k/\rho(x;\boldsymbol{\beta})$ and variance $(\sigma_k/\rho(x;\boldsymbol{\beta}))^2$. The negative log-likelihood function (2.2) can be rewritten as

$$(2.3) \qquad \mathcal{L}_{\boldsymbol{\beta}}(\Theta) = -\int_{\Omega} \log \sum_{k=1}^{K} \frac{\gamma_k \rho(x;\boldsymbol{\beta})}{\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{[\rho(x;\boldsymbol{\beta})I(x)-c_k]^2}{2\sigma_k^2}\right\} dx.$$

In GMM, the prior probability of any pixel, irrespective of its intensity, belonging to the kth Gaussian is the stationary mixing proportion γ_k. For medical images, additional information usually can be obtained from other subjects' images with tissue probability maps. These maps give the probability of any pixel being of any of the tissue classes. For example, in the brain image segmentation, the tissue probability maps $s_k(x), k = 1, 2, 3, 4$, could represent the probability of pixel x being of classes of grey matter, white matter, cerebrospinal fluid, and "other," respectively. The $s_k(x)$ can be further deformed according to a parameter vector $\boldsymbol{\alpha}$, which matches $s_k(x)$ to the image to be segmented. The deformed tissue probability map $s_k(x;\boldsymbol{\alpha})$ therefore gives a prior probability of any pixel in a registered image being of any of the tissue classes. To combine with this prior, in [3], the authors replace the stationary mixing proportion γ_k by $\frac{\gamma_k s_k(x;\boldsymbol{\alpha})}{\sum_{l=1}^{K} \gamma_l s_l(x;\boldsymbol{\alpha})}$, which introduces a segmentation prior at each pixel x. Then the model (2.3) is modified to

$$\mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\Theta) = -\int_{\Omega} \log \sum_{k=1}^{K} \frac{\gamma_k s_k(x;\boldsymbol{\alpha})\rho(x;\boldsymbol{\beta})}{\left[\sum_{l=1}^{K} \gamma_l s_l(x;\boldsymbol{\alpha})\right]\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{[\rho(x;\boldsymbol{\beta})I(x)-c_k]^2}{2\sigma_k^2}\right\} dx.$$

Furthermore, to guarantee the smoothness of the bias field and the deformations, two regularization terms $p(\boldsymbol{\beta})$ and $p(\boldsymbol{\alpha})$ are considered in [3]. $p(\boldsymbol{\beta})$ and $p(\boldsymbol{\alpha})$ are zero-mean multivariate Gaussian probability density functions with covariances $\Sigma_{\boldsymbol{\beta}}$, $\Sigma_{\boldsymbol{\alpha}}$, respectively.

Above all, the objective function to be minimized in [3] can be written as

$$(2.4) \qquad \mathcal{E}(\Theta) = \mathcal{L}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\Theta) - \log p(\boldsymbol{\beta}) - \log p(\boldsymbol{\alpha}).$$

The unknown parameters are

$$\Theta = \{\gamma_1, \ldots, \gamma_K, c_1, \ldots, c_K, \sigma_1^2, \ldots, \sigma_K^2, \boldsymbol{\alpha}, \boldsymbol{\beta}\}.$$

The optimization of (2.4) involves the partial derivatives of the objective function and alternating among classification, bias correction, and registration [3].

This method effectively integrates intensity inhomogeneity and image registration into GMM, which can be a great help for tissue classification. However, this method still has some drawbacks. First, the assumption in GMM that pixels are independent is inappropriate because it ignores the smoothness of clusters. Although the spatial priors embody a degree of spatial dependency, it is not enough, as illustrated in the simulations in [3]. Second, although this method could accurately model the intensity inhomogeneity, the EM algorithm lacks spatial regularization and the additional regularization term for the bias parameter $\boldsymbol{\beta}$ is computationally time-consuming. Third, the registration in this model is performed by only matching the segmentation map with tissue probability prior, which is somehow a weak registration. Therefore it needs a preprocessing step to roughly align the images with the tissue probability prior to avoid strange results. Last, this model is embedded in the probabilistic framework of GMM, which makes it less straightforward to incorporate with spatial regularity. In this work, we will try to derive a more general framework to tackle these problems by considering the contributions of [20], introduced in the next subsection 2.3.

2.3. Image segmentation using a local GMM in a variational framework. A variational framework to solve GMM-based methods was proposed in [20]. Under this framework, the GMM-based method can be extended more easily; regularization terms, for instance, can be added. To achieve this, the following conclusion of convex analysis [33] is used.

Theorem 2.1 (commutativity of log-sum operations). *Given a function $\mathcal{A}_k(x) > 0$, for any function $\mathcal{B}_k(x) > 0$, we have*

$$-\log \sum_{k=1}^{K} \mathcal{A}_k(x) \exp[-\mathcal{B}_k(x)] = \min_{\boldsymbol{u} \in \mathbb{U}} \left\{ \sum_{k=1}^{K} [\mathcal{B}_k(x) - \log \mathcal{A}_k(x)] u_k(x) + \sum_{k=1}^{K} u_k(x) \log u_k(x) \right\},$$

where $\mathbb{U} = \left\{ \boldsymbol{u}(x) = (u_1(x), u_2(x), \ldots, u_K(x)) \in [0,1]^K : \sum_{k=1}^{K} u_k(x) = 1 \right\}.$

By defining $\mathcal{A}_k(x) := \frac{\gamma_k}{\sqrt{2\pi}\sigma_k}$, $\mathcal{B}_k(x) := \frac{[I(x) - c_k]^2}{2\sigma_k^2}$ and applying Theorem 2.1, we have that minimizing the negative log-likelihood function (2.2) is equivalent to minimizing functional

$$(2.5) \qquad \mathcal{E}(\Theta, \boldsymbol{u}) = \int_{\Omega} \sum_{k=1}^{K} \left[\frac{[I(x) - c_k]^2}{2\sigma_k^2} - \log \frac{\gamma_k}{\sqrt{2\pi}\sigma_k} \right] u_k(x) dx + \int_{\Omega} \sum_{k=1}^{K} u_k(x) \log u_k(x) dx.$$

This problem can be solved by alternately solving $\boldsymbol{u}$ and Θ with closed forms. One can easily check that the process for solving (2.5) is the same as the EM algorithm for solving (2.2).

In this variational form, the spatial regularization can be easily handled. In [20], the authors adopted a special smoothness term which is related to the boundary length of the clusters, called threshold dynamics regularization [24, 25, 11],

$$(2.6) \qquad \mathcal{R}(\boldsymbol{u}) = \int_\Omega \sum_{k=1}^{K} u_k(x)(\omega * (1 - u_k))(x)dx.$$

The symbol "*" stands for convolution and ω is a kernel function (usually a Gaussian kernel). Compared to the classical total variation regularization, a superiority of (2.6) is that it has no singularity and $\boldsymbol{u}$ still has an explicit updating formulation by using a linearization in some sense.

Based on the variational framework (2.5) of GMM, combining with regularization (2.6) and a bias field for intensity inhomogeneity, an effective segmentation model is proposed in [20].

2.4. The motivation of our work. As mentioned above, on the one hand, the unified segmentation model in [3] integrates GMM with intensity inhomogeneity and registration prior. However, it is not easy to combine with some good variational priors, such as spatial regularization, under such a probabilistic framework. Besides, the registration in this model only considers the matching of tissue probability maps, without considering other information such as image intensity, thus making it a rough registration.

On the other hand, the local GMM model in [20] combines GMM with spatial regularization and intensity inhomogeneity in a variational framework. Compared to [3], this model can flexibly combine some additional excellent variational properties. However, it lacks spatial priors from an atlas, which is indispensable in some difficult segmentation tasks, such as thigh muscle segmentation.

Therefore, a combination of both advantages will be more powerful. Inspired by the technique used in [3], one can introduce registration into the variational framework in a similar way. First, let us see what the replacement used in [3] means under the variational framework. For simplicity, denote $\mathcal{H}(\boldsymbol{u}) = \int_\Omega \sum_{k=1}^{K} u_k(x) \log u_k(x)dx$, and then replace γ_k with $\frac{s_k(x;\boldsymbol{\alpha})}{\sum_{l=1}^{K} s_l(x;\boldsymbol{\alpha})}$ in (2.5) (this is equivalent to the replacement used in [3]). One has

$$\tilde{\mathcal{E}}(\Theta, \boldsymbol{u}) = \int_\Omega \sum_{k=1}^{K} \left[\frac{[I(x) - c_k]^2}{2\sigma_k^2} - \log \frac{s_k(x;\boldsymbol{\alpha})}{\left[\sum_{l=1}^{K} s_l(x;\boldsymbol{\alpha})\right] \sqrt{2\pi}\sigma_k} \right] u_k(x)dx + \mathcal{H}(\boldsymbol{u})$$

$$= \int_\Omega \sum_{k=1}^{K} \left[\frac{[I(x) - c_k]^2}{2\sigma_k^2} - \log \frac{1}{\sqrt{2\pi}\sigma_k} \right] u_k(x)dx - \int_\Omega \sum_{k=1}^{K} u_k(x) \log \frac{s_k(x;\boldsymbol{\alpha})}{\sum_{l=1}^{K} s_l(x;\boldsymbol{\alpha})} dx + \mathcal{H}(\boldsymbol{u}).$$

It is not difficult to find that replacing γ_k with $\frac{s_k(x;\boldsymbol{\alpha})}{\sum\limits_{l=1}^{K} s_l(x;\boldsymbol{\alpha})}$ in the probabilistic framework is equivalent to modifying the cross entropy metric in the variational framework, i.e., replacing $-\int_\Omega \sum_{k=1}^{K} u_k(x) \log \gamma_k dx$ with $-\int_\Omega \sum_{k=1}^{K} u_k(x) \log \frac{s_k(x;\boldsymbol{\alpha})}{\sum\limits_{l=1}^{K} s_l(x;\boldsymbol{\alpha})} dx$. The former cross entropy only considers the volume of classes, not including information at pixels (γ_k only depends on k but not x). But the latter introduced a prior (deformed tissue probability map) at each pixel x ($s_k(x;\boldsymbol{\alpha})$ depends on both k and x) by combining with a registration problem.

With this observation, in the variational form, the registration model can be easily extended. In subsection 3.2, we will present a more general variational framework for joint segmentation and registration, and the two components are connected by a cross entropy metric.

2.5. Image registration. Before presenting the proposed framework, for completeness, we would like to first introduce image registration briefly. Image registration is a process to establish spatial correspondence between different images. The goal of registration is to estimate an optimal displacement field $T : \Omega \subset \mathbb{R}^d \to \mathbb{R}^d$ that maps a moving image $I_m : \Omega \to \mathbb{R}$ to a fixed image $I_f : \Omega \to \mathbb{R}$. Image registration can commonly be formulated as the following problem:

$$\min_T E_D(I_f, I_m \circ T) + \eta E_R(T),$$

where $I_m \circ T$ is the deformed image of I_m. E_D is an image dissimilarity metric that quantifies the level of alignment between I_f and $I_m \circ T$. E_R is a regularization term to favor specific properties that the user needs and tackle the difficulty of an ill-posed problem, while η controls the amount of regularization. Much research has been dedicated to image registration, such as [35, 34, 43], to name a few. Comprehensive reviews of image registration techniques can be found in [47, 37]. The selection of registration method depends on the anatomical properties of tissues and the features of images to be registered. In this paper, we mainly focus on the registration of human thigh muscle images, which usually have small deformation between different subjects. Therefore, in this paper, we just choose a simple registration model to demonstrate the effectiveness of the proposed joint segmentation and registration framework. Specifically, we use the SSD dissimilarity metric and the Tikhonov regularity, that is,

$$(2.7) \qquad \min_T \frac{1}{2} \int_\Omega [I_f(x) - I_m(x + T(x))]^2 \, dx + \frac{\eta}{2} \int_\Omega |\nabla T(x)|^2 dx.$$

In this case, any gradient-based solver can be used to minimize (2.7) by computing the derivatives of the smooth metric E_D and regularizer E_R with respect to the displacement field T. Further, we combine this registration model with a Gaussian pyramid and parametric technique for robustness. In the small deformation case, this registration method can usually provide a not bad spatial shape prior. More sophisticated registration models, such as diffeomorphic registration, feature-based registration, may obtain further benefits, and lead to better overall results, but that is not the focus of this work; it is reserved for follow-up work.

3. The proposed joint segmentation and registration model. When there is not enough boundary information to separate one region from another, one needs some prior spatial

information. We adopt an automatic spatial prior by registering the image to be segmented to an image with ground truth (GT) segmentation. Let $I_m : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ be such an image with GT segmentation $\boldsymbol{s} = (s_1, \ldots, s_K)$; here K stands for the number of clusters, and s_k the segmentation map for the kth cluster, $k = 1, \ldots, K$. Denote $I_f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ the given image to be segmented. First, let's introduce the GMM-based segmentation model for intensity inhomogeneity image.

3.1. The GMM with intensity inhomogeneity. As mentioned in the introduction and shown in Figure 1, thigh muscle MR images usually have the problem of intensity inhomogeneity, which is often caused by a smooth, spatially varying illumination, etc. Although it is not usually a problem for visual inspection, it can hinder the performance of automatic intensity-based image segmentation. According to its contributing factors, many algorithms assume that the bias is multiplicative. In this work, similar to [20], we model the intensity inhomogeneity image as

$$(3.1) \qquad\qquad\qquad I_f(x) = \beta(x)g(x),$$

where $I_f(x)$ is the observed image, $g(x)$ represents the GT image, and $\beta(x)$ is a smoothly varying intensity bias field.

By assuming that the true image g can be well segmented by the GMM like (2.1), with the relationship (3.1), the probability of pixel intensity $I_f(x)$ can be modeled as

$$p(I_f(x)) = \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(x)} \exp\left\{ -\frac{[I_f(x) - c_k\beta(x)]^2}{2\sigma_k^2\beta^2(x)} \right\}.$$

Besides, since the bias field β is varying smoothly, i.e., $\beta(x) \approx \beta(y)$ when x is in a small neighborhood O_y centered at y, we can assume that the intensities of image I_f within the neighborhood O_y share the same probability density function parameterized by $\gamma_k, c_k, \sigma_k^2, \beta(y)$. By the independence assumption, we have the local negative log-likelihood function in O_y

$$\mathcal{L}_y(\Theta) = -\int_{O_y} \log \sum_{k=1}^{K} \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} \exp\left\{ -\frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} \right\} dx;$$

here $\Theta = \{\gamma_1, \ldots, \gamma_K, c_1, \ldots, c_K, \sigma_1^2, \ldots, \sigma_K^2, \beta(y)\}$.

If we consider different contributions of each $x \in O_y$ to the local cost function $\mathcal{L}_y$ in terms of the distance to centering point y, we can assign a weight for each pixel. A common choice is to use a Gaussian kernel G_σ with an appropriate standard deviation σ, such that $G_\sigma(y - x) \approx 0$ when $x \notin O_y$. Then the local cost function in O_y becomes

$$(3.2) \qquad \mathcal{L}_y(\Theta) = -\int_\Omega G_\sigma(y - x) \log \sum_{k=1}^{K} \exp\left\{ \log \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} - \frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} \right\} dx.$$

Note that the above function (3.2) has the form of log-sum-exp. Now we introduce the dual formulation of log-sum-exp functional, which can be regarded as a generalization of Theorem 2.1.

Theorem 3.1. *Let functional*

$$\mathcal{F}(\boldsymbol{z}) = \varepsilon \log \sum_{k=1}^{K} \exp\left\{\frac{z_k}{\varepsilon}\right\},$$

and then its Fenchel–Legendre transformation is

$$\mathcal{F}^*(\boldsymbol{u}) = \max_{\boldsymbol{z}} \left\{ < \boldsymbol{z}, \boldsymbol{u} > -\mathcal{F}(\boldsymbol{z}) \right\}$$

$$= \begin{cases} \varepsilon \sum_{k=1}^{K} u_k \log u_k, & \boldsymbol{u} \in \mathbb{U}, \\ +\infty & else, \end{cases}$$

where $\mathbb{U} = \{\boldsymbol{u} = (u_1, u_2, \ldots, u_K) \in [0,1]^K : \sum_{k=1}^{K} u_k = 1\}$. *Moreover,* $\mathcal{F}(\boldsymbol{z})$ *is convex with respect to* $\boldsymbol{z}$ *and thus*

$$\mathcal{F}(\boldsymbol{z}) = \mathcal{F}^{**}(\boldsymbol{z}) = \max_{\boldsymbol{u} \in \mathbb{U}} \left\{ < \boldsymbol{z}, \boldsymbol{u} > -\varepsilon \sum_{k=1}^{K} u_k \log u_k \right\}.$$

The proof of this theorem is a standard argument of convex optimization; we leave it to the readers.

With Theorem 3.1, problem (3.2) has the following variational form:

$$\mathcal{L}_y(\Theta) = \min_{\boldsymbol{u} \in \mathbb{U}} \int_\Omega G_\sigma(y - x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} - \log \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} \right) u_k(x) \right] dx$$

$$+ \int_\Omega \sum_{k=1}^{K} u_k(x) \log u_k(x) dx.$$

Considering the global information, then the total cost function can be written as

$$\mathcal{L}(\Theta) = \min_{\boldsymbol{u} \in \mathbb{U}} \int_\Omega \int_\Omega G_\sigma(y - x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} - \log \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} \right) u_k(x) \right] dx dy$$

$$+ \int_\Omega \sum_{k=1}^{K} u_k(x) \log u_k(x) dx;$$

here parameter set $\Theta = \{\gamma_1, \ldots, \gamma_K, c_1, \ldots, c_K, \sigma_1^2, \ldots, \sigma_K^2, \cup_y \beta(y)\}$.

In this formulation, we can easily combine with a regularization like (2.6). Finally, we have the following image segmentation model:

$$(3.3) \qquad \min_{\Theta, \boldsymbol{u} \in \mathbb{U}} \hat{\mathcal{E}}(\Theta, \boldsymbol{u}) = \mathcal{F}(\Theta, \boldsymbol{u}) + \lambda \mathcal{R}(\boldsymbol{u}),$$

where

$$\mathcal{F}(\Theta, \boldsymbol{u}) = \int_\Omega \int_\Omega G_\sigma(y - x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} - \log \frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} \right) u_k(x) \right] dxdy$$
$$+ \int_\Omega \sum_{k=1}^{K} u_k(x) \log u_k(x) dx,$$

$$\mathcal{R}(\boldsymbol{u}) = \int_\Omega \sum_{k=1}^{K} u_k(x)(\omega * (1 - u_k))(x) dx.$$

The minimizer of $\hat{\mathcal{E}}(\Theta, \boldsymbol{u})$ can be computed by the following alternating algorithm:

$$(3.4) \qquad \begin{cases} \Theta^{t+1} = \underset{\Theta}{\arg\min}\ \mathcal{F}(\Theta, \boldsymbol{u}^t), \\ \boldsymbol{u}^{t+1} = \underset{\boldsymbol{u} \in \mathbb{U}}{\arg\min}\ \mathcal{F}(\Theta^{t+1}, \boldsymbol{u}) + \lambda\mathcal{R}(\boldsymbol{u}; \boldsymbol{u}^t), \end{cases}$$

where $t = 0, 1, 2, \dots$ stands for the iteration number, and

$$\mathcal{R}(\boldsymbol{u}; \boldsymbol{u}^t) = \int_\Omega \sum_{k=1}^{K} u_k(x)(\omega * (1 - 2u_k^t))(x) dx$$

is the linearization of $\mathcal{R}(\boldsymbol{u})$ at $\boldsymbol{u}^t$. We can get the following energy descent theorem.

Theorem 3.2 (energy descent). *The sequence $(\Theta^t, \boldsymbol{u}^t)$ produced by iteration scheme (3.4) satisfies*

$$\hat{\mathcal{E}}(\Theta^{t+1}, \boldsymbol{u}^{t+1}) \leq \hat{\mathcal{E}}(\Theta^t, \boldsymbol{u}^t).$$

Proof. See Appendix A. ∎

Compared with segmentation model (2.4) in [3], the proposed segmentation model (3.3) integrates GMM with a variational spatial regularization, and here the smoothness of the bias field is guaranteed by a Gaussian kernel and thus doesn't need an extra regularization term as in [3]. In addition, since both are deduced from the property of log-sum-exp function described in Theorem 3.1, and both use the threshold dynamics regularization, the proposed segmentation model (3.3) is very close to the soft threshold dynamics model proposed in [19]. The two models are both composed of a similarity term, the entropy regularization, and the threshold dynamics regularization. However, here in (3.3), the similarity term is derived from GMM combined with a bias field, which is very important for medical images with intensity inhomogeneity. Besides, as shown in the next subsection 3.2, we integrate this model (3.3) with an adaptive shape prior from joint registration, instead of combining a geometric prior of the segmented region itself (e.g., star shape prior) as in [19]. These differences make the proposed model in this paper more suitable for segmenting images like thigh muscle MRI.

3.2. The proposed joint model. With the segmentation model introduced above and the motivation described in subsection 2.4, our proposed variational framework for joint segmentation and registration is to optimize the following energy function:

$$(3.5) \qquad \mathcal{E}(\Theta, \boldsymbol{u}, T) = \mathcal{E}_{\mathrm{Seg}}(\Theta, \boldsymbol{u}) + \mathcal{E}_{\mathrm{CE}}(\boldsymbol{u}, \boldsymbol{s} \circ T) + \mathcal{E}_{\mathrm{Reg}}(\Theta, T),$$

where

$$\mathcal{E}_{\mathrm{Seg}}(\Theta, \boldsymbol{u}) = \int_{\Omega} \int_{\Omega} G_{\sigma}(y - x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k \beta(y)]^2}{2\sigma_k^2 \beta^2(y)} - \log \frac{\gamma_k}{\sqrt{2\pi} \sigma_k \beta(y)} \right) u_k(x) \right] dx dy$$

$$+ \varepsilon \int_{\Omega} \sum_{k=1}^{K} u_k(x) \log u_k(x) dx + \lambda \int_{\Omega} \sum_{k=1}^{K} u_k(x)(\omega * (1 - u_k))(x) dx,$$

$$\mathcal{E}_{\mathrm{CE}}(\boldsymbol{u}, \boldsymbol{s} \circ T) = -\xi \int_{\Omega} \sum_{k=1}^{K} u_k(x) \log \frac{s_k(x + T(x))}{\sum\limits_{l=1}^{K} s_l(x + T(x))} dx,$$

$$\mathcal{E}_{\mathrm{Reg}}(\Theta, T) = \frac{\zeta}{2} \int_{\Omega} \left[\frac{I_f(x)}{\beta(x)} - I_m(x + T(x)) \right]^2 dx + \frac{\eta}{2} \int_{\Omega} |\nabla T(x)|^2 dx,$$

and $\Theta = \{\gamma_1, \ldots, \gamma_K, c_1, \ldots, c_K, \sigma_1^2, \ldots, \sigma_K^2, \cup_y \beta(y)\}$. Parameters $\varepsilon, \lambda, \xi, \zeta, \eta$ are used to balance the weights of different terms.

In this model, the first part $\mathcal{E}_{\mathrm{Seg}}(\Theta, \boldsymbol{u})$ is designed to segment the fixed image I_f through an extended GMM, which integrates intensity inhomogeneity and spatial regularization into the classical GMM. After explaining GMM-based methods in this variational formulation, it is more flexible to cooperate with other variational models, such as image registration, as shown in this framework.

The third part $\mathcal{E}_{\mathrm{Reg}}(\Theta, T)$ is a registration model that maps the moving image I_m to the intensity corrected image I_f/β. Here the registration model can come in various forms and be chosen according to the characteristics of the specific dataset. In this work, we just use a simple registration model (i.e., composed of a modified SSD and Tikhonov regularization) as an example to show the effectiveness of this joint framework.

The second part $\mathcal{E}_{\mathrm{CE}}(\boldsymbol{u}, \boldsymbol{s} \circ T)$ plays a bridge role between segmentation and registration, which enforces $\boldsymbol{u}$ (the segmentation map) and $\boldsymbol{s} \circ T$ (the normalized deformed tissue probability map) as similar as possible under the cross entropy metric. For the segmentation, $\boldsymbol{s} \circ T$ can be viewed as a prior to guide the segmentation process. In turn, the segmentation map $\boldsymbol{u}$ can promote the registration procedure by providing a reliable geometric structure, since the registration procedure considers not only the alignment of image intensities but also segmentation maps. Here, we choose cross entropy as the similarity metric for several reasons. First, as interpreted in subsection 2.4, this cross entropy constraint is equivalent to replacing the constant mixture ratio γ_k in GMM by tissue probability priors $\frac{s_k(x+T(x))}{\sum\limits_{l=1}^{K} s_l(x+T(x))}$, which has

a close relationship with the GMM-based segmentation model. Second, it can well measure the similarity between distributions and is widely used in machine learning. Last, the cross entropy is linear with respect to the segmentation map $\boldsymbol{u}$ and derivable with respect to the displacement field T, and thus won't bring additional difficulties in solving $\boldsymbol{u}$ and T.

Overall, we propose a joint model for segmentation and registration to leverage their positive mutual influence. Compared with other joint models, the proposed model has some advantages. First, for the segmentation component, it efficiently combines GMM with intensity correction (doesn't need an extra regularization for the bias field), spatial regularization (based on variational regularization), and adaptive spatial priors from variational-based registration. This model integrates the merits of probabilistic-based and variational-based methods and can be used for multiclass segmentation. Second, for the registration component, two levels of matching are considered, i.e., geometrical structure matching (between the segmentation map and the normalized deformed probability map) and intensity matching (between the intensity corrected image and the moving image). Thus this registration can provide more reliable priors and further promote the segmentation. Therefore, with intensity correction, spatial regularization, and spatial priors from joint registration, this proposed model has the potential to find an accurate segmentation of images which are degraded by intensity inhomogeneity, noise, and weak boundaries, such as thigh muscle MR images.

3.3. Optimization. In this section, the optimization scheme for minimizing objective function (3.5) is described. This is implemented within a constraint optimization framework, using the following iterated conditional modes:

$$
\begin{cases}
\Theta^{t+1} = \arg\min_{\Theta} \ \mathcal{E}(\Theta, \boldsymbol{u}^t, T^t) & \text{(i)}, \\
\boldsymbol{u}^{t+1} = \arg\min_{\boldsymbol{u} \in \mathbb{U}} \ \mathcal{E}(\Theta^{t+1}, \boldsymbol{u}, T^t) & \text{(ii)}, \\
T^{t+1} = \arg\min_{T} \ \mathcal{E}(\Theta^{t+1}, \boldsymbol{u}^{t+1}, T) & \text{(iii)},
\end{cases}
$$

where t is the iteration number.

(i) The Θ-subproblem is to solve the following optimization problem, which is similar to the M-step in EM algorithm.

$$
\begin{aligned}
\arg\min_{\Theta} \quad & \int_{\Omega}\int_{\Omega} G_\sigma(y-x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k\beta(y)]^2}{2\sigma_k^2\beta^2(y)} - \log\frac{\gamma_k}{\sqrt{2\pi}\sigma_k\beta(y)} \right) u_k^t(x) \right] dxdy \\
& + \frac{\zeta}{2}\int_{\Omega} \left[\frac{I_f(x)}{\beta(x)} - I_m(x + T^t(x)) \right]^2 dx.
\end{aligned}
$$

Recalling that $\sum_{k=1}^{K} \gamma_k^{t+1} = 1$, one can easily get the updating formulations,

$$
(3.6) \quad
\begin{cases}
\gamma_k^{t+1} = \dfrac{\int_{\Omega} u_k^t(x)dx}{\int_{\Omega} 1 dx}, \\[2ex]
c_k^{t+1} = \dfrac{\int_{\Omega} u_k^t(x)I_f(x)\int_{\Omega} G_\sigma(y-x)\frac{1}{\beta^t(y)}dydx}{\int_{\Omega} u_k^t(x)dx}, \\[2ex]
(\sigma_k^2)^{t+1} = \dfrac{\int_{\Omega} u_k^t(x)\int_{\Omega} G_\sigma(y-x)\left[\frac{I_f(x)}{\beta^t(y)} - c_k^{t+1}\right]^2 dydx}{\int_{\Omega} u_k^t(x)dx}, \\[2ex]
\beta^{t+1}(y) = \dfrac{-[s^{t+1}(y) + p^{t+1}(y)] + \sqrt{[s^{t+1}(y) + p^{t+1}(y)]^2 + 4[v^{t+1}(y) + \zeta I_f^2(y)]}}{2},
\end{cases}
$$

where

$$s^{t+1}(y) = \sum_{l=1}^{K} \frac{c_l^{t+1}}{(\sigma_l^2)^{t+1}} \int_\Omega G_\sigma(y-x) I_f(x) u_l^t(x) dx,$$

$$v^{t+1}(y) = \sum_{l=1}^{K} \frac{1}{(\sigma_l^2)^{t+1}} \int_\Omega G_\sigma(y-x) I_f^2(x) u_l^t(x) dx,$$

$$p^{t+1}(y) = \zeta I_f(y) I_m(y + T^t(y)).$$

(ii) The $\boldsymbol{u}$-subproblem is to optimize the following problem, which can be considered as the E-step (with a regularization and shape prior) in EM algorithm:

$$\underset{\boldsymbol{u} \in \mathbb{U}}{\arg\min} \quad \int_\Omega \int_\Omega G_\sigma(y-x) \left[\sum_{k=1}^{K} \left(\frac{[I_f(x) - c_k^{t+1}\beta^{t+1}(y)]^2}{2(\sigma_k^{t+1})^2(\beta^{t+1}(y))^2} - \log \frac{\gamma_k^{t+1}}{\sqrt{2\pi}\sigma_k^{t+1}\beta^{t+1}(y)} \right) u_k(x) \right] dx dy$$

$$+ \varepsilon \int_\Omega \sum_{k=1}^{K} u_k(x) \log u_k(x) dx + \lambda \int_\Omega \sum_{k=1}^{K} u_k(x)(\omega * (1 - 2u_k^t))(x) dx$$

$$- \xi \int_\Omega \sum_{k=1}^{K} u_k(x) \log \frac{s_k(x + T^t(x))}{\sum\limits_{l=1}^{K} s_l(x + T^t(x))} dx.$$

Here we used the linearization of the regularization term. Now the $\boldsymbol{u}$-subproblem is convex with respect to $\boldsymbol{u}$, and one can easily get the updating formulation

$$(3.7) \qquad u_k^{t+1}(x) = \frac{q_k^{t+1}(x)}{\sum\limits_{\ell=1}^{K} q_\ell^{t+1}(x)},$$

where

$$q_k^{t+1}(x) = \left(\frac{\gamma_k^{t+1}}{\sigma_k^{t+1}} \right)^{\frac{1}{\varepsilon}} \left(\frac{s_k(x + T^t(x))}{\sum\limits_{l=1}^{K} s_l(x + T^t(x))} \right)^{\frac{\xi}{\varepsilon}} \exp\left\{ -\frac{1}{2\varepsilon(\sigma_k^{t+1})^2} \int_\Omega G_\sigma(y-x) \left[\frac{I_f(x)}{\beta^{t+1}(y)} - c_k^{t+1} \right]^2 dy \right.$$

$$\left. - \frac{\lambda}{\varepsilon}(\omega * (1 - 2u_k^t))(x) \right\}.$$

(iii) The T-subproblem is to solve the following registration problem:

$$(3.8) \qquad \underset{T}{\arg\min} \quad -\xi \int_\Omega \sum_{k=1}^{K} u_k^{t+1}(x) \log \frac{s_k(x + T(x))}{\sum\limits_{l=1}^{K} s_l(x + T(x))} dx + \frac{\zeta}{2} \int_\Omega \left[\frac{I_f(x)}{\beta^{t+1}(x)} - I_m(x + T(x)) \right]^2 dx$$

$$+ \frac{\eta}{2} \int_\Omega |\nabla T(x)|^2 dx.$$

The partial derivative with respect to $T(x)$ is

$$
(3.9) \quad \frac{\partial \mathcal{E}(\Theta^{t+1}, \boldsymbol{u}^{t+1}, T)}{\partial T(x)} = -\xi \sum_{k=1}^{K} u_k^{t+1}(x) \frac{\nabla s_k(x + T(x))}{s_k(x + T(x))} + \xi \frac{\sum_{l=1}^{K} \nabla s_l(x + T(x))}{\sum_{l=1}^{K} s_l(x + T(x))}
$$

$$
+ \zeta \left[I_m(x + T(x)) - \frac{I_f(x)}{\beta^{t+1}(x)} \right] \nabla I_m(x + T(x)) - \eta \triangle T(x).
$$

It is a smooth optimization problem that can be solved using any gradient-based optimization technique. In this work, we solve it with the quasi-Newton limited-memory BFGS (LBFGS) method as in [43]. Since nonparametric registration methods are highly susceptible to local minima and not robust during optimization, we employ a parametric method to reduce the dimensionality of the registration optimization problem, thus providing physically plausible transformations in robust schemes with large displacement capture ranges. Concretely, as in [43] and [42], we parametrize the displacement field T using bilinear (for 2D images) or trilinear (for 3D images) interpolation with displacements D on control points that are placed with N-pixel spacing. Then we impose regularization on the displacement control grid points D instead of the displacement field T itself. In addition, a Gaussian pyramid is used to downsample images, and the registration is implemented from coarse-to-fine. At each consecutive image level, the displacement is initialized by interpolating from the previous level's control grid displacements. We implement the registration details based on a modified version of the MATLAB toolbox pTVreg [41].

In this optimization process, the first two steps (i) and (ii) are to solve the segmentation problem, and the last step (iii) is to solve the registration problem. Therefore, this process can be considered as alternating update segmentation and registration, which is summarized in Algorithm 3.1.

3.4. Modification for thigh muscle segmentation. In our proposed model (3.5), the cross entropy metric enforces the segmentation map $\boldsymbol{u}$ and the normalized deformed tissue probability map $\boldsymbol{s} \circ T$ as similar as possible for every class k ($k = 1, \cdots, K$). For different applications, the proposed model can be modified to accommodate specific needs.

In thigh muscle MR images, as shown in Figure 1, there are three main intensity classes: "dark" (background, cortical bone), "gray" (muscle), and "light" (fat, bone marrow) [2]. In our proposed model, the segmentation model $\mathcal{E}_{\text{Seg}}$ is an intensity-based method. So it is more suitable to classify a muscle image into three classes (dark, gray, light) but not four classes (one for each of the three muscle groups and one for others), which is actually the goal of muscle segmentation, while for the registration model, the classification is based on a deformed atlas as the shape prior, and thus it can distinguish different muscles and classify a muscle image into the desired four classes. Besides, in the proposed model, segmentation and registration play different roles. In some sense, the role of the segmentation model is to obtain accurate segmentation boundaries, and the registration model is to provide accurate segmentation categories, which can be illustrated by the later simulations in subsection 4.2. Therefore, the classification numbers of segmentation and registration could be different. In subsection 4.3,

Algorithm 3.1 The proposed joint segmentation and registration algorithm.

Input: I_f, I_m, s, ε, λ, ξ, ζ, η; L—number of image pyramid levels; M_{iter}—max number of alternate iterations; M_{LGMM}—max number of LGMM iterations; M_{LBFGS}—max number of LBFGS iterations; Tol—tolerance error.

Output: Segmentation map $\boldsymbol{u}$ or deformed tissue probability map $s \circ T$.

Initialization: u^0 by K-means, $\beta^0 = 1$, $T^0 = 0$.

   Compute objective function $\mathcal{E}^0$.

   for $t = 0, 1, 2, \ldots, M_{\text{iter}} - 1$ **do**

      $u^{t+1,0} = u^t$, $\beta^{t+1,0} = \beta^t$.

      for $i = 0, 1, 2, \ldots, M_{\text{LGMM}} - 1$ **do**

         Compute $\Theta^{t+1,i+1}$ by (3.6) with $u^{t+1,i}$, $\beta^{t+1,i}$, T^t.

         Compute $u^{t+1,i+1}$ by (3.7) with $u^{t+1,i}$, $\Theta^{t+1,i+1}$, T^t.

      end for

      $u^{t+1} = u^{t+1,M_{\text{LGMM}}}$, $\beta^{t+1} = \beta^{t+1,M_{\text{LGMM}}}$.

      Initialize parametric displacement field $D^{t+1,L}$(the coarsest level) with T^t.

      for $j = 1, \ldots, L$ **do**

         Compute I_m^j: $(L - j + 1)$th Gaussian pyramid level of I_m.

         Compute I_f^j: $(L - j + 1)$th Gaussian pyramid level of $\frac{I_f}{\beta^{t+1}}$.

         Compute s^j: $(L - j + 1)$th Gaussian pyramid level of s.

         Compute $u^{t+1,j}$: $(L - j + 1)$th Gaussian pyramid level of u^{t+1}.

         Compute $D^{t+1,L-j+1}$: solution of (3.8) with LBFGS optimization scheme (using I_m^j, I_f^j, s^j, $u^{t+1,j}$, M_{LBFGS}).

         Initialize $D^{t+1,L-j}$: linealy upsample $D^{t+1,L-j+1}$.

      end for

      Compute T^{t+1} with parametric displacement field $D^{t+1,1}$.

      Compute objective function $\mathcal{E}^{t+1}$.

      Convergence check. If $\frac{(\mathcal{E}^{t+1} - \mathcal{E}^t)^2}{(\mathcal{E}^t)^2} < Tol$, break.

   end for

   Segmentation function:

$$label(I_f(x)) = \arg \max_k \{u_k(x)\} \quad \text{or} \quad label(I_f(x)) = \arg \max_k \{s_k(x + T(x))\}.$$

   return Segmentation result $label(I_f(x))$.

for the thigh muscle segmentation, we set the classification number to be three for segmentation and four for registration.

In order to achieve this, we need to modify the cross entropy metric in the proposed model (3.5) to the following form:

$$\mathcal{E}_{\text{CE}}(u, s \circ T) = -\xi \int_\Omega \sum_{k=1}^2 p_k(x; u) \log q_k(x; s \circ T) dx,$$

where

$$p_1(x; u) = u_2(x), \quad p_2(x; u) = u_1(x) + u_3(x),$$

$$q_1(x; s \circ T) = \frac{s_1(x + T(x)) + s_2(x + T(x)) + s_3(x + T(x))}{\sum\limits_{l=1}^{4} s_l(x + T(x))}, \quad q_2(x; s \circ T) = \frac{s_4(x + T(x))}{\sum\limits_{l=1}^{4} s_l(x + T(x))}.$$

Here, according to intensity, we set u_1, u_2, u_3 to be the segmentation maps for "dark," "gray," and "light," respectively. Meanwhile, we set $s_1 + s_2 + s_3$, s_4 to be the GT segmentation maps of the moving image for "the whole muscle" and "others," respectively. Therefore, the modified cross entropy metric guarantees the similarity of the segmentation map and normalized deformed atlas for "the whole muscle" and "others." With this modification, it is not difficult to get all of the updating formulations. This also reflects the flexibility of our proposed model in specific applications.

4. Numerical experiments.

4.1. Parameter selection and implementation details.
For the segmentation component, the Gaussian kernel G_σ in the data fidelity term controls the smoothness of the bias field $\boldsymbol{\beta}$. Bigger σ leads to a smoother bias field. The entropy parameter ε controls the smoothness of the segmentation map $\boldsymbol{u}$. Small ε tends to lead to a binary segmentation, and large ε tends to lead to a smooth segmentation according to the maximum entropy principle. The kernel function ω in the regularization term controls the scale of the spatial regularization. The regularization parameter λ balances the data fidelity term, the regularization term, and the cross entropy term in the segmentation procedure, which is important for obtaining a desired segmentation result. In this paper, without specification, we set $\sigma = 20$, $\varepsilon = 1$, and the kernel function ω to be a 7×7 matrix with equal weights. The regularization parameter λ is image dependent, usually between 0.001 and 0.01.

The cross entropy parameter ξ acts as a bridge between segmentation and registration. According to our experience, the value between 0.001 and 0.03 is suitable for most of the cases.

For the registration component, the fidelity parameter ζ and the regularization parameter η control the weight of the modified SSD and the regularization term, respectively. In this paper, we set $\zeta = 1$ and make the value of η depend on images, usually between 0.001 and 0.05. For the Gaussian pyramid multiscales, the number of levels L can be determined by the size of the input images. In our experiments, we set $L = 8$ and 0.7 for the downscaling factor. For the displacement parameterization in each level, the control grid spacing (the gap in pixels between knots) controls the flexibility of the displacement field, and we set it to be 4 and use bilinear interpolation for 2D (or trilinear for 3D) image warping.

In our experiments, both the moving and the fixed images are converted to double precision and mapped to $[0, 1]$. We set the maximum iteration number $M_{\text{iter}} = 10$, $M_{\text{LGMM}} = 100$, $M_{\text{LBFGS}} = 100$, and the tolerance $Tol = 10^{-5}$. Although the proposed model is a joint segmentation and registration work, which is more complex and more difficult to implement than a sequential one (registration followed by segmentation), the running time is still acceptable due to several reasons. First, the convolution that appears in the equations can be efficiently calculated by FFT. Second, note that the proposed algorithm has inner and outer loops. Most of the computation time is spent on dealing with the first outer loop of segmentation

and registration. The following segmentation and registration processes take little time, as only a few number of inner loops are needed when given good initializations. The computation is done on an Intel processor (Core i5, 2GHz, 16GB RAM) with nonoptimized MATLAB code. The average computation time is about two minutes for a 3D MR image of size $256 \times 256 \times 15$, which is much less than manual segmentation.

4.2. Test on synthetic images. To show the validity of our proposed model, we first test it on a pair of synthetic images, which are shown in Figure 2(a) and (b). The moving image I_m and the fixed image I_f are both composed of a star shape and a heart shape, while the fixed image is polluted by noise, intensity inhomogeneity, and broken shapes. The segmentation task is to separate the two shapes from the background, individually, as shown in Figure 2(c) and (d). This is difficult for the fixed image as the two objects have similar intensities. Before giving the segmentation result of the fixed image by our proposed model with given (a), (b), and (c), we first do some numerical experiments to show the effectiveness of the components by turning off others in the proposed model.

The first experiment is to show the effectiveness of the segmentation component. Figure 3 shows the comparison of different segmentation energy terms; here "LEM" stands for the EM algorithm combined with bias correction, i.e., local EM, "RegEM" stands for the regularized EM algorithm, "LRegEM" stands for the local and regularized EM algorithm [20], and "LRegEMPrior" stands for the segmentation method of our proposed model, i.e., LRegEM combined with a GT prior provided by the cross entropy constraint. From images (c)–(l), we can clearly see that bias correction and regularization have great beneficial effects on segmentation. What's more, in the bottom-left of (g), (h), (k), and (l), one can see that the spatial regularization can greatly promote bias correction, i.e., the result of intensity correction depends on the promotion of a regularizer to segmentation. However, all of these methods shown in (c)–(l) cannot segment the two shapes from the background individually. As shown in (m) and (n), our proposed segmentation method can well accomplish this task, showing the validity of spacial priors provided by the cross entropy.

The second numerical experiment is to test the performance of the registration component. Figure 4 shows the registration results with different regularization parameters η. As can be seen from this figure, the smaller η is the more irregular the displacement field, and the bigger η is the smaller the displacement, which tends to keep the grid on its original location.
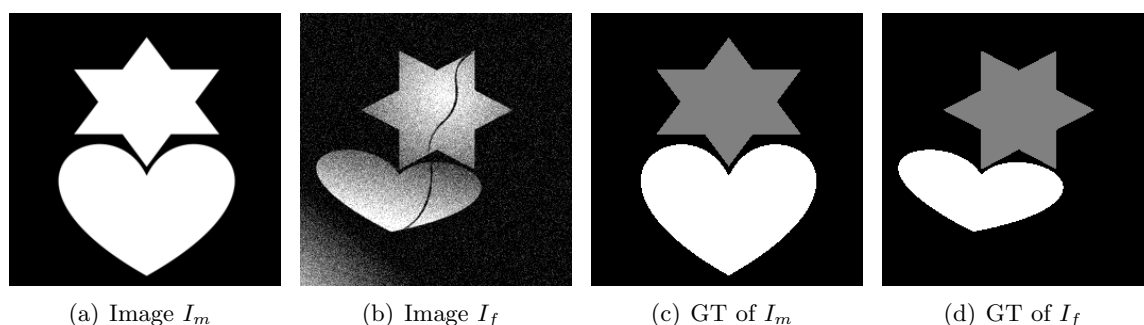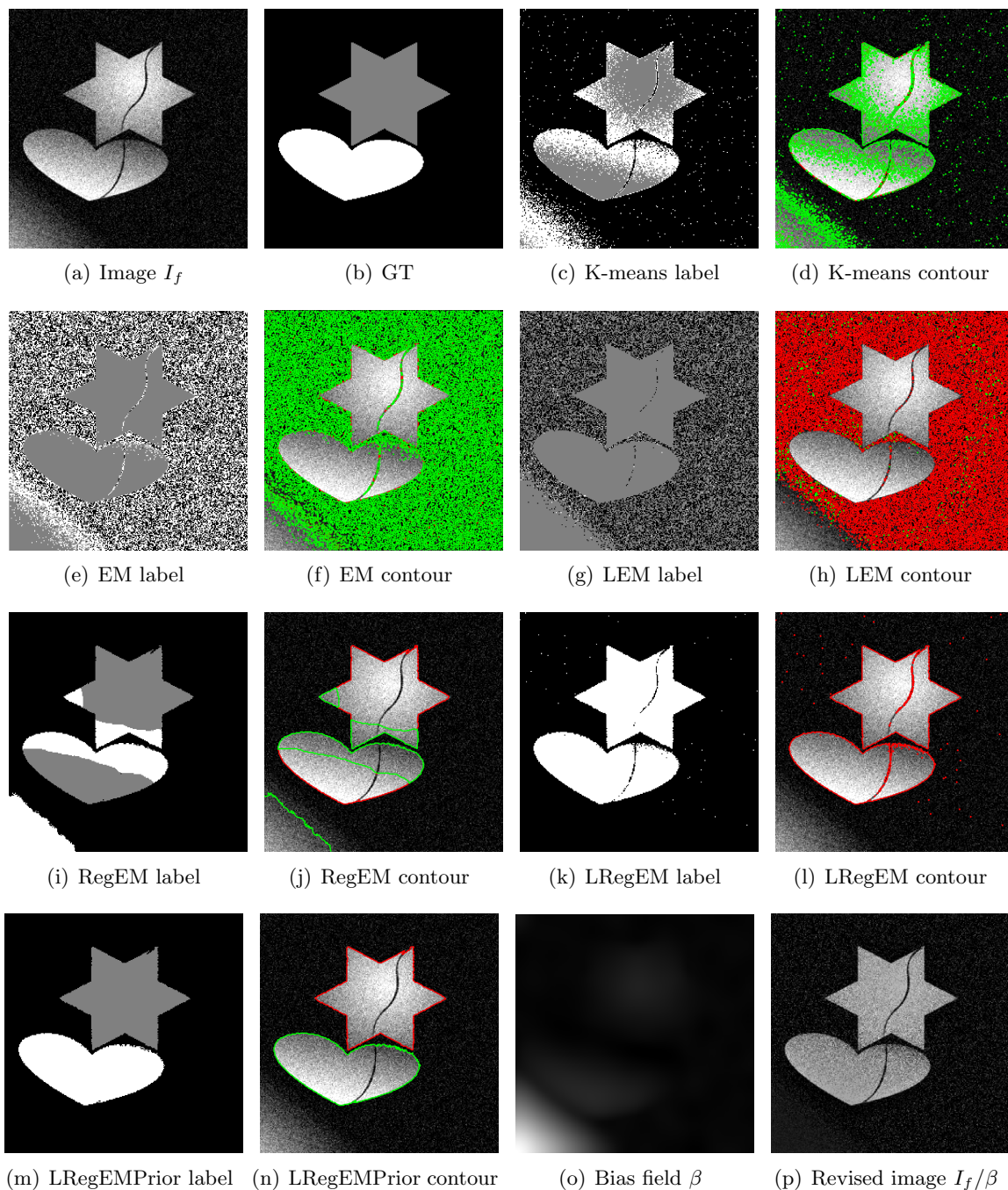

(a) Image I_m (b) Image I_f (c) GT of I_m (d) GT of I_f

Figure 2. *Synthetic images and the GT segmentations.*

(a) Image I_f (b) GT   (c) K-means label   (d) K-means contour

(e) EM label   (f) EM contour   (g) LEM label   (h) LEM contour

(i) RegEM label   (j) RegEM contour   (k) LRegEM label   (l) LRegEM contour

(m) LRegEMPrior label   (n) LRegEMPrior contour   (o) Bias field β (p) Revised image I_f/β

Figure 3. *Segmentation model test.* (a) *the fixed image* I_f, (b) *the GT segmentation of* I_f, (c)–(d) *K-means,* (e)–(f) *EM algorithm,* (g)–(h) *local EM algorithm, i.e., EM combined with bias correction,* (i)–(j) *regularized EM algorithm,* (k)–(l) *local and regularized EM algorithm* [20], (m)–(n) *proposed method, i.e., local and regularized EM algorithm combined with a GT prior provided by the cross entropy constraint,* (o)–(p) *estimated bias field and revised image* I_f/β *by the proposed segmentation model.*

(a) $\eta = 0$ (b) $\eta = 0.1$ (c) $\eta = 0.5$
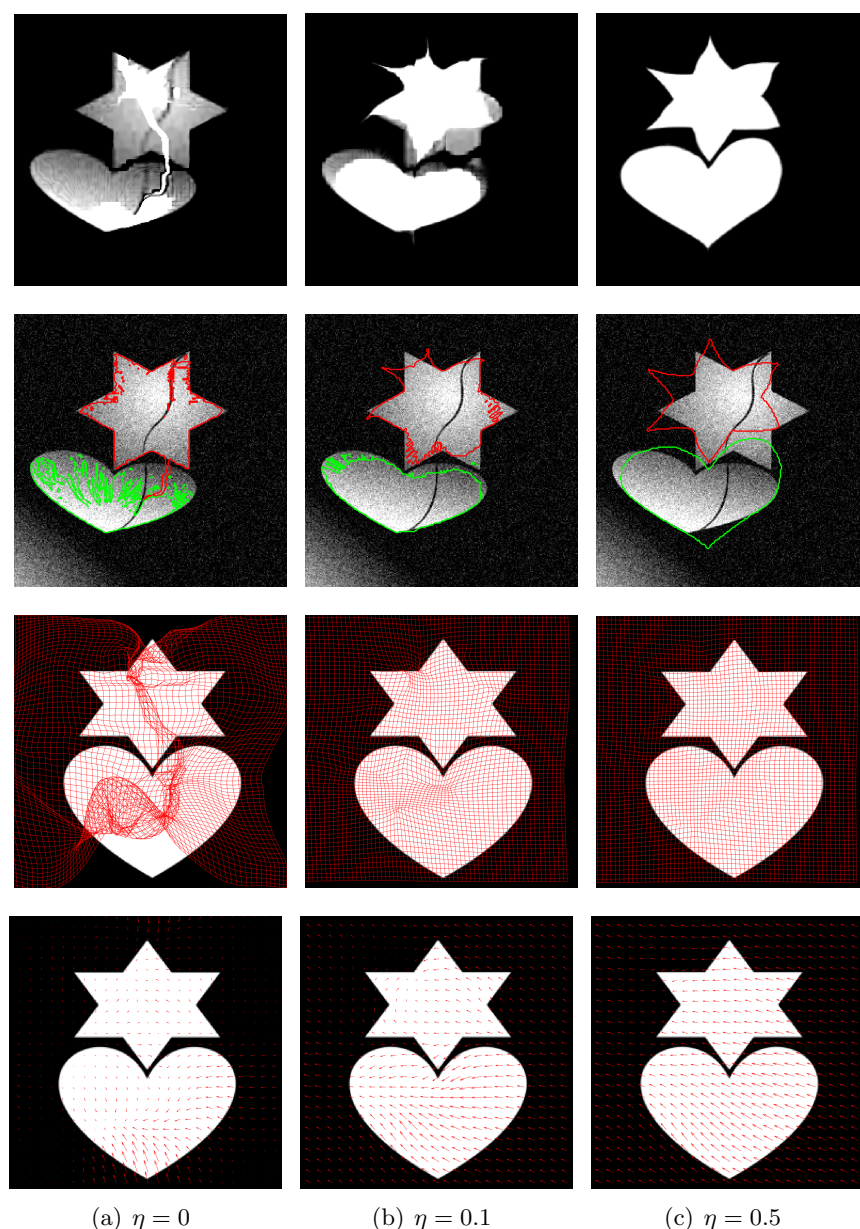
Figure 4. *Test of different regularization parameters η. First line: deformed image of I_m; second line: deformed contour; third line: deformed mesh grid; fourth line: displacement field.*

As mentioned in subsection 3.3, the registration is implemented from coarse level to fine level. In Figure 5, we show the registration results on odd levels, which indicates that the registration algorithm is stable and progressive.

The third experiment is to show the bridge function, cross entropy, between segmentation and registration. Figure 6 shows the results with different parameters ξ. We can see that

Level 9   Level 7   Level 5   Level 3   Level 1

Figure 5. *Registration results on odd levels. First line: deformed image of I_m; second line: deformed contour; third line: deformed mesh grid; fourth line: displacement field.*

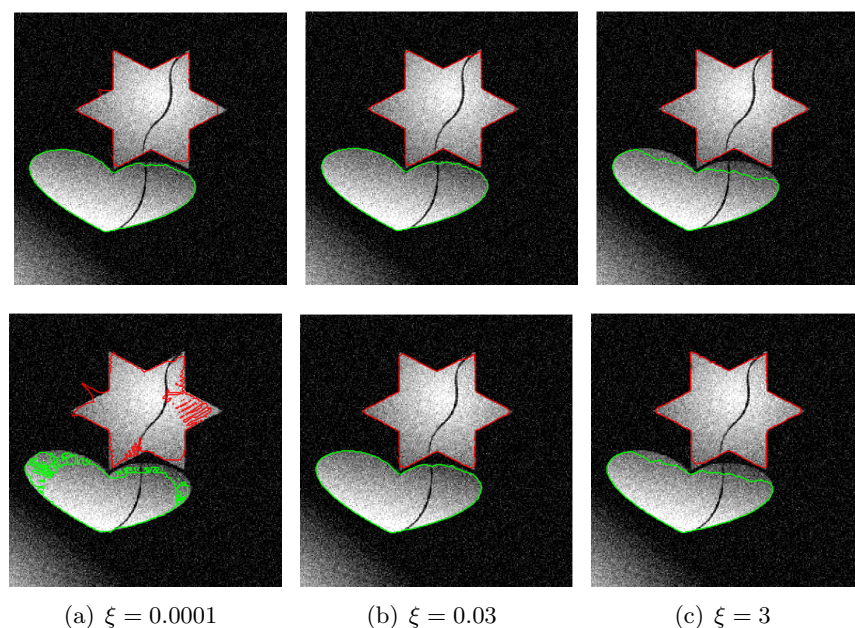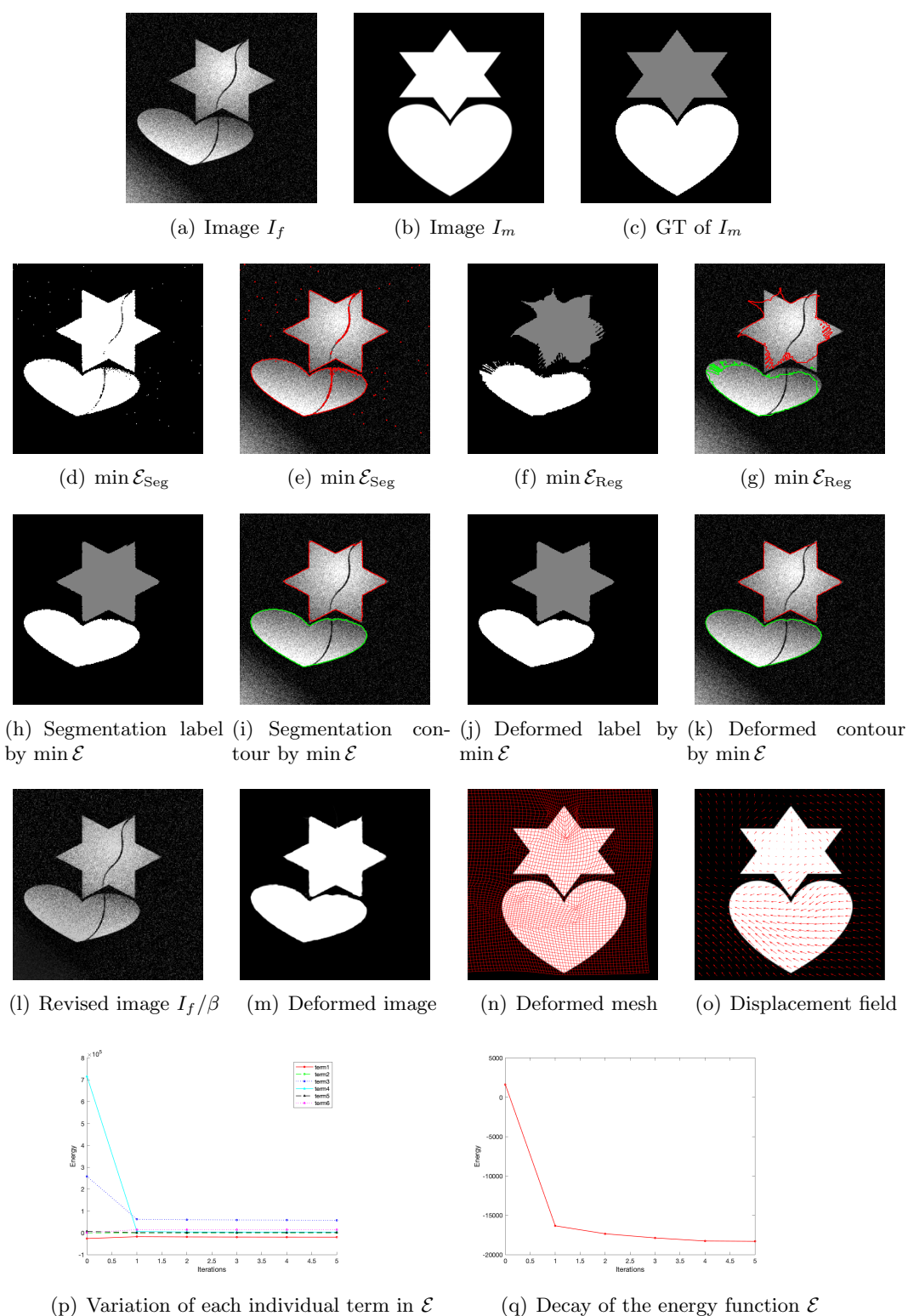when ξ is small, the segmentation result is less similar to the registered template. However, when ξ is too large, the strong interaction would force the segmentation result to be too close to the registered template and end up with missing the true boundaries.

(a) $\xi = 0.0001$ (b) $\xi = 0.03$ (c) $\xi = 3$

Figure 6. *Test of different cross entropy parameters ξ. First line: the contour of segmentation map $\boldsymbol{u}$; second line: the contour of deformed atlas by registration.*
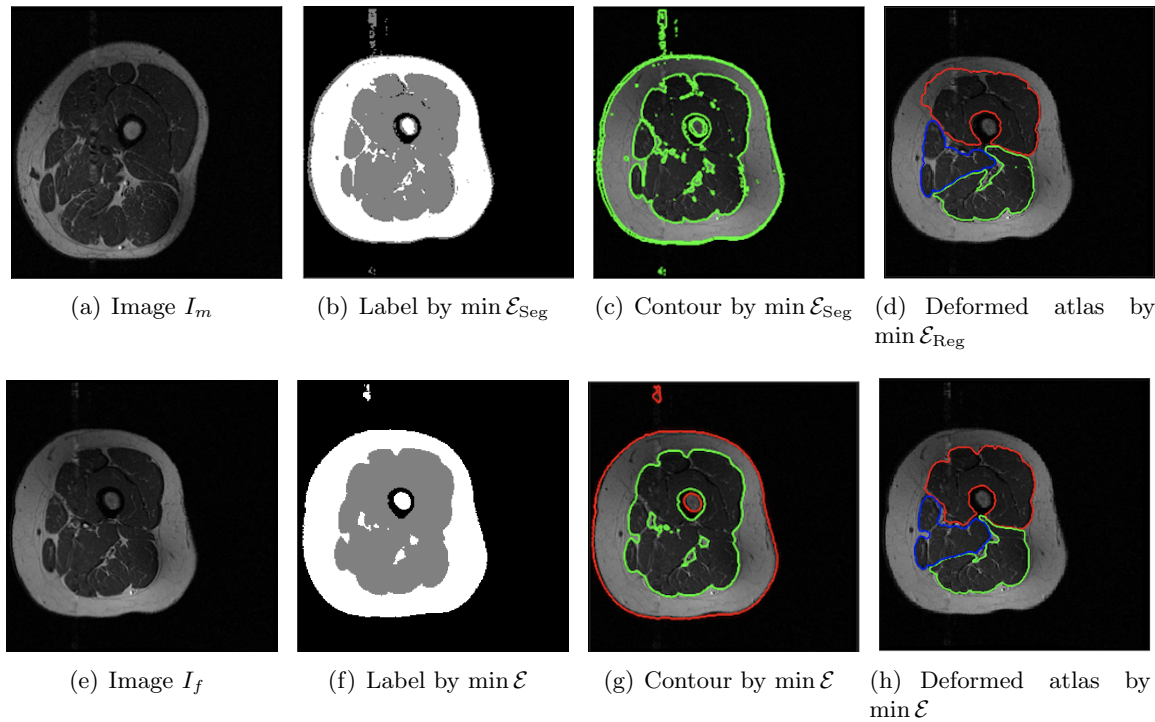
The fourth experiment is to illustrate the superiority of the proposed joint model compared with the segmentation only or registration only method. In Figure 7(d)–(e), we can see that the segmentation only method cannot separate the two shapes individually as they have similar intensities. Figure 7(f)–(g) show that the registration only method cannot well match the shapes of the objects. However, our proposed joint model can well accomplish this segmentation task, as shown in (h)–(k). Figure 7(l)–(q) display the bias corrected image, deformed image, deformed mesh, displacement field, and decay of the energy function of the proposed joint model.

4.3. Test on thigh muscle MR images. We evaluate our proposed model on five T1-weighted MRI 3D volumes. We choose one of them as the moving image which has GT segmentation by professional doctors through manual annotation and the other four as fixed images to be segmented for validation. The proposed algorithm is implemented on a MATLAB R2019b 9.7 environment. The four segmentation clusters are quadriceps (vastus medials, vastus lateralis, vastus intermedius, rectus femoris), hamstrings (semimembranosus, semitendinosus, biceps femoris long head, biceps femoris short head), other muscle groups (adductor group, gracilis, sartorius), and the rest (background, fat, cortical bone, bone marrow). For 3D thigh muscle data, we choose the middle axial cross section to visualize the segmentation results. Figure 8 shows the comparison of the proposed model with segmentation only or registration only method. The (b)–(c), (f)–(g) are segmentation results of minimizing $\mathcal{E}_{\mathrm{Seg}}$ and $\mathcal{E}$, respectively. We can see that minimizing $\mathcal{E}_{\mathrm{Seg}}$ can hardly separate the whole muscle from others, as the skin layer of the thigh is also divided into the muscle cluster. However,

(a) Image I_f (b) Image I_m (c) GT of I_m

(d) $\min \mathcal{E}_{\mathrm{Seg}}$ (e) $\min \mathcal{E}_{\mathrm{Seg}}$ (f) $\min \mathcal{E}_{\mathrm{Reg}}$ (g) $\min \mathcal{E}_{\mathrm{Reg}}$

(h) Segmentation label by $\min \mathcal{E}$ (i) Segmentation contour by $\min \mathcal{E}$ (j) Deformed label by $\min \mathcal{E}$ (k) Deformed contour by $\min \mathcal{E}$

(l) Revised image I_f / β (m) Deformed image   (n) Deformed mesh   (o) Displacement field

(p) Variation of each individual term in $\mathcal{E}$ (q) Decay of the energy function $\mathcal{E}$

Figure 7. *Comparison of the proposed joint model with segmentation or registration only method.*

(a) Image I_m (b) Label by $\min \mathcal{E}_{\mathrm{Seg}}$ (c) Contour by $\min \mathcal{E}_{\mathrm{Seg}}$ (d) Deformed atlas by $\min \mathcal{E}_{\mathrm{Reg}}$

(e) Image I_f (f) Label by $\min \mathcal{E}$ (g) Contour by $\min \mathcal{E}$ (h) Deformed atlas by $\min \mathcal{E}$
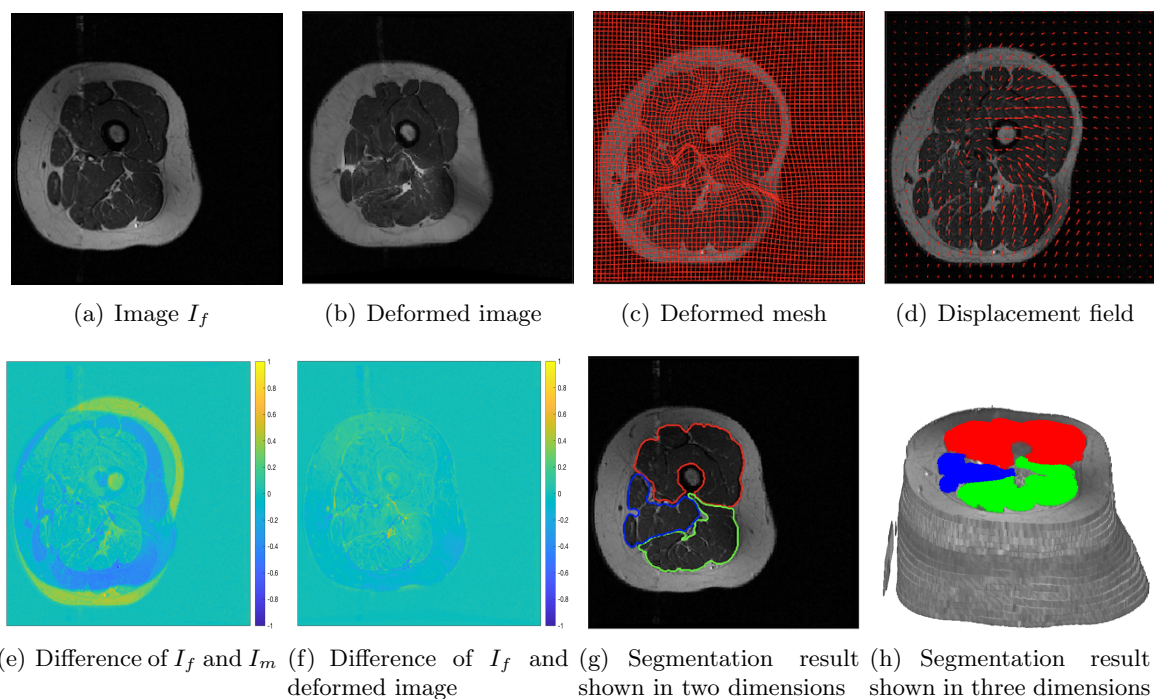
Figure 8. *Comparison of the proposed model with segmentation or registration only method.*

minimizing $\mathcal{E}$ can well separate out the whole muscle as a prior provided by the registration. Figure 8(d) and (h) are deformed atlas by minimizing $\mathcal{E}_{\mathrm{Reg}}$ and $\mathcal{E}$, respectively. We can clearly see that the proposed model obtains an accurate segmentation result of thigh muscles with the geometric structure provided by the segmentation, but the method with registration only can't. This figure further illustrates the advantages of a joint model. In the following experiments, without specification, we use the data shown in Figure 8(a) as the moving image. Figure 9 visualizes the results of our proposed model from different aspects. We can see that the deformed image is very similar with the fixed image and the segmentation result is accurate in the visual. Figure 10 lists the segmentation results of all slices by the proposed model, which shows that all of the cross sections have an accurate segmentation, thus illustrating the reliability and effectiveness of the proposed model.

In the last, we compare our proposed model with three other methods: a nonparametric diffeomorphic image registration algorithm based on Thirion's demons algorithm [40] (we denote it as M1), a nonrigid image registration based on B-spline composition and level sets [7] (denoted as M2), and an unified segmentation [3] introduced in subsection 2.2 (denoted as M3). For convenience, we denote our proposed model as M4 in the following. Note that M1 and M2 are both atlas-based segmentation methods as they depend on image registration only. M3 is another joint segmentation and registration method. Figures 12 and 13 show the segmentation results of these methods evaluated on two subjects (shown in Figure 11). These visual results demonstrate the superiority of the proposed method over others.

(a) Image I_f　　　(b) Deformed image　　　(c) Deformed mesh　　　(d) Displacement field

(e) Difference of I_f and I_m　(f) Difference of I_f and deformed image　(g) Segmentation result shown in two dimensions　(h) Segmentation result shown in three dimensions

Figure 9. *Results of the proposed model (the moving image I_m is shown in Figure 8(a)).*
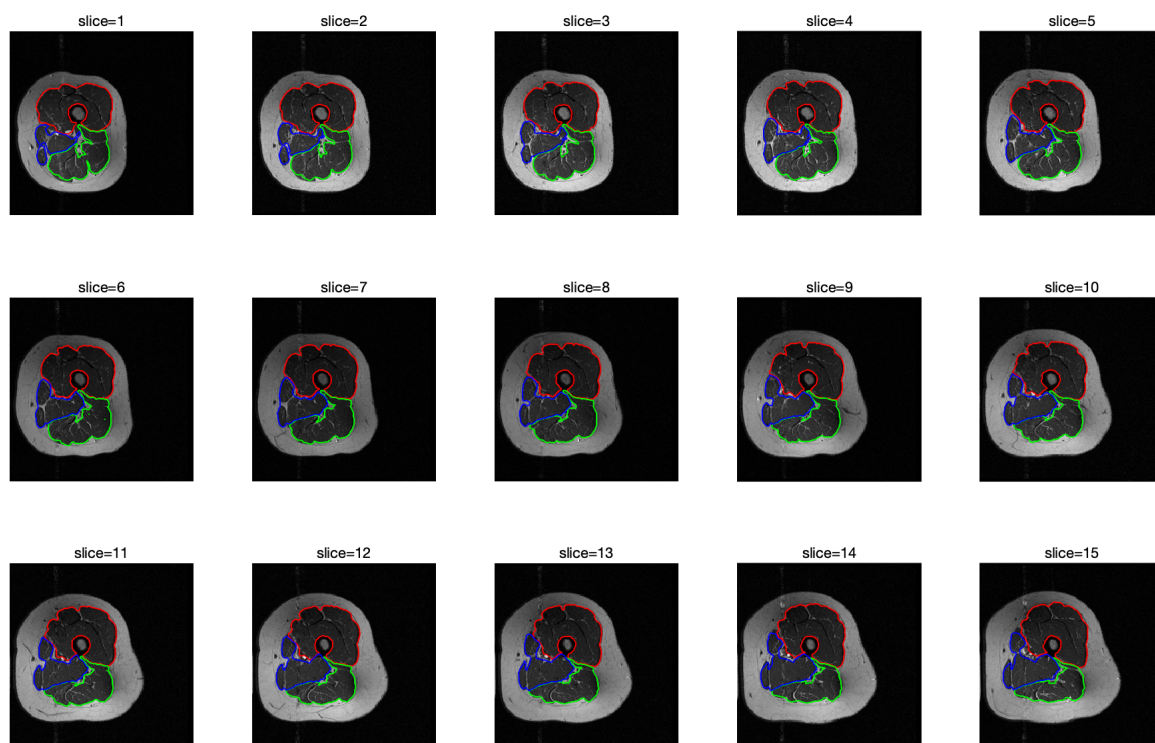
Figure 10. *The segmentation results of all slices by the proposed model.*

(a) Moving image   (b) Subject 1   (c) Subject 2
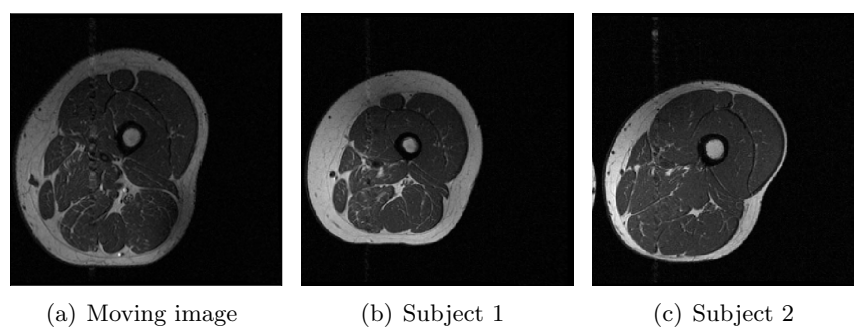
Figure 11. *Visualization of middle slices of the moving image, Subjects 1 and 2.*

(a) M1 [40]   (b) M2 [7]   (c) M3 [3]   (d) M4 (proposed)
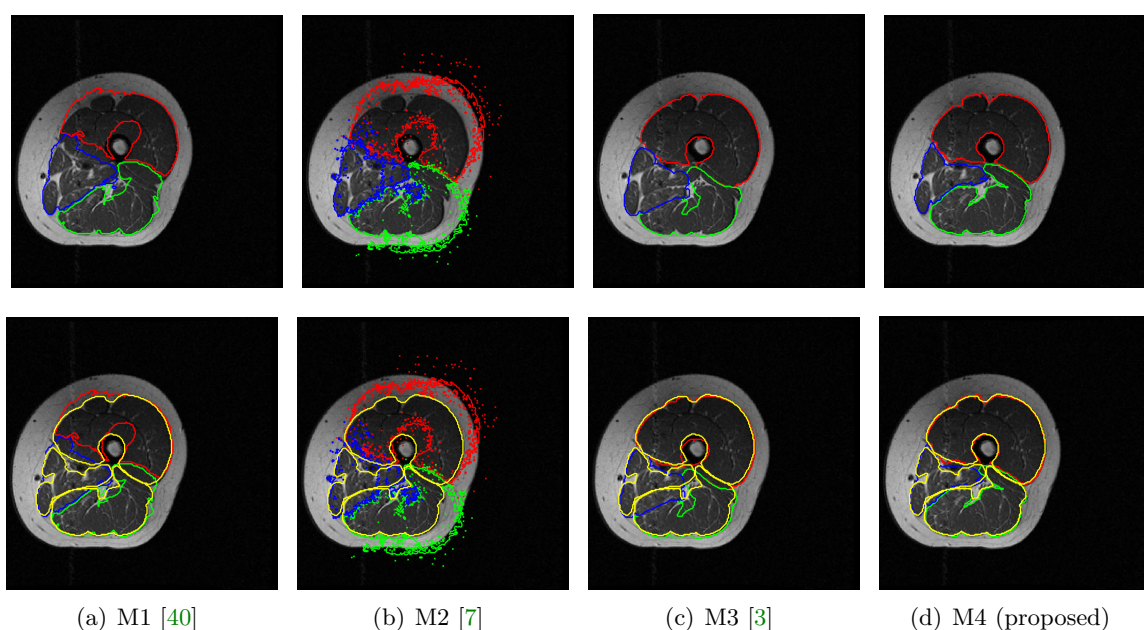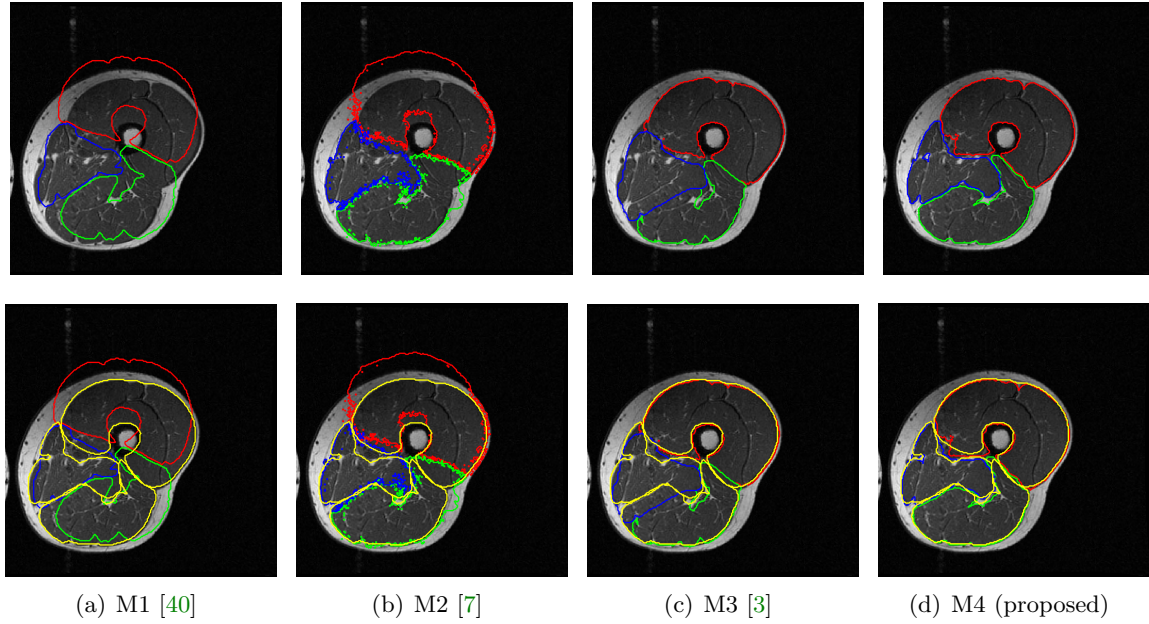
Figure 12. *Comparison of segmentation results of M1, M2, M3, and M4 on Subject 1 (see Figure 11). The first line: segmentation results of different methods. The second line: segmentation results with GTs overlaying in yellow color.*

To evaluate the performance of M1, M2, M3, and the proposed M4, we calculate the segmentation accuracy compared with GTs (obtained from manual segmentation). We adopt modified Jaccard (mJ), Dice similarity coefficient (DSC), and average surface distance (ASD) as the evaluation indexes. Let X be the segmentation result from the algorithm and Y the GT, ∂X the segmentation boundary of X, and ∂Y the boundary of Y. We have the following definitions:

- mJ:

$$\mathrm{mJ}(X, Y) = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|}.$$

(a) M1 [40]   (b) M2 [7]   (c) M3 [3]   (d) M4 (proposed)

Figure 13. *Comparison of segmentation results of M1, M2, M3, and M4 on Subject 2 (see Figure 11). The first line: segmentation results of different methods. The second line: segmentation results with GTs overlaying in yellow color.*

- DSC:

$$\mathrm{DSC}(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2\mathrm{TP}}{\mathrm{FN} + 2\mathrm{TP} + \mathrm{FP}}.$$

Here FP represents the number of false positives (i.e., the total number of the misclassified voxels of the background), FN is the number of false negatives (i.e., the total number of the misclassified voxels of the object), and TP is the true positive (i.e., total number of the correctly classified pixels).

- ASD:

$$\mathrm{ASD}(\partial X, \partial Y) = \frac{\sum\limits_{x \in \partial X} \min\limits_{y \in \partial Y} \|x - y\| + \sum\limits_{y \in \partial Y} \min\limits_{x \in \partial X} \|y - x\|}{|\partial X| + |\partial Y|}.$$

The means and standard deviations of mJ, DSC, and ASD by methods M1, M2, M3, and M4, on four data sets, are listed in Table 1, respectively. This table shows that our proposed method has the highest mJ and DSC scores and smallest ASD, which illustrates the superiority of the proposed method from the perspective of numerical evaluation.

5. Conclusion and discussion. We propose a novel image segmentation model with adaptive spatial priors from joint registration, which can exploit the strong correlation between segmentation and registration, thus achieving more accurate results than sequential treatment. Besides, this framework absorbs the merits of both variational and statistical methods. The segmentation process combines GMM with spatial smoothness, intensity inhomogeneity, and shape prior from registration under a variational framework, and the registration process

Table 1

mJ, DSC, and ASD (mean ± std) for M1 [40], M2 [7], M3 [3], and M4 (proposed).

		Quadriceps	Hamstrings	Other muscles	Average
M1	mJ	0.3627 ± 0.2139	0.5203 ± 0.1027	0.4563 ± 0.1439	0.4464 ± 0.1462
	DSC	0.5425 ± 0.2352	0.6602 ± 0.0913	0.5987 ± 0.1523	0.6005 ± 0.1527
	ASD	5.9953 ± 2.7591	4.8623 ± 1.7181	4.6537 ± 2.8222	5.1704 ± 2.4203
M2	mJ	0.4421 ± 0.1521	0.4933 ± 0.0811	0.3975 ± 0.1783	0.4443 ± 0.1322
	DSC	0.6007 ± 0.1570	0.6578 ± 0.0696	0.5522 ± 0.1749	0.6036 ± 0.1289
	ASD	6.1155 ± 2.3083	6.0264 ± 1.6284	5.3183 ± 3.2121	5.8201 ± 2.3625
M3	mJ	0.7834 ± 0.1077	0.6070 ± 0.1167	0.4741 ± 0.1526	0.6215 ± 0.1228
	DSC	0.8753 ± 0.0719	0.7502 ± 0.0965	0.6313 ± 0.1539	0.7523 ± 0.1055
	ASD	2.5174 ± 1.0276	4.3736 ± 1.2004	4.7323 ± 2.0927	3.8744 ± 1.3975
M4	mJ	$\mathbf{0.9190 \pm 0.0071}$	$\mathbf{0.8828 \pm 0.0181}$	$\mathbf{0.7594 \pm 0.0412}$	$\mathbf{0.8537 \pm 0.0154}$
	DSC	$\mathbf{0.9578 \pm 0.0039}$	$\mathbf{0.9377 \pm 0.0102}$	$\mathbf{0.8628 \pm 0.0261}$	$\mathbf{0.9194 \pm 0.0097}$
	ASD	$\mathbf{0.8171 \pm 0.1185}$	$\mathbf{1.1712 \pm 0.1606}$	$\mathbf{1.4448 \pm 0.1751}$	$\mathbf{1.1444 \pm 0.0336}$

considers two levels matching, thus making more accurate results. The numerical experiments evaluated on synthetic and thigh muscle MR images demonstrate the superiority of this proposed model. The application to automatic segmentation of thigh muscle only needs a small number of manually segmented label mapping, which is significant for clinical research as obtaining manual segmentation of thigh muscle is a time-consuming task.

There are also some aspects that can be further improved. For example, the registration model can be designed more precisely according to the characteristics of specific dataset, such as considering diffeomorphic registration or adding edge information. We just take a simple model as an example, and more elaborate registration may obtain further benefits. Moreover, the variability among individuals can be significant in medical images, and a more meaningful atlas will be welcomed. In this paper, we simply use one person's GT segmentation as the atlas. This can be improved by choosing a best one from multiple atlases or constructing an "average" atlas, which is a meaningful statistical atlas of the global underlying anatomy of thigh muscle from a set of atlas [9]. A more excellent registration method and more meaningful atlas would benefit the segmentation result, which will be considered in our future work.

Appendix A. Proof of Theorem 3.2. Since $\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^{t+1}) = \mathcal{R}(\boldsymbol{u}^{t+1})$ and $\mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t) = \mathcal{R}(\boldsymbol{u}^t)$, to prove $\hat{\mathcal{E}}(\Theta^{t+1}, \boldsymbol{u}^{t+1}) \leq \hat{\mathcal{E}}(\Theta^t, \boldsymbol{u}^t)$ is to prove

$$\mathcal{F}(\Theta^{t+1}, \boldsymbol{u}^{t+1}) + \lambda\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^{t+1}) \leq \mathcal{F}(\Theta^t, \boldsymbol{u}^t) + \lambda\mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t).$$

According to the second and first formulation of iteration scheme (3.4), we have

$$\mathcal{F}(\Theta^{t+1}, \boldsymbol{u}^{t+1}) + \lambda\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^t) \leq \mathcal{F}(\Theta^{t+1}, \boldsymbol{u}^t) + \lambda\mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t) \leq \mathcal{F}(\Theta^t, \boldsymbol{u}^t) + \lambda\mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t).$$

Next, it suffices to prove that

$$\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^{t+1}) \leq \mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^t).$$

It is easy to check that $\mathcal{R}(\boldsymbol{u})$ is concave if the kernel function ω is semipositive definite. The variation of $\mathcal{R}(\boldsymbol{u})$ is $\delta\mathcal{R}(\boldsymbol{u}) = \omega * (1 - 2\boldsymbol{u})$ if ω is a symmetric kernel function such as a Gaussian kernel. By the property of a concave function, one can have

$$\mathcal{R}(\boldsymbol{u}^{t+1}) - \mathcal{R}(\boldsymbol{u}^t) \leq \langle \boldsymbol{u}^{t+1} - \boldsymbol{u}^t, \omega * (1 - 2\boldsymbol{u}^t) \rangle.$$

That is,

$$\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^{t+1}) - \mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t) \leq \langle \boldsymbol{u}^{t+1} - \boldsymbol{u}^t, \omega * (1 - 2\boldsymbol{u}^t) \rangle = \mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^t) - \mathcal{R}(\boldsymbol{u}^t; \boldsymbol{u}^t).$$

Thus we have

$$\mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^{t+1}) \leq \mathcal{R}(\boldsymbol{u}^{t+1}; \boldsymbol{u}^t),$$

which completes the proof. ∎

REFERENCES

[1] E. Ahmad, M. H. Yap, H. Degens, and J. S. McPhee, *Atlas-registration based image segmentation of MRI human thigh muscles in* 3*D space*, in Proceedings of Medical Imaging 2014: Image Perception, Observer Performance, and Technology Assessment, SPIE, 2014, pp. 424–435.

[2] S. Andrews and G. Hamarneh, *The generalized log-ratio transformation: Learning shape and adjacency priors for simultaneous thigh muscle segmentation*, IEEE Trans. Med. Imaging, 34 (2015), pp. 1773–1787.

[3] J. Ashburner and K. J. Friston, *Unified segmentation*, Neuroimage, 26 (2005), pp. 839–851.

[4] T. Atta-Fosu and W. Guo, *Joint segmentation and nonlinear registration using fast Fourier transform and total variation*, in Research in Shape Analysis, Assoc. Women Math. Ser. 12, A. Genctav, K. Leonard, S. Tari, E. Hubert, G. Morin, N. El-Zehiry, and E. Chambers, eds., Springer, New York, 2018, pp. 111–132.

[5] P. Y. Baudin, N. Azzabou, P. G. Carlier, and N. Paragios, *Prior knowledge, random walks and human skeletal muscle segmentation*, in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, New York, 2012, pp. 569–576.

[6] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, 1995.

[7] C. L. Chan, C. Anitescu, Y. Zhang, and T. Rabczuk, *Two and three dimensional image registration based on B-spline composition and level sets*, Commun. Comput. Phys., 21 (2017), pp. 600–622.

[8] Y. Chen, W. Guo, F. Huang, D. Wilson, and E. A. Geiser, *Using prior shape and points in medical image segmentation*, in Proceedings of the International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, Springer, New York, 2003, pp. 291–305.

[9] N. Debroux, J. Aston, F. Bonardi, A. Forbes, C. L. Guyader, M. Romanchikova, and C. B. Schonlieb, *A variational model dedicated to joint segmentation, registration, and atlas generation for shape analysis*, SIAM J. Imaging Sci., 13 (2020), pp. 351–380.

[10] N. Debroux and C. L. Guyader, *A joint segmentation/registration model based on a nonlocal characterization of weighted total variation and nonlocal shape descriptors*, SIAM J. Imaging Sci., 11 (2018), pp. 957–990.

[11] S. Esedo–Lu and F. Otto, *Threshold dynamics for networks with arbitrary surface tensions*, Comm. Pure Appl. Math., 68 (2015), pp. 808–864.

[12] S. Ghosh, N. Ray, and P. Boulanger, *A structured deep-learning based approach for the automated segmentation of human leg muscle from* 3*D MRI*, in Proceedings of the 14th Conference on Computer and Robot Vision, IEEE, 2017, pp. 117–123.

[13] A. Gooya, K. M. Pohl, M. Bilello, L. Cirillo, G. Biros, E. R. Melhem, and C. Davatzikos, *GLISTR: Glioma image segmentation and registration*, IEEE Trans. Med. Imaging, 31 (2012), pp. 1941–1954.

[14] Z. Guo, H. Zhang, Z. Chen, E. van der Plas, L. Gutmann, D. Thedens, P. Nopoulos, and M. Sonka, *Fully automated* 3*D segmentation of MR-imaged calf muscle compartments: Neighborhood relationship enhanced fully convolutional network*, Comput. Med. Imaging Graph., 87 (2021), 101835.

[15] C. L. Guyader and L. A. Vese, *A combined segmentation and registration framework with a nonlinear elasticity smoother*, Comput. Vis. Image Underst., 115 (2011), pp. 1689–1709.

[16] E. Jolivet, E. Dion, P. Rouch, G. Dubois, R. Charrier, C. Payan, and W. Skalli, *Skeletal muscle segmentation from MRI dataset using a model-based approach*, Comput. Methods Biomech. Biomed. Eng. Imaging Vis., 2 (2014), pp. 138–145.

[17] J. Kemnitz, C. F. Baumgartner, F. Eckstein, A. Chaudhari, A. Ruhdorfer, W. Wirth, S. K. Eder, and E. Konukoglu, *Clinical evaluation of fully automated thigh muscle and adipose tissue segmentation using a U-Net deep learning architecture in context of osteoarthritic knee pain*, Magn. Reson. Mater. Phy., 33 (2020), pp. 483–493.

[18] J. Kemnitz, F. Eckstein, A. G. Culvenor, A. Ruhdorfer, T. Dannhauer, S. Ring-Dimitriou, A. M. Sänger, and W. Wirth, *Validation of an active shape model-based semi-automated segmentation algorithm for the analysis of thigh muscle and adipose tissue cross-sectional areas*, Magn. Reson. Mater. Phy., 30 (2017), pp. 489–503.

[19] J. Liu, X. Wang, and X.-C. Tai, *Deep convolutional neural networks with spatial regularization, volume and star-shape priors for image segmentation*, J. Math. Imaging Vision, 2022, pp. 1–21.

[20] J. Liu and H. Zhang, *Image segmentation using a local GMM in a variational framework*, J. Math. Imaging Vision, 46 (2013), pp. 161–176.

[21] J. M. Lötjönen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, D. Rueckert, and A. D. N. Initiative, *Fast and robust multi-atlas segmentation of brain magnetic resonance images*, Neuroimage, 49 (2010), pp. 2352–2365.

[22] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, New York, 2007.

[23] G. J. McLachlan, S. X. Lee, and S. I. Rathnayake, *Finite mixture models*, Annu. Rev. Stat. Appl., 6 (2019), pp. 355–378.

[24] B. Merriman, J. K. Bence, and S. Osher, *Diffusion Generated Motion by Mean Curvature*, Department of Mathematics, University of California, Los Angeles, 1992.

[25] B. Merriman, J. K. Bence, and S. Osher, *Motion of multiple junctions: A level set approach*, J. Comput. Phys., 112 (1994), pp. 334–363.

[26] S. Mesbah, A. M. Shalaby, S. Stills, A. M. Soliman, A. Willhite, S. J. Harkema, E. Rejc, and A. S. El-Baz, *Novel stochastic framework for automatic segmentation of human thigh MRI volumes and its applications in spinal cord injured individuals*, PLoS One, 14 (2019), e0216487.

[27] M. Molaie and R. A. Zoroofi, *A knowledge-based modality-independent technique for concurrent thigh muscle segmentation: Applicable to CT and MR images*, J. Digit. Imaging, 33 (2020), pp. 1122–1135.

[28] R. Ni, C. H. Meyer, S. S. Blemker, J. M. Hart, and X. Feng, *Automatic segmentation of all lower limb muscles from high-resolution magnetic resonance imaging using a cascaded three-dimensional deep convolutional neural network*, J. Med. Imaging, 6 (2019), 044009.

[29] A. Ogier, L. Heskamp, C. P. Michel, A. Fouré, M.-E. Bellemare, A. Le Troter, A. Heerschap, and D. Bendahan, *A novel segmentation framework dedicated to the follow-up of fat infiltration in individual muscles of patients with neuromuscular disorders*, Magn. Reson. Med., 83 (2020), pp. 1825–1836.

[30] A. Ogier, M. Sdika, A. Foure, A. Le Troter, and D. Bendahan, *Individual muscle segmentation in MR images: A 3D propagation through 2D non-linear registration approaches*, in Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2017, pp. 317–320.

[31] A. Pawar, Y. J. Zhang, C. Anitescu, and T. Rabczuk, *Joint image segmentation and registration based on a dynamic level set approach using truncated hierarchical B-splines*, Comput. Math. Appl., 78 (2019), pp. 3250–3267.

[32] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, *A Bayesian model for joint segmentation and registration*, NeuroImage, 31 (2006), pp. 228–239.

[33] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[34] D. Rueckert and P. Aljabar, *Non-rigid registration using free-form deformations*, in Handbook of Biomedical Imaging, N. Paragios, J. Duncan, and N. Ayache, eds., Springer, New York, 2015, pp. 277–294.

[35] T. W. Sederberg and S. R. Parry, *Free-form deformation of solid geometric models*, in Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques, ACM, 1986, pp. 151–160.

[36] M. K. Sharma, M. Jas, V. Karale, A. Sadhu, and S. Mukhopadhyay, *Mammogram segmentation using multi-atlas deformable registration*, Comput. Biol. Med., 110 (2019), pp. 244–253.

[37] A. Sotiras, C. Davatzikos, and N. Paragios, *Deformable medical image registration: A survey*, IEEE Trans. Med. Imaging, 32 (2013), pp. 1153–1190.

[38] P. Swierczynski, B. W. Papież, J. A. Schnabel, and C. Macdonald, *A level-set approach to joint image segmentation and registration with application to CT lung imaging*, Comput. Med. Imaging Graph., 65 (2018), pp. 58–68.

[39] A. L. Troter, A. Fouré, M. Guye, S. Confort-Gouny, J.-P. Mattei, J. Gondin, E. Salort-Campana, and D. Bendahan, *Volume measurements of individual muscles in human quadriceps femoris using atlas-based segmentation approaches*, Magn. Reson. Mater. Phys. Biol. Med., 29 (2016), pp. 245–257.

[40] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, *Diffeomorphic demons: Efficient non-parametric image registration*, NeuroImage, 45 (2009), pp. S61–S72.

[41] V. Vishnevskiy, *pTVreg Toolbox*, https://github.com/visva89/pTVreg, 2018.

[42] V. Vishnevskiy, T. Gass, G. Székely, and O. Goksel, *Total variation regularization of displacements in parametric image registration*, in Proceedings of the International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging, Springer, 2014, pp. 211–220.

[43] V. Vishnevskiy, T. Gass, G. Szekely, C. Tanner, and O. Goksel, *Isotropic total variation regularization of displacements in parametric image registration*, IEEE Trans. Med. Imaging, 36 (2016), pp. 385–395.

[44] P. P. Wyatt and J. A. Noble, *MAP MRF joint segmentation and registration of medical images*, Med. Image Anal., 7 (2003), pp. 539–552.

[45] A. Yezzi, L. Zollei, and T. Kapur, *A variational framework for joint segmentation and registration*, in Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, IEEE, 2001, pp. 44–51.

[46] F. Yokota, Y. Otake, M. Takao, T. Ogawa, T. Okada, N. Sugano, and Y. Sato, *Automated muscle segmentation from CT images of the hip and thigh using a hierarchical multi-atlas method*, Int. J. Comput. Assist. Radiol. Surg., 13 (2018), pp. 977–986.

[47] B. Zitova and J. Flusser, *Image registration methods: A survey*, Image Vision Comput., 21 (2003), pp. 977–1000.