MACHINE LEARNING BASED IMAGE SEGMENTATION FOR LARGE-SCALE OSTEOARTHRITIS ANALYSIS

By

CHAOWEI TAN

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Computer Science

Written under the direction of

Dimitris N. Metaxas

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2020

ABSTRACT OF THE DISSERTATION

Machine Learning based Image Segmentation for Large-scale Osteoarthritis Analysis

by CHAOWEI TAN

Dissertation Director:

Dimitris N. Metaxas

Osteoarthritis (OA) is the most common degenerative joint disease worldwide, tending to occur in the joints of hip and knee. Large adult population in the United States have been affected by OA, and by 2030, an estimated 20 percent of Americans (about 70 million people) may be at increased risk for this disease. Effective medical image segmentation methods play fundamental roles in the clinical analysis of the disease. In this dissertation, three machine learning based segmentation for knee cartilage, femoral head-neck junction and thigh muscular/adipose tissue are discussed, respectively. Furthermore, large-scale OA analysis on the knee and hip joints could be further implemented based on these segments.

Knee cartilages (i.e., femoral, tibial, and patellar cartilage) are essential tissue for

knee radiographic OA diagnosis. Effective segmentation of knee cartilages in large-sized and high-resolution 3D magnetic resonance (MR) data is firstly proposed. The key contribution is an adversarial learning based collaborative multi-agent network. The method employs three parallel segmentation agents to label cartilages in their respective region of interest (ROI), and then fuses the three cartilages by a ROI-fusion layer and drive a collaborative learning by an adversarial sub-network. The ROI-fusion layer not only fuses the individual cartilages, but also backpropagates the training loss from the adversarial sub-network to each agent to enable joint learning of shape and spatial constraints. The proposed scheme is shown robust and accurate in knee cartilage segmentation, and it is effective for cartilage biomarkers (e.g., surface area, volume) estimation in large-scale quantitative tests. Second, a deep multi-task learning network is exploited for the shape-preserved segmentation of the proximal part of femur (i.e., femoral head and neck) in 2D MR images. This method combines the tasks of region identification and boundary distance regression, and thus enables the task-specific feature learning for continuous segmented object with smooth boundary. This bone joint depiction could help the measurements of the femoral head-neck morphology and reflect the evolution of hip OA. In the last part of the dissertation, the muscular and adipose tissue extraction in 3D MR thigh data is investigated by an integrated framework. Specifically, deformable models and learning based detection/classification are integrated into the framework to enable robust tissue quantification for a large-scale analysis of OA-related thigh tissue changes.

Acknowledgements

First, I would like to express my sincere gratitude to my advisor, Professor Dimitris N. Metaxas, for his kind guidance during my Ph.D. study. Prof. Metaxas provided excellent research environment to explore challenging tasks in medical image analysis. None of the work in this dissertation would have been achieved without his support.

My sincere thanks also goes to my co-advisor, Prof. Kang Li, for his mentorship in the orthopedics-related research. Without his precious support, it would be very difficult to conduct the innovative work in orthopedics.

Besides my advisor and co-advisor, I would like to thank the rest members in my thesis committee: Prof. Jingjin Yu, and Dr. Xin Dou, for their insightful comments and advice regarding this thesis. It is my great honor to invite all of them in my committee and help me to improve my research.

Moreover, I would like to thank Prof. Shaoting Zhang's profound inspiration during my Ph.D. study. I also appreciate other collaborators, especially to Dr. Zhennan Yan, Dr. Yiqiang Zhan, Dr. Liang Zhao, Dr. Shu Liao, Dr. Boubakeur Belaroussi and Dr. Hui Jing Yu. They provided me the internship opportunities to join their teams, and inspired my ideas. I would like to thank all the labmates in Rutgers CBIM, especially to Dr. Bo Liu, Dr. Jingjing Liu and Dr. Dong Yang. Thanks for the helpful discussions.

Last but not least, I never forget to thank my families to support and encourage me in the unforgettable Ph.D. journey.

Dedication

This dissertation is dedicated to my wife, my children, my parents and parents-in-law.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Osteoarthritis (OA), the most prevalent arthritis, is a type of disease on degenerative joints, including articular cartilage and subchondral bone. OA causes mechanical abnormalities of joints, and then results in activity limitation and physical disability [71]. OA may affect any joint in the body, but it is particularly common in the knee and hip joints. In the United States at the present day, more than 30 million U.S. adults have the radiological evidence of OA [21, 80]. By 2030, an estimated 20 percent of Americans (about 70 million people) may be at increased risk for this disease [86].

As the development of computing technologies, computer-aided image processing is fast growing and has been widely utilized in medical imaging in recent decades. Medical imaging is the technique of generating visual signal/image information to represent the interior of a body for medical analysis, clinical intervention and physiological presentation [124]. Magnetic resonance (MR) is one of the medical imaging technologies in radiology to form the anatomical images and physiological procedures [123]. MR scanners employ strong magnetic fields to generate organic pictures of in the body, and thus this imaging is non-invasive. MR technique can produce three dimensional (3D) detailed anatomical images and is often used for disease detection, diagnosis, and treatment monitoring. MR is the most promising imaging modality to detect the changes of bony joints in structure and intensity, as it provides direct and noninvasive images

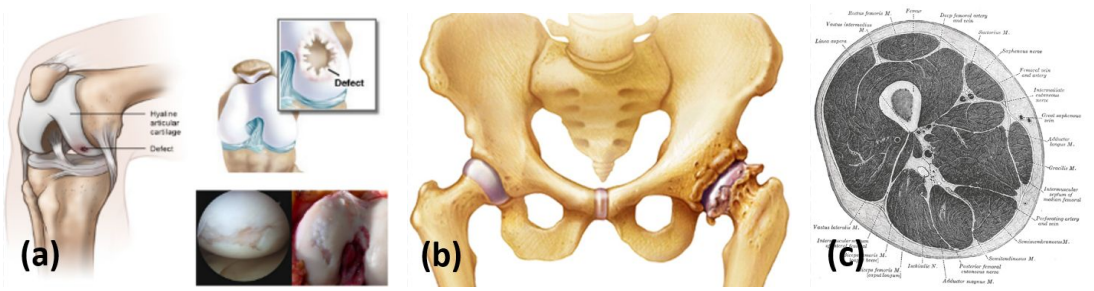


Figure 1.1: Schematic view of tissue anatomy (from [54, 79, 126]).

of the whole joint, including the soft tissue (e.g., cartilage) [61, 97]. On the other hand, computed tomography (CT) scans make use of computer-guided X-ray imaging to measure multiple planes of the body, and allow volumetric representations to be formatted for the demonstration of various bodily structures [122]. CT has the risks of ionizing radiation exposure, yet has still been applied as a modality showing the structural mechanisms of joints, and has acquired incremental usage in imaging bone and identifying osteoarthritis in recent years [14, 117].

1.1 Background

Developed from computer vision field, image segmentation is the process of automatic or semi-automatic partitioning a digital image into multiple segments (sets of pixels, also known as objects) within 2D or 3D images [6, 40, 90, 108]. In medical image segmentation, these segments often represent different tissue, organs, biologically relevant structures, or pathological conditions, and can simplify and/or change the representation of a medical image into something that is more meaningful and easier to analyze [50, 95]. Medical image segmentation is typically used to locate objects' masks or boundaries (e.g., lines and curves) in images. Since different imaging systems have highly diverse characteristics on body organs, numerous algorithms of organ segmentation have been

developed [73, 110, 146, 148] to assist clinicians or researchers to interpret and assess medical images [27, 30, 101].

Modern machine learning models mostly rely on massive human annotation for supervised training, and on the other hand, unsupervised learning builds a self-organized learning to find the interior patterns in dataset without annotated labels. Machine learning has been successfully applied in many applications of medical imaging [110, 111, 136]. With the fast increase of GPU computational ability and substantial growth of medical data/annotation, deep learning [64] has become the mainstream or even default machine-learning technique for medical image analysis [37, 134, 139, 137]. In all the deep learning based derivatives, convolutional neural networks (CNNs) have received special attention. By exploiting a series of convolutional layers, normalization layers, activation function, pooling layers, and fully connected layers, many state-of-the-art network models have been established for medical image classification, detection and segmentation [47, 88, 113, 134, 141].

In clinical trials, OA may affect any joint in the body, but it is particularly common in the knee. Knee OA is a chronic joint disease associated with the degeneration of knee cartilages. Cartilages are thin and elastic tissue, and they cover the ends of knee bones and supple the movement of knee joint. As the progresses of knee OA, the knee cartilages get thinner and may even completely wear away in severe cases, and eventually lose their functionality and cause some symptons (e.g., pain and stiffness in knee joint) [109, 127]. The schematic diagram of knee cartilage anatomy is shown in Fig. 1.1 (a). The assessment of morphological parameters (e.g., volume, surface area and thickness) of knee cartilages with MR imaging is important to diagnose the severity of OA and further analyze the treatments [20, 34, 44].

As the body's secondary weight-bearing joint, the hip joint is also commonly affected by OA, and the corresponding schematic figure is demostrated in Fig. 1.1 (b). The hip joint consists of the femoral head-neck bone and the acetabulum. Femoroacetabular impingement (FAI) or known as hip impingement is a symptom with bone overgrowth (called bone spurs) around the hip joint. FAI causes the joint bones an irregular shape and results in a structural disorder of hip joint [93]. In recent years, FAI has been growingly recognized as a potential forerunner of hip pain and main contributor to hip OA later in people's life [12, 82]. Insufficient femoral head-neck offset is common in FAI and reflected by the alpha (or beta) angle, a validated measurement for quantifying this anatomic deformity in patients with FAI [10, 87].

Additionally, the relevance of skeletal muscle and adipose tissue parameters (e.g., absolute or relative volume) in the thigh (shown in Fig. 1.1 (c)) are important modifiable factors that have been identified as determinants of both radiographic (structural) and symptomatic knee OA progression [9, 81, 106, 143]. Thus, as an auxiliary processing, quantitative assessment of thigh tissue composition is also helpful in studies of relevance between thigh tissue composition and knee OA [58, 111, 112].

1.2 Organization

In the following chapters, we focus on three segmentation modules for knee cartilage, femoral head-neck junction and thigh tissue extraction, which play important roles in knee and hip OA analysis, respectively. The methods are all designed to be robust and scalable for 2D/3D MR imaging modality.

In Chapter 2, we introduce the collaborative multi-agent learning for MR knee articular cartilage segmentation. The 3D morphology and quantitative assessment of knee

cartilages in MR imaging is of great importance for knee radiographic OA diagnostic decision making. In order to capture the wide range and thin structure of cartilages in detail, the MR data is usually acquired in large size (millions of voxels) and high resolution. The proposed framework is an effective and efficient delineation of all the knee articular cartilages in such imaging conditions. The key contribution is the adversarial learning based collaborative multi-agent segmentation network. In the proposed network, we use three parallel segmentation agents to label cartilages in their respective region of interest (ROI), and then fuse the three cartilages by a novel ROI-fusion layer. The collaborative learning is driven by an adversarial sub-network. The ROI-fusion layer not only fuses the individual cartilages from multiple agents, but also backpropagates the training loss from the adversarial sub-network to each agent to enable joint learning of shape and spatial constraints.

The first work is a specialized segmentation algorithm focusing on the knee cartilage segmentation problem. In Chapter 3, we propose a femoral head-neck bone segmentation method using deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In this work, we introduce a deep end-to-end network with two task-specific branches to ensure continuousness, smoothness and shape-preservation in segmented structure without additionally sophisticated shape adjustment, e.g., dense conditional random fields. First, we formulate the organ segmentation as a multi-task learning process that combines both region and boundary identification tasks, which can alleviate spatially isolated segmentation errors. Second, we use boundary distance regression to ensure the smoothness of the segmented contours, instead of formulating boundary identification as a classification problem [1]. Third, our deep network is designed to have a "Y" shape, i.e., the first half of the

network is shared by both region and boundary identification tasks, while the second half is branched for each task independently. This architecture enables the task-specific feature learning for better region and boundary identification, and offers information for segmentation refinement based on a fusion scheme using energy functional.

In Chapter 4, another application of thigh tissue quantification is discussed. The work focus on large scale MR thigh image analysis via accurately quantifying major tissue composition in the thigh by an integrated framework. Specifically, the framework is able to distinguish muscular tissue and different types of adipose tissues, i.e. subcutaneous adipose tissue(SAT), inter- and intra-muscular adipose tissue (IMAT and IAMAT), efficiently. Deformable models and learning based techniques are integrated in the framework to enable robust quantification. At last, we conclude the proposed modules in Chapter 5, and discuss future work to produce robust results towards large scale medical image data analytics.

Chapter 2

Collaborative Multi-agent Learning for MR Knee Articular Cartilage Segmentation

2.1 Introduction

The knee joint consists of knee bones, cartilages and menisci, and these tissue has complex structure and may undergo low-contrast imaging quality. The schematic diagram of knee anatomy is shown in Fig. 2.1 (a). Specially, the radiographic demonstrations of knee articular cartilages (i.e., femoral, lateral tibial, medial tibial and patellar cartilage), shown in Fig. 1 (b) and (c), have a large coverage scope and are the essential tissue components in the knee joint. Eckstein et al. [35] indicated that the cartilage morphology outcomes (e.g., cartilage thickness, cartilage surface area) by measuring 3D Magnetic Resonance (MR) knee data can help to identify the symptomatic and structural severity of knee OA. Bricca et al. [17] investigated the impact of knee joint loading exercise on people at risk of articular cartilage defect, which is an important hallmark of knee OA. On the other hand, the 3D cartilaginous labels are potential criterias for extensive quantitative measures in the knee joint. For instance, knee cartilage marks could help to compute the width of joint space narrowing and the derived distance map, which are considered as reference to assess the change of structure in knee OA [19]. Hunter et al. [53] investigated the MR Osteoarthritis Knee Score (MOAKS) by evolving a semi-quantitative scoring method based on the geometric relationships

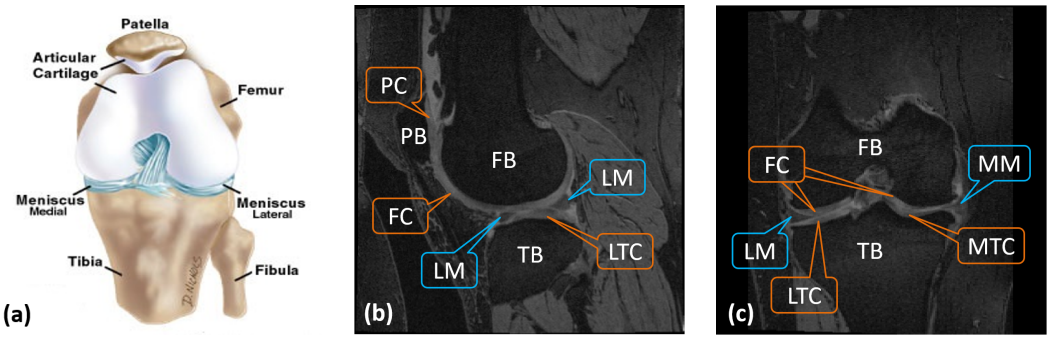


Figure 2.1: (a) Schematic view of knee anatomy (from [54]). (b) and (c) Demonstrative images from the sagittal and coronal directions of a 3D magnetic resonance (MR) knee data, respectively. They show a knee joint, including femoral bone (FB), tibial bone (TB), patellar bone (PB), femoral cartilage (FC), lateral tibial cartilage (LTC), medial tibial cartilage (MTC), patellar cartilage (PC), lateral meniscus (LM) and medial meniscus (MM).

between cartilaginous masks.

In order to capture the wide-coverage and exhaustive structure of cartilages for accurate assessment of knee OA, MR data is usually scanned with large size (millions of voxels) and high resolution. Fig. 2.2 exhibits a 3D MR knee data from the experimental dataset, which is from the public Osteoarthritis Initiative (OAI) database. All the volumetric MR scans in the set have $0.375mm \times 0.375mm$ in-plane resolution and $0.7mm$ between-slice resolution, the in-plane size is 384×384, and the slice numbers are all 160. These 3D MR data with high pixel resolution can reveal detailed-grand organ shape, structure and intensity information. Their large physical range in space also guarantees the data can capture all the crucial cartilaginous tissue in the knee joint region for the 3D based processing and clinical metrics analysis. Furthermore, in a diagnosing routine of knee OA, radiologists need to review the 3D medical images slice by slice to detect clues of joint degeneration, and measure the corresponding quantitative parameters manually. However, it is difficult to visually determine the symptoms of knee OA, because the radiographic representations may vary a lot in different individuals.

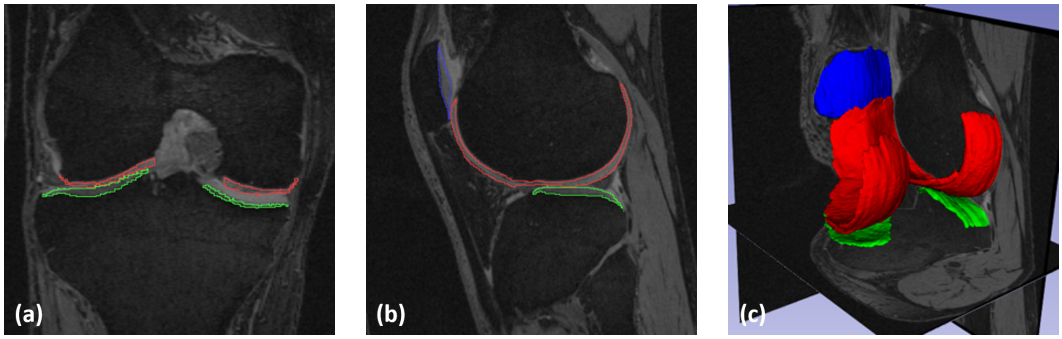


Figure 2.2: (a) and (b) show the coronal and sagittal slices of a 3D MR knee data. The red, green and blue contours indicate the femoral cartilage (FC), tibial cartilage (TC) and patellar cartilage (PC), respectively. (c) demonstrates the cartilage labels in 3D.

To facilitate this procedure, in OA research, some automated techniques of knee cartilage segmentation have been proposed. Atlas-based automated segmentation has been used to effectively segment objects of interest in biomedical images, which are estimated by using image registration. Shan et al. [107] considered the spatial relation of femur and tibia in knee area, and utilized the multi-atlas-based bone and cartilage registrations to initialize the spatial tissue priors. Then these priors are used for a joint decision of cartilage classification. Despite promising results by the atlas-driven cartilage segmentation, the accuracy of the above approaches highly relies on the results of registration (which may perform poorly when the training atlases can not represent the target subjects well) and the parameter setting of cartilage shape refinement. In addition, for large-size and large-scale dataset, the multi-atlas bone/cartilage registration requires high computation cost and relatively long segmentation time. Classification-oriented approach is another type of cartilage mask initialization.

In recent years, deep convolutional neural networks (DCNNs) has the ability to learn multi-level contextual information from raw input data, and thus DCNNs have shown superior potency in classification and segmentation problems. Prasoon et al. [98] proposed a tibia cartilage classifier by learning a joint objective function from multi-plane

2D DCNNs. This classifier can be trained with a limited memory budget, and slightly outperforms previous 3D non-learning/learning voxel categorization approaches. But its 2.5D feature learning strategy may not be sufficient for comprehensive information representation in 3D space for organ/tissue segmentation. As a progressive development, fully convolutional network (FCN) [72] has been highlighted as a fundamental for anatomy segmentation in medical images. Liu et al. [70] developed a fully automated framework for knee cartilage extraction by training a 2D FCN-based tissue probability predictor to drive a 3D deformable simplex-mesh-based cartilage reconstruction. Although this method demonstrates good performance, its shape representations by deformable models are limited to the use of pairwise forces which have high sensitivity in parameter setting and uncertainty in global/local image information assigning. Additionally, the mesh deformations and regularization iterations are still high computation consuming. It may cause boundary delineation errors at regions with high variability of cartilage shapes and weak appearance imaging quality, which may be common in MR knee joint images. The over-the-counter deep learning method, VNet [83], has shown superior performances in many 3D segmentation tasks, however, simply applying VNet to the MR knee data may have low accuracy and result in crash of training due to huge GPU memory consumption. Besides, the task of multi-cartilage classification suffers from severe class imbalance problem. Xu et al. [132] showed a contextual additive network focusing on the boost of memory efficiency for cartilage segmentation. The approach is based on small overlapping patches (a patch may only capture partial target) which may sacrifice certain accuracy. Some previous methods [46, 114] present multi-task networks. They introduce the distinctive boundary features of organ to improve accuracy. But the tissue of cartilage has very thin structure and its topology may

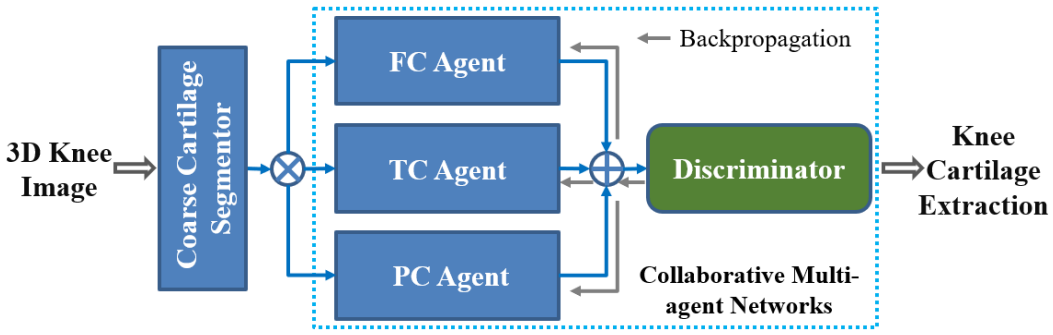


Figure 2.3: Flowchart of the collaborative multi-agent learning for segmentation.

change in degenerative conditions. Xu et al. [130] segmented thin objects in 2D images through a myocardial infarction segmentation. Yet this 2D task-specific strategy may still suffer from the memory issue when applying for the 3D knee data.

In this chapter, we propose a segmentation framework with collaborative multi-agent learning (shown in Fig. 2.3) for the task of knee cartilage labeling in large-sized and high-resolution 3D MR data. Through region of interest (ROI) extraction, 3 high-resolution cartilage ROIs are fed into different segmentation agents. The multiple agents collaborate by the help of discriminator and produce multi-class cartilage labels at the end. The ROI-fusion layer not only fuses the individual cartilages from multiple agents for discriminator, but also backpropagates the training errors from the adversarial sub-network to each agent to enable joint learning of shape and spatial constraints. Such collaborative multi-agent framework can obtain fine-grained segmentation in each ROI, and ensure the spatial constraints between different cartilages. It satisfies the limits of GPU resources and enables smooth training on the challenging data. The experimental results show that the proposed method can extract all cartilages accurately. The framework is also applied to test a large dataset with one thousand of subjects, to investigate the knee OA related cartilaginous parameters (i.e., volume, surface area).

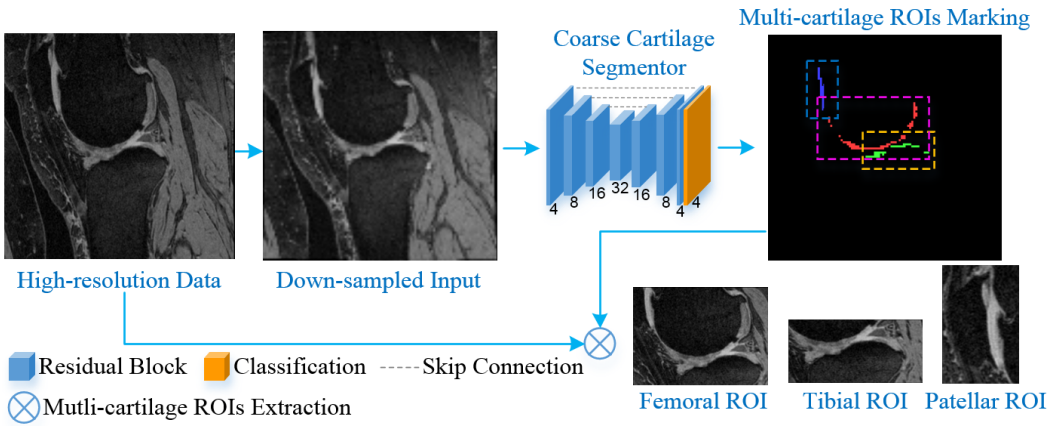


Figure 2.4: Overview of the multiple cartilage ROIs extraction (only show the sagittal view). The number of feature maps in the network is displayed under each block.

2.2 Methods

The overview of the proposed framework is shown in Fig. 2.3. The coarse cartilage segmentor and ROI extraction (i.e., $\otimes$) steps aim to efficiently localize and extract three local regions of FC, TC and PC, and feed the ROIs to segmentation agents respectively. The blue dashed box shows the collaborative multi-agent cartilage segmentation module, which consists of three segmentation agents, one ROI-fusion layer (i.e., $\oplus$), and one joint-label discriminator.

2.2.1 ROI extraction

In order to initialize the collaborative multi-agent learning, we first extract the ROIs of three cartilages. As shown in Fig. 2.4, by utilizing the location information of the multi-cartilage marks from the coarse segmentor, the image and label ROIs of FC, TC and PC are extracted from the original data. The segmentor's structure is like VNet [83], i.e., encoding-decoding. The encoding part contains 3 down-samplings (by convolutions of filter size 2 and stride 2) to obtain 3 different scales of feature maps. The decoding part has 3 up-samplings (by deconvolutions of filter size 2 and stride 2) to restore the

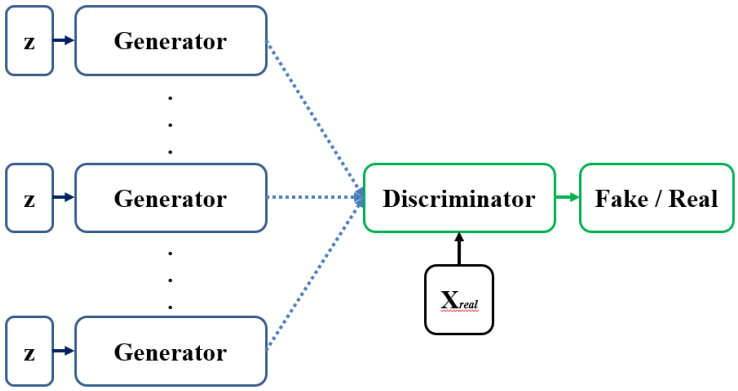


Figure 2.5: Illustration of the multi-agent diverse GANs in [41, 51].

scale of feature maps to reach the original input size. The blue block in this figure represents residual block followed by a down-sampling or up-sampling layer mentioned above when changing resolution. All the convolutional layers in the residual blocks have filter size 3, stride 1 and zero-padding 1. PReLU activation and batch normalization follow the convolutional and deconvolutional layers. The coarse cartilage segmentor is trained based on multi-class cross entropy loss ℓ_{mce} to obtain cartilage masks from the down-sampled MR data (e.g., $192 \times 192 \times 160$).

2.2.2 Collaborative multi-agent learning

Generative adversarial Networks (GANs) [43] have emerged as a powerful data synthesis/augmentation approach in various applications. The advent of GANs suggests that images can be synthesized by training a generative network with a discriminative network. The generative network uses the input z with a fixed distribution (e.g. Gaussian), and learn to transform it to a real sample distribution X_{real}. The GANs-based approaches are capable of generating "real" data because of its competitive mechanism. For the stability in GANs' training on diverse-class dataset, Ghosh et al. [41] proposed

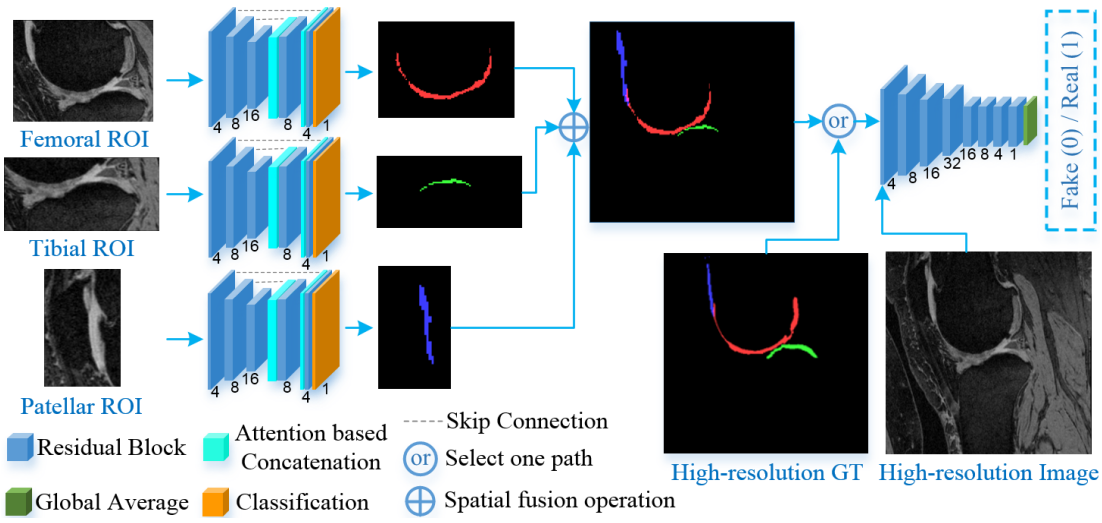


Figure 2.6: Demonstration of the collaborative multi-agent learning framework for fine-grained cartilage segmentation. The agents yield binary labels and the spatial fusion operation outputs a 4-channel result (FC, TC, PC and background).

the multi-agent diverse GANs which disposes multiple generators for one discriminator to adapt the diversity of samples in Fig. 2.5. This framework enables each agent (generator) better captures the intra-class variances, and meanwhile makes the discriminator jointly finds the inter-class differences. On ther other hand, GANs have attracted marked attention in the segmentation field [28, 74, 138]. In the segmentation with GANs, the segmentor incorporates a discrimination (critic) network to discriminate the segments from the ground truth labels. Through this adversarial process, the discrimination part learns the higher order regularities from shaped structure and spatial position, and effectively propagates this global information back to the segmentation part to improve the reality of segmented outcomes.

Inspired by the aforementioned theories, in this learning stage (shown in Fig. 2.6), we construct one big network by three individual segmentation agents, one ROI-fusion layer, and one adversarial sub-network. The segmentation agent $A_{c=\{f,t,p\}}$ (f, t and p stand for FC, TC and PC, respectively) aims to generate fine cartilage binary mask

$A_c(\mathbf{x}_{i,c})$ in the respective ROI $\mathbf{x}_{i,c}$ (its ground truth (GT) ROI is $\mathbf{y}_{i,c}$ and i is the data index). Each ROI is small enough to cover only one cartilage in it. Since the large portion of background and other cartilages are excluded, the class imbalance problem is relieved significantly. The small ROIs also reduce the requirement for the computational resources (i.e., GPU memories) and enable fine-grained segmentation in high-resolution data. All the segmentation agents have similar VNet-like pattern as the coarse segmentor. To balance the receptive field of neurons and the GPU memory consumption, we further reduce the down- and up-sampling operations to 2. Considering the thin characteristics and unclear boundary of cartilage, we need to better utilize the multi-resolution contextual features to capture its fine details. In VNet, skip connection is designed to merge the up-sampled high-level features I_h^{up} in decoding path and the equivalent-resolution low-level features I_l in symmetrical encoding path by simple concatenation. Here, we apply an attention mechanism [55] to extend the skip connections. Formally, the connecting operation becomes $o\left(\alpha \odot I_l, I_h^{up}\right)$, where o denotes concatenation along the channel dimension, and $\odot$ is element-wise multiplication. The attention mask $\alpha = m\left(\sigma_r\left(c_l\left(I_l\right) + c_h\left(I_h^{up}\right)\right)\right)$ serves as a weight map that guides the learning to focus on desired region. Here, c_h and c_l are two convolutions of filter size 1 and stride 1; σ_r is an activation function (e.g., ReLU); m is another convolution of filter size 1 and stride 1 with sigmoid to contract the features to a single-channel mask. The light blue blocks in Fig. 2.6 represent the attention based concatenation, and its schematic diagram is illustrated in Fig. 2.7.

Although individual agent can obtain fine segmentation in its ROI, the individual learning losses the mutual constraints between cartilages. In order to make the agents collaborate together to make use of the mutual position and shape priors of all the

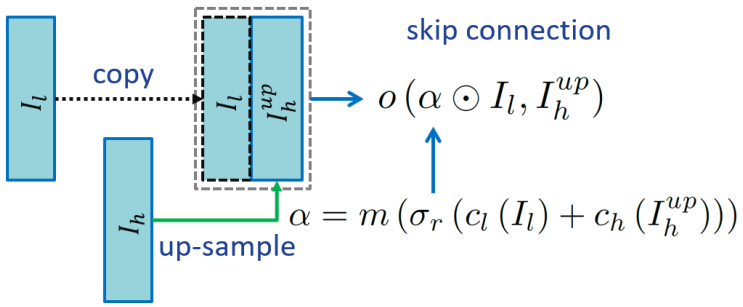


Figure 2.7: Illustration of the attention mechanism in skip connection.

cartilages for better delineations, we propose a collaborative learning strategy. This strategy utilizes a ROI-fusion layer $\mathcal{F}$ to restore the single-cartilage output from each agent back to the original knee joint space where the mutual constraints and priors can be encoded. $\mathcal{F}(A_f, A_t, A_p)$ is implemented by using the location information of the three input ROIs to fuse the fine cartilage masks back to the original space. Then, the multi-cartilage priors are learned implicitly by adversarial learning strategy. We utilize a discriminator sub-network D to classify the fused multi-cartilage mask as "fake" and the whole GT label $\mathbf{y}_i$ as "real". In adversarial learning, the agents and the discriminator are trained alternatively. The parameters of agents are fixed when training the discriminator, and vice verse. In this way, discriminator sub-network can learn joint priors of multiple cartilages and guide the agents to produce better segmentation. It is important to note that the layer $\mathcal{F}$ not only fuses ROIs by their coordinates, but also passes the gradient updates from the discriminator to the agents during backpropagation, so that the two parts can be optimized in this alternating fashion. Since it is not intuitive to judge the labels without seeing the input in segmentation task, we borrow the idea of conditional generative adversarial nets, and treat the input MR knee image $\mathbf{x}_i$ as the conditioning variable. Fig. 2.6 shows that the discriminator sub-network consists of 4 down-sampling convolutional layers, and the same residual block in the

agents is also employed under each resolution level for contextual information learning. The input to the discriminator is a pair of MR knee image $\mathbf{x}_i$ and multi-label cartilage mask (either the GT label $\mathbf{y}_i$ or $\mathcal{F}(A_f, A_t, A_p)$). A global average layer is utilized at the end to generate a probability value for fake/real mask discrimination.

The loss functions of discriminator and agents are defined in Eq. 2.1 and Eq. 2.2. Here, ℓ_b indicates the binary cross entropy loss. In Eq. 2.2, the first term $L_s = \ell_b \left[A_c \left(\mathbf{x}_{i,c}\right), \mathbf{y}_{i,c}\right]$ is to train each single segmentation agent. The second term $L_m = \ell_{mce} \left[\mathcal{F}(A_f, A_t, A_p), \mathbf{y}_i\right]$ and the third one are applied on the fused multi-cartilage mask for joint-label learning. The discriminator D and segmentation agents $A_{c=\{f,t,p\}}$ are alternatively trained by minimizing Eq. 2.1 and Eq. 2.2.

$$\sum\nolimits_i \left\{\ell_b \left[D\left(\mathbf{x}_i, \mathbf{y}_i\right), 1\right] + \ell_b \left[D\left(\mathbf{x}_i, \mathcal{F}\left(A_f, A_t, A_p\right)\right), 0\right]\right\} \tag{2.1}$$

$$\sum\nolimits_i \left\{\sum\nolimits_{c=\{f,t,p\}} L_s\left(\mathbf{x}_{i,c}, \mathbf{y}_{i,c}\right) + L_m + \ell_b \left[D\left(\mathbf{x}_i, \mathcal{F}\left(A_f, A_t, A_p\right)\right), 1\right]\right\} \tag{2.2}$$

2.3 Experiments

2.3.1 Experimental settings

We validate our proposed method on the iMorphics dataset from the OAI database. This set includes 176 3D MR (sagittal DESS sequences) knee images. The set is splitted into training: 120, validation: 26, testing: 30. Patients are randomly and exclusively used in the three subsets. Fixed ROI size of each type of cartilage is pre-defined based on adequate evaluation on the training data. We compare the proposed method with the state-of-the-art dense atrous spatial pyramid pooling (DenseASPP) for semantic segmentation [140]. It integrates the ASPP architecture in a dense connection manner,

Table 2.1: Quantitative comparisons of approaches: mean and std of evaluation metrics.

	Femoral Cartilage			**Tibial Cartilage**			**Patellar Cartilage**			**All Cartilages**		
	DSC	*VOE*	*ASD*	*DSC*	*VOE*	*ASD*	*DSC*	*VOE*	*ASD*	*DSC*	*VOE*	*ASD*
$D1$	0.862	24.15	0.103	0.869	22.93	0.104	0.844	26.65	0.107	0.866	23.59	0.095
	0.024	3.621	0.042	0.034	5.184	0.061	0.052	7.429	0.049	0.023	3.475	0.026
$D2$	0.832	28.64	0.131	0.879	21.38	0.088	0.861	23.69	0.091	0.851	25.94	0.111
	0.025	3.618	0.059	0.038	5.972	0.055	0.040	6.027	0.051	0.023	3.393	0.036
$C0$	0.814	31.30	0.205	0.806	32.42	0.199	0.771	35.74	0.350	0.809	31.99	0.213
	0.029	4.155	0.095	0.033	4.577	0.055	0.132	14.56	0.129	0.031	4.350	0.095
P1	0.868	23.19	0.108	0.854	25.17	0.126	0.824	28.78	0.201	0.862	24.24	0.110
	0.023	3.514	0.067	0.029	4.173	0.059	0.104	12.45	0.439	0.023	3.457	0.048
P2	**0.900**	**18.82**	**0.074**	**0.889**	**19.81**	**0.082**	**0.880**	**21.19**	**0.075**	**0.893**	**19.19**	**0.073**
	0.037	6.006	0.041	0.038	6.072	0.051	0.043	6.594	0.038	0.034	5.434	0.034

which is able to generate large receptive field and multi-scale features for segmentation tasks. We also evaluate performances of the proposed coarse segmentor and individual agents to show the effectiveness of the collaborative learning. Dice similarity coefficient (DSC), volumetric overlap error (VOE) and average surface distance (ASD) between the GT labels and segmented results are reported. In the training (no pre-trained weights used), we set the batch size to 1 and multiply a factor of 0.95 every 10 epochs to reduce the learning rate (LR). The Adam (with initial LR 0.001) and stochastic gradient descent (SGD, with initial LR 0.0002) solvers are used for each agent and the discriminator. All the networks are trained and tested by a 12GB-RAM Titan X GPU.

2.3.2 Experimental results

Quantitative comparisons are shown in Table 2.1. $C0$ represents the coarse cartilage extraction by the segmentor in Fig. 2.4. **P1** denotes the fused results generated by the proposed segmentation agents, without the joint learning by the adversarial sub-network. **P2** represents results from the proposed method by employing the collaborative multi-agent learning framework as in Fig. 2.6. For comparison, we integrate two variants of DenseASPP into the collaborative multi-agent framework. In the first

variant $D1$, the residual blocks and skip connections are replaced by DenseASPP blocks in the two down-sampled levels of the agent network. While in the second variant $D2$, only the deepest level is replaced with DenseASPP block.

From the table, we can see that the proposed segmentation **P2** achieves the best performance in all metrics. The mean results of $C0$ (i.e., a similar implementation of VNet) are relatively good and have no gross failure in our experiments. This shows that the coarse stage is reliable initialization. The proposed **P2** obviously outperforming **P1** shows that segmentation agents are improved with the help of the proposed collaborative learning strategy. The overall performances of the DenseASPP based variants $D1$ and $D2$ are close to that of **P1**. It indicates that the proposed agent network with the attention based concatenation is effective enough, compared to the DenseASPP blocks which have more complicated architecture. In addition, the results of the proposed method are comparable to those reported in some recent studies [4, 132]. Xu et al. [132] reported a total DSC (0.887 ± 0.024) value of FC and TC. Ambellan et al. [4] utilized both 2D and 3D deep learning based segmentations with statistical shape models as shape refinement postprocessing for femoral and tibial cartilages extraction. Using a similar set from OAI, they achieved[2] DSC (0.893 ± 0.024), VOE (19.4 ± 3.87) and ASD (0.19 ± 0.09) for FC, DSC (0.881 ± 0.038), VOE (21.05 ± 5.808) and ASD (0.223 ± 0.143) for TC. Without the sophisticated shape adjustment step, the proposed method acquires the comparable DSC and VOE scores, and much lower surface distance errors. Hence, the proposed framework can be used to automatically generate reliable assessments of all important articular cartilages in quantitative analysis for knee OA.

[2] [4] separately presents the results of FC, medial TC and lateral TC at two timepoints. For convenience, we average these results and get the approximate mean/std metrics.

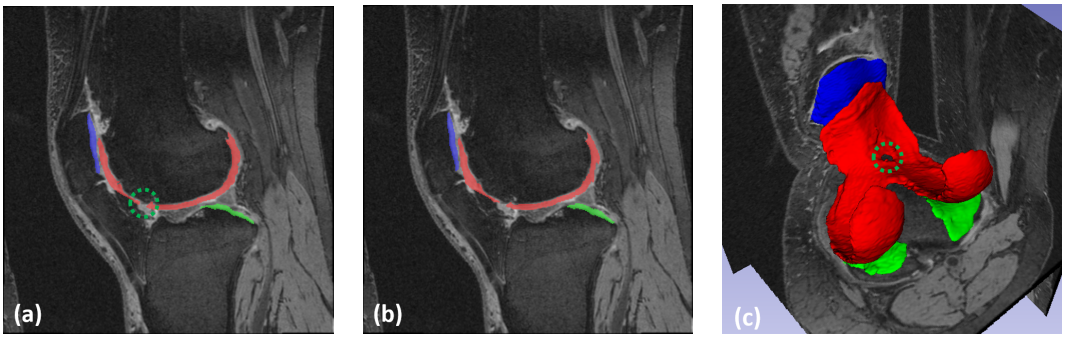


Figure 2.8: Results of subject 1. (a) and (b) show the segmentation and GT labels for FC (red), TC (green), and PC (blue) in sagittal view. (c) is the segmented 3D cartilages.

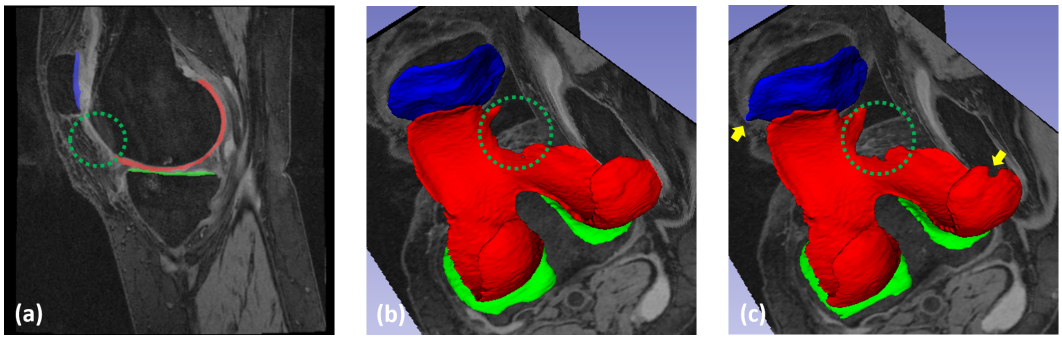


Figure 2.9: Results of subject 2. (a) shows the segmented cartilages in sagittal view. (b) and (c) demonstrate the GT and segmentation results in 3D view.

Visualization results (two examples) of the proposed method are showed in Fig. 2.8 and 2.9. The two patients have obvious shape variance of cartilages. In Fig. 2.8 (a)-(c), the proposed method can accurately extract most of the cartilage regions and obtain smooth tissue boundaries. Furthermore, as indicated by green dashed circles in Fig. 2.8 (a) and (c), our method can effectively capture a small cartilage defect. The green dashed circles in Fig. 2.9 (a) and (c) indicate a possible cartilage damage/miss symptom well captured by our method. The 3D view exhibiting accurate 3D pattern of cartilage defects could be very useful in visual study of cartilage-related diseases. The yellow arrows in Fig. 2.9 (c) show some minor errors occurred in some neighborhood areas due to unclear boundaries.

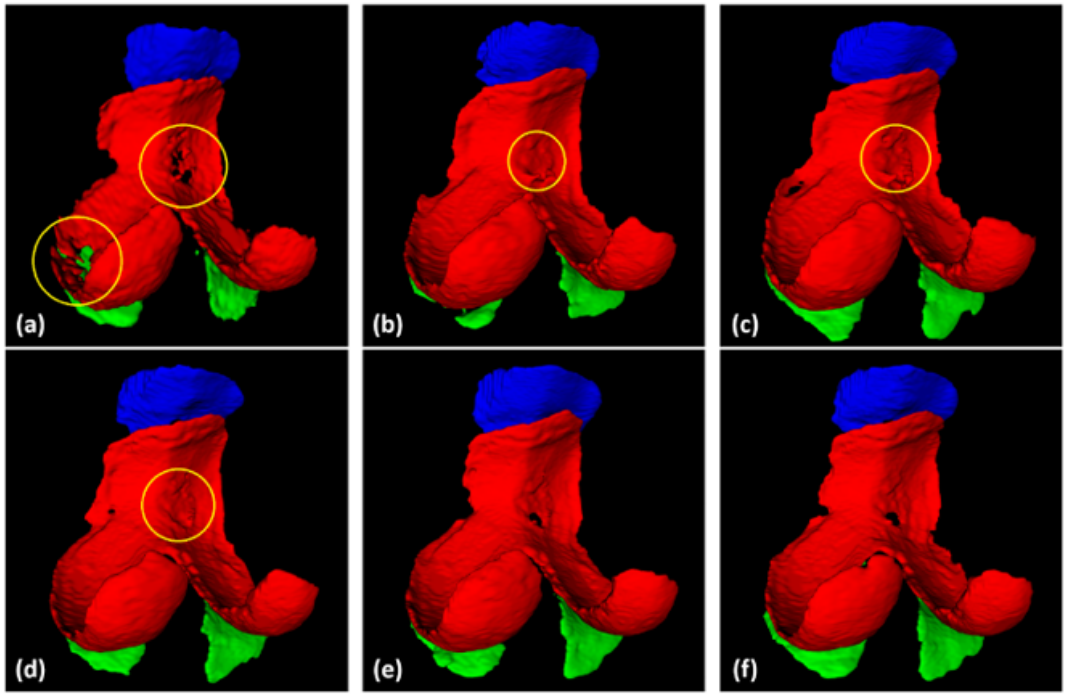


Figure 2.10: Visual comparisons between different methods for subject 1. (a) Segmented results of $C0$. (b) $D1$. (c) $D2$. (d) **P1**. (e) **P2**. (f) GT labels.

To further validate the performance of the proposed method, we compare the segments between different methods and show the visual results of subject 1 and 2. The result of $C0$ shown in Fig. 2.10 (a) affords a reliable initialization of cartilage labeling, yet it can not capture the detailed grained cartilage labels. The down-sampled MR data used in $C0$ may degenerate the structure of cartilage at a point with very thin tissue layer, and causes the incomplete segmentation (indicated by the yellow circles in Fig. 2.10 (a)) From Fig. 2.10 (b) to (d), the segmented labels have the similar shape patterns and the accuracy is obviously increased compared to $C0$, but all of them over-segment the defect area (indicated by the yellow circles in Fig. 2.10 (b) to (d)). **P2** gives the closest cartilages reconstruction comparing with the GT labels. In Fig. 2.11 (a), $C0$ still suffers from the similar incomplete segmentation problem, although it could segment the main structure of cartilages. $D1$ and $D2$ in Fig. 2.11 (b) and (c) show good

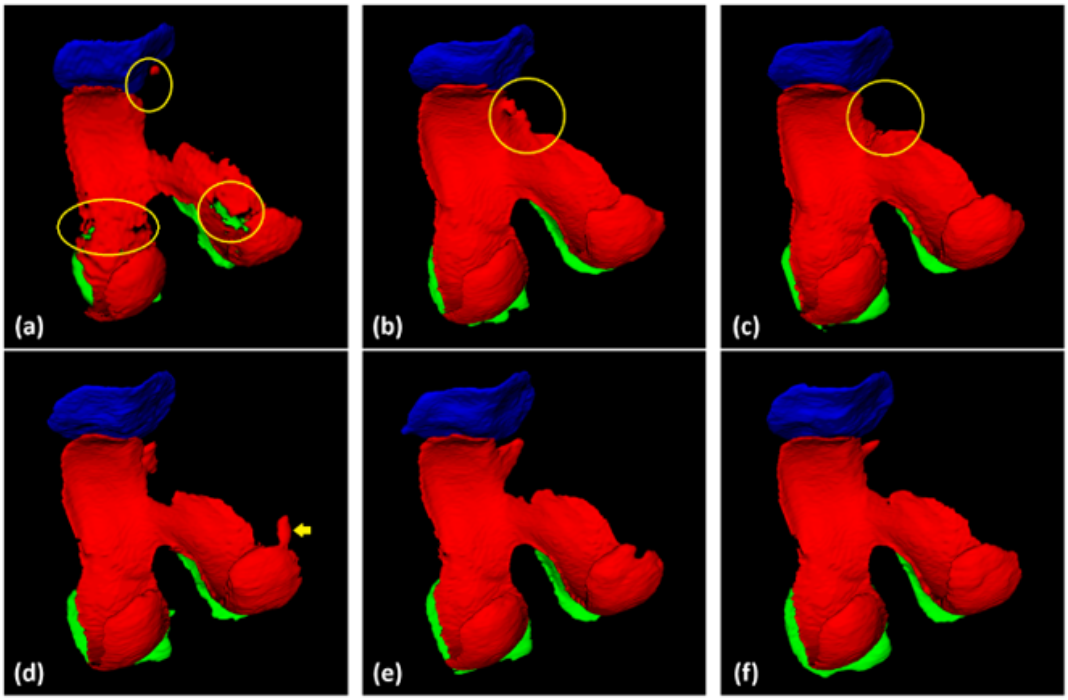


Figure 2.11: Visual comparisons between different methods for subject 2. (a) Segmented results of $C0$. (b) $D1$. (c) $D2$. (d) **P1**. (e) **P2**. (f) GT labels.

outcomes, but they still over-segment the damage/miss symptom located in the yellow circles of Fig. 2.11 (b) and (c). **P1** trends to restore the cartilage missing symptom, yet it has an extra leaking issue, because of the absence of the adversarial sub-network. On the other hand, with the help of the proposed collaborative learning, **P2** obtains the best result in Fig. 2.11 (e).

2.3.3 Extensive evaluations for knee OA analysis

The proposed scheme is effective for cartilage biomarkers (e.g., surface area and volume) estimation in large-scale quantitative tests. We have applied the scheme on a 36-month subset (1000 3D MR knee data) of OAI for OA-MOAKS analysis [53] and the segmented results are visually evaluated (gross error $< 10\%$). In the MOAKS metric, higher score of index means the patient has more severe OA degree in the knee, and score 0 represent

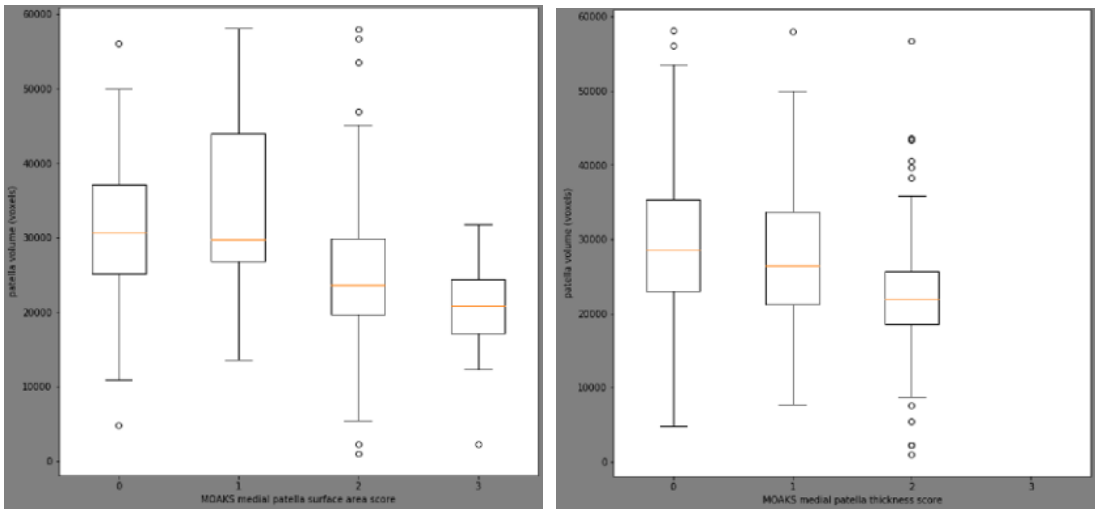


Figure 2.12: Correlation between MOAKS scores and mean volume/surface area of patellar cartilage. In (a) and (b), the x-axis represents the MOAKS scores. Y-axis denotes the value of surface area (in (a)) and volume (in (b)).

the patient is normal. From Fig. 2.12, the plots show a clear correlation between the MOAKS scores and the mean volume/surface area of patellar cartilage.

2.4 Summary

In this chapter, we present a fully automatic method to segment three knee cartilages in 3D MR images based on a collaborative multi-agent learning architecture. Each segmentation agent depicts the high-resolution cartilage mask in its coarsely (but efficiently) located ROI. A novel skip connection by multi-resolution attention mechanism is introduced to enhance the feature extraction of target, while suppressing confusing information in neighborhood areas. Then, the depicted multiple ROIs are spatially fused into the original space to form a multi-cartilage label image for collaborative learning. The collaboration of agents is implemented by the novel ROI-fusion layer followed by an adversarial discriminator to ensure the shape and position constraints. Learning of the agents and discriminator are conducted in an alternating fashion. In our experiments,

the proposed method achieves robust and accurate segmentation for all important articular cartilages in high resolution and large 3D MR knee data. Moreover, in the extensive test on a large-scale dataset, the outputs show that the proposed method is not only effective for the experimental dataset, but also practical for extensive quantitative analysis on knee OA. In future we will apply the method for quantifing cartilage biomarkers (e.g., volume, thickness, surface area) in large-scale studies and detecting cartilage defects for lesion estimation [35, 53]. Besides the cartilages, the proposed framework could also be extended for other multi-organ segmentation tasks [118].

Chapter 3

Shape Preserved Femoral Head-neck Bone Segmentation

3.1 Introduction

In recent years, studies have recognized femoroacetabular impingement (FAI) as an important pathogenic mechanism to contribute the progression of hip OA [11, 38, 59, 100]. The impingement of this disease is conditioned on the influence of the abnormality of bony shape on the femoral head-neck junction and/or the acetabulum. In the adult population, FAI is one the major reasons to cause hip joint pain or tenderness, and may eventually result to hip OA later in people's life. There are three types of FAI, i.e., cam-type, pincer-type and combined-type [45, 100, 104], in Fig. 3.1). In the cam-type impingement, the proximal femoral neck or the femoral head-neck junction suffers from irregular osseous prominence. The excess and abnormal bone growth grinds the cartilage inside the acetabulum, and further causes non-smooth rotation of the femoral head inside the acetabulum. In the pincer-type impingement, excessive acetabular bone grows over its normal rim and over-covers the femoral head. Over time, this situation could result in tears of the labrum, under the prominent rim of the acetabulum. Furthermore, many FAI cases are mixed by the two typs of impingement, begetting varying degrees of abnormal bone morphology. All of these FAI cases will injure the contours at the femoral head-neck junction of patients, and thus it is highly imperative to develop examination and treatment for the disease [63].

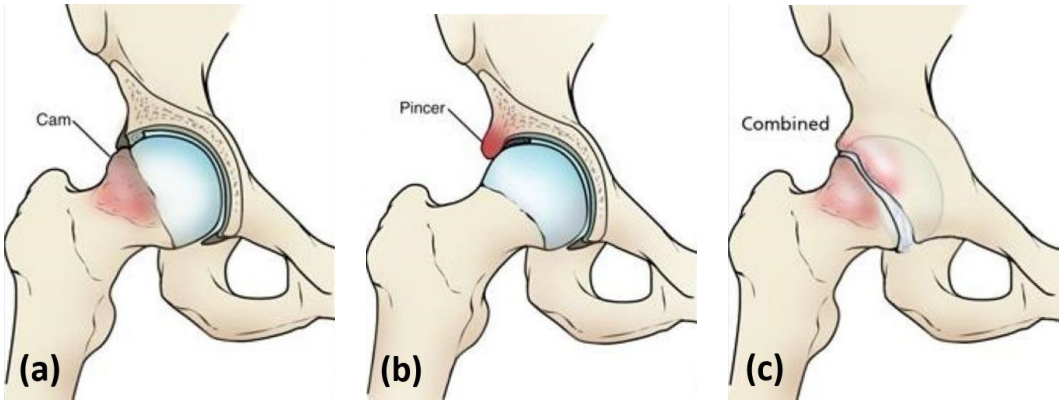


Figure 3.1: Schematic view of FAI [1]. (a), (b) and (c) show cam-type, pincer-type and combined-type FAI, respectively.

Although clinical history and physical examination findings are helpful for the diagnosis of FAI, radiologic studies are the most convincing way. Common radiologic imaging methods for hip joint imaging include radiography [49, 85, 115], computed tomography (CT) [15, 91, 116], and magnetic resonance (MR) [5, 65, 94]. Radiography (or x-rays) could show good outline of bone in images. CT provides more detailed bone imaging than x-ray to diagnose the abnormally shaped bones of FAI. MR image acquisition has been considered as an important method for hip bones imaging in multiplane. Besides displaying the contour of bone, MR could offer better soft tissue imaging, such as cartilage and labrum, and thus enable comprehensive evaluation of the femoral head-neck junction shape [36, 129].

In order to quantify the degree of femoral deformity implicated by FAI, alpha (or beta) angle is defined as a measure of femoral head asphericity, and another parameter, anterior (or posterior) offset, is assessed between the femoral head and neck [32, 87, 145]. Based on the manual or automatic femoral head-neck bone segmentation (built from the input images sampled in the coronal plane), to measure the alpha angle [3, 103], a 2D circle fitting is firstly placed over the femoral head. Then, the intersection point p_{ca}

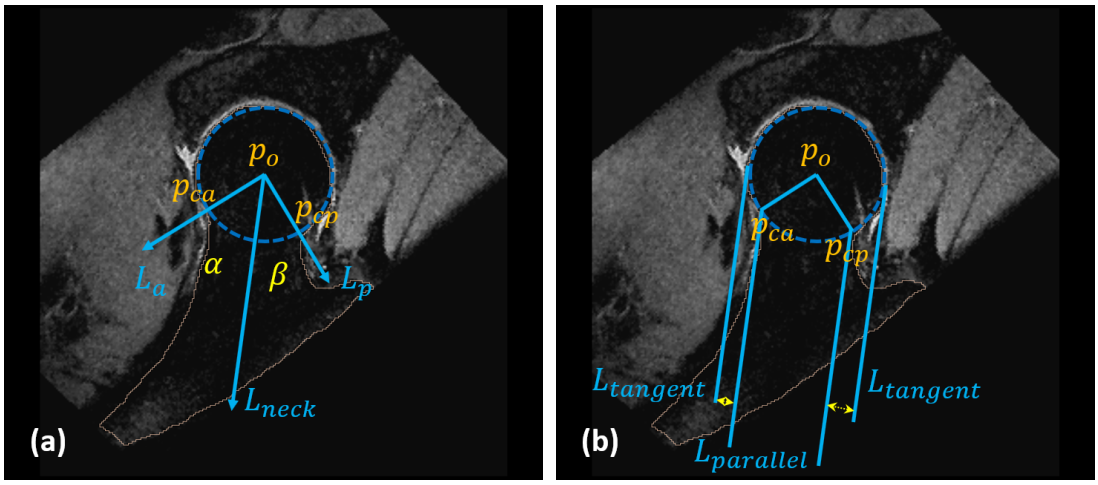


Figure 3.2: Measurements of the alpha (or beta) angle and anterior (or posterior) offset are shown in (a) and (b), respectively.

between the fitted circle and the anterior femoral neck is marked. Third, the central axis of the femoral neck, L_{neck}, is described by drawing a line to connect the femoral head center (approximated by the fitted circle center p_o) and the femoral neck center. After depicting the line L_a between the points p_{ca} and p_o, the angle between L_a and L_{neck} is defined as the alpha angle. The definition of the beta angle is similar to the alpha one, just replacing p_{ca} to p_{cp} (the intersection point between the fitted circle and the posterior femoral neck). To measure head-neck offset [3], a line $L_{tangent}$ is drawn which is tangent to the anterior (or posterior) femoral neck and meanwhile parallel to L_{neck}. At the concavity of the anterior (or posterior) femoral neck, another line $L_{parallel}$ parallel to $L_{tangent}$ is obtained. The perpendicular distance between $L_{parallel}$ and $L_{tangent}$ is defined as the head-neck offset. The measurements of the two types of parameters based on a segmented contour are demonstrated in Fig. 3.2. Cases with large alpha (or beta) angle, or narrow head-neck offset, would be considered abnormal [7].

Manual delineation and measurement of the angle and offset parameters work well in single image or small dataset [8, 84]. However, automatic assessments are very necessary

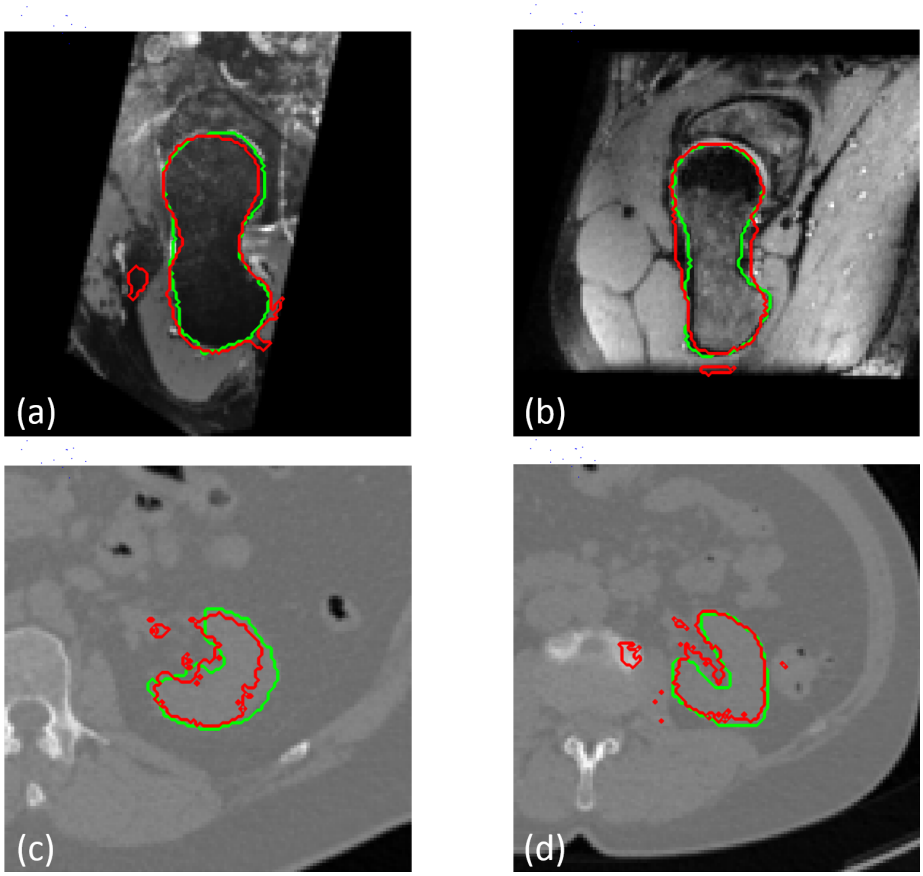


Figure 3.3: (a)-(d) show the ground truth (green lines) and segmented contours (red), respectively. The FCN method is tested on two MR femoral head-neck images in (a)-(b). In order to extensively validate the method, two CT kidney cases are also tested. The FCN-based results exist spatially isolated errors and smoothless segmented boundaries.

for numerous multi-slices. Therefore, the femoral head-neck bone segmentation plays the key role in the measuring process. An automated segmentation performs statistical shape modeling (SSM) to extract the proximal femoral surfaces, and further utilizes the results to determine the variation of bone shape [22]. Xia et al. [128] incorporated Atlas-based initialization and SSM to segment both proximal femoral bone and innominate bone. Although SSM is a basic module to capture the range of shape variability, it is a very challenging task to build the shape correspondence among all training labels. The modeling has high computational intensity and may mismatch structure in occasional.

In recent progress of machine learning, fully convolutional network (FCN) [72] has

been highlighted as a fundamental segmentation approach for anatomy delineation in medical images. It exploits the deep convolutional neural networks (DCNN) for coarse-to-fine inference and makes a prediction at every pixel. Without manually setting handcrafted features, DCNN has the ability to learn a hierarchical representation of raw input data. However, FCN is limited for lower-level tasks requiring precise localization, e.g., semantic segmentation, since the DCNN-based inferences inside FCN build invariance to spatial transformations and provide only abstraction of spatial details. In Fig. 3.3, FCN is more likely to have predicted outliers due to the high variability of organic shapes and low-contrast imaging quality in medical images. It may not produce a continuous segmented object with smooth boundary.

Some current researches employ a new strategy called multi-task learning for organ segmentation. The main task of this strategy is to optimize target extraction by leveraging auxiliary information from a set of correlated tasks (e.g., background classification or bounding box investigation). Zhang et al. [149] studied an effective facial landmark detection with multi-task deep representation that combined heterogeneous but subtly correlated tasks. The fast RCNN [42] jointly trains for classification and bounding-box regression, and obtain superior object detection. Chen et al. [24] presented a multi-task deep representation that combines region and boundary classifications to obtain continuous tissue description. In a diabetic macular edema grading approach, Li et al. [68] utilized a multi-task framework to promote the performance of each individual task.

Based on these previous studies, we formulate the organ segmentation as a multi-task network consisting of two parallel end-to-end branches, as shown in Fig. 3.4, to alleviate the spatially isolated segmentation errors in Fig. 3.3. Each task involves two symmetrical parts, i.e., encoding and decoding. The first task is a conventional fully

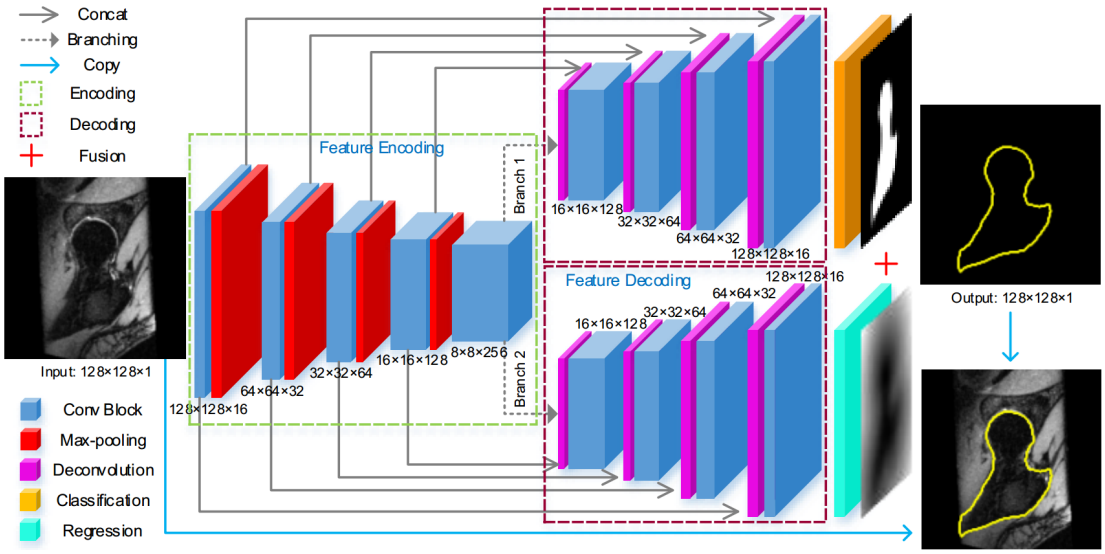


Figure 3.4: Proposed deep multi-task and task-specific feature learning network. The left half is encoding for the abstraction of multi-level contextual feature. The right half has two decoding branches for the task-specific learning. Each blue convolution block includes two convolutional layers with filter size of 3×3 and zero-padding of 1. Parametric rectified linear unit and batch normalization are also adopted in all convolutional and deconvolutional layers.

convolutional network for the inference of organ probability map. The second task is a novel deep regression network (DRN) that regresses the distance constraint information of organic boundary. During the distance regressing, it produces continuously numerical constraints of spatial information. Thus the DRN obtains better potential to preserve the smoothness of boundary, comparing with the discrete classified labels obtained by boundary classification. Gao et al. [39] proposed a boundary extractor by learning a joint objective function from displacement estimation and organ classification, and these two tasks only share parameters in the final loss. But in Fig. 3.4, the two task-specific branches in the proposed network share parameters in the encoding process in the first half of Y-shape network, while having their own decoding parameters to represent the features for the classification and regression, respectively. This structure ensures a balanced and sufficient parameters learning to represent the

task-specific features for region and boundary identification. During the training, the cumulative loss is optimized by jointly investigating the two losses from each single branch. Adding the distance regression task can effectively regularize the smoothness of segmented boundary and reduce the isolated segmentation errors. Simultaneously, the classification branch is efficient for locating and extracting target. Finally, the proposed method also explores a unified segmentation architecture which incorporates a shape refinement. We combine the inferred organ probability map and the regressed boundary distance map based on a fusion scheme using energy functional. This scheme can efficiently refine organ boundary, and avoid the complex parameter tuning of additionally sophisticated shape adjustment (e.g. dense conditional random fields[25]).

The proposed architecture has multi-fold benefits. (1) Its multi-task architecture combines both region and boundary identification, and mitigates spatially isolated segmentation errors. (2) The usage of boundary distance regression is able to ensure the smoothness of the segmented contours. (3) The Y-shape network structure enables sufficient learning of task-specific feature for better region and boundary identification.

3.2 Method

In this section, we elaborate the design of multi-task learning architecture, and then we specify the objective function and discusses its advantages in solving. Subsection 3.2.1 describes the background knowledge of multi-level contextual feature representation. Subsection 3.2.2 presents the design of multi-task learning architecture. Subsection 3.2.3 gives the objective function and discusses its advantages in solving. In the last subsection, a further refinement of the two parallel end-to-end branches is discussed.

3.2.1 Multi-level contextual feature representation

FCN-based approaches have led to significant quantitative improvements for the task of semantic image segmentation. Given a certain size of receptive field, the predication scores from FCN are generated to guide the final discrimination for object extraction. However, the network with single receptive field can not properly deal with the complex conditions of high variability of organ shape in medical images. For instance, at different axial position of a medical volume data, the organ's size in each sampled slice highly varies. A receptive field with fixed size (e.g. 32×32) may completely capture the targets at both ends of an organ (e.g., liver, kidney or bone), while larger receptive size is required for slices in the intermediate locations since the organ's area enlarges. Hence multi-size contextual information is helpful to receive the integral interior structure of tissue and the sufficient background knowledge surrounding it, and then improve the recognition performance.

3.2.2 Deep end-to-end network with multi-task learning

In this subsection, we present a deep end-to-end network branched by two task-specific learning for the organ segmentation. As shown in Fig. 3.4, the network takes the entire 2D image as input, and the first task is a conventional FCN for the inference of organ probability map, and the second one is the novel DRN regressing the distance constraint information of organic boundary.

The main structure of each task is designed as a symmetric way, i.e., encoding-decoding, similar to [105]. The two branches share the encoding part, which contains 4 max-pooling layers with stride 2 to obtain 4 different resolutions of raw image. Under

each resolution, two convolutional layers are utilized for feature abstraction. This successive encoding allows to obtain multi-size contextual information which is helpful to receive the integral interior structure of tissue and the sufficient background knowledge surrounding it, and then improve the recognition performance. In the decoding part of each task, we deploy 4 deconvolutional layers in a cascaded way for up-sampling feature maps. Each deconvolutional layer adopts stride 2, so it avoids the usage of large up-sampling factors (16, or 32), and effectively reduces computation and details missing in deconvolution. These deconvolutional operations restore input image's resolution from lower to higher, and finally reach the original size. Each upscaling operation also follows two convolutional layers, playing the same role to abstract features. The design of the proposed network firstly makes the encoder extract high-level abstraction features, and then the two decoders acquire pixel-wise organ probability map and boundary distance map, respectively. Furthermore, building a multi-level features extractor is very necessary for the end-to-end segmentation task of medical image. Under the conditions of high variability of organic shapes and low-contrast imaging quality in medical images, the multi-level features extractor can capture discriminative contextual information.

For the facilitation from correlated tasks, we do not treat the segmentation as a single-task problem, but investigate the distance regression as regularization, and formulate the problem as a multi-task learning framework. The deep contour-aware network [24] made its boundary identification branch as classification. This approach provides strong boundary constraints, and can increase the overall segmentation accuracy and reduce spatially isolated errors. However the classification-based auxiliary task offers discrete boundary labels, which may cause some non-smooth segmented contours. Hence, boundary distance regression is a reasonable criteria to produce continuously

numerical constraints of spatial information. The distance regression performs as complementary cues to the probability of organ classification, and then it regularizes the trouble of classification outliers.

Another question is where to branch the regression path from our baseline segmentation framework. A simple way to carry out the regression loss layer is directly from the convolutional layer that provides input for the classification score map. But this kind of network structure results in excessive overlapping layers between the main and auxiliary tasks. Accordingly, the correlated tasks can not sufficiently train their respective up-sampling path to represent the specific features. We address this issue by setting the path branching at the end of the down-sampling path, as shown in Fig. 3.4. It balances the parameters in the shared and non-shared parts of each task, and guarantees each task's up-samplings have a large number of individual feature channels for better task-specific feature learning. Meanwhile, the missing of spatial information in the shared network can be regularized by the regression network through parameters updating in backpropagation.

3.2.3 Formulation

We define the loss of the classification branch L_{cls} by applying multi-class cross entropy loss to each pixel of the output probability map. In the regression branch, we formulate the loss term L_{dis} based on the following loss:

$$L_{dis} = \frac{1}{2K} \sum_{i=1}^{K} \sum_{\mathbf{x} \in \Omega} w\left(\mathbf{x}\right) \left\| \hat{D}_i\left(\mathbf{x}\right) - D_i\left(\mathbf{x}\right) \right\|_2^2 \tag{3.1}$$

where K is the number of classes. For the i-th class, $\hat{D}_i\left(\mathbf{x}\right)$ and $D_i\left(\mathbf{x}\right)$ are the predicted and the ground truth distance maps at pixel location $\mathbf{x} \in \Omega$, where Ω is the image space, and $\subset Z^2$. w is a weight function that gives higher penalty weight to the pixels that

are farther away the organ border.

Thus, the objective function of the network is as follows:

$$\underset{\{\mathbf{W}_t\}_{t=1}^{T}}{\arg\min} \left\{ L_{cls} + \alpha L_{dis} + \beta \sum_{t=1}^{T} \|\mathbf{W}_t\|_2^2 \right\} \tag{3.2}$$

where α and β are the balance weights. T represents the number of tasks. $\mathbf{W}_\mathrm{t}$ denotes the parameters of t-th task.

Based on the proposed multi-task network and losses, the objective function will not suffer from the non-trivial solving issue mentioned in [60]. We can easily derive a standard solution for the novel ℓ_2-based regression loss. Moreover, except the two loss layers in the two branches, the rest layers use the same decoding network design, and thus Eq. 3.2 holds low model complexity and can be effectively solved through, for instance, the stochastic gradient descent (SGD) solver.

3.2.4 Energy functional based fusion scheme

A further refinement could fuse the predicted organ probability map $\hat{P}$ and boundary distance map $\hat{D}$ to obtain the final segmentation. Here, the fusion scheme minimizes an energy functional $F(\hat{P}, \hat{D})$ based on the Chan-Vese model [23]. The scheme treats the probability map $\hat{P}$ as the optimization target, and the distance map $\hat{D}$ is treated as the signed distance function φ for the model deformation. Because $\hat{D}$ has already been very close to the real organ, it can be used as a good and straightforward initialization for the fusion. With few iterations, the fusion can be quickly solved by the Euler-Lagrange equation and narrow band method. For one image, this step would only take 0.02s.

3.3 Experiment

3.3.1 Experimental settings

Besides testing the proposed method on the femur dataset, we validate two more datasets. We first test a synthetic dataset including 6000 2D simulated images (training: 4000, validating: 1000, testing: 1000). We use 3 types of geometry elements (circle, triangle and square) to construct the toy examples. The target for segmentation is combined by a circle and a triangle, which is initially located in the center of image with a stochastic offset in the x and y directions. In order to simulate complex varieties of shape, the angle, length and direction of the target are randomly set. Each toy image also includes some interferences by randomly placing several squares and circles with various sizes surrounding the target. Meanwhile, heavy Gaussian noises are added to blur all shapes. The second dataset includes 2304 2D MR femur-head junction images (training: 1368, validating: 468, testing: 468). They are radially acquired from 64 3D MR femur scans with voxel spacing $(1mm, 1mm, 1mm)$. On each 3D femur volume, we rotationally sample 36 slices around the femoral shaft, with 5 degrees interval angle. Fig. 3.5 shows the radial imaging (a common analysis method for FAI [32, 103]) to obtain the 2D femur-head junction series. The third set is built from 107 computed tomography (CT) 3D kidney images. Each kidney image is resampled and cropped, and has the same physical size $(20cm \times 20cm \times 15cm)$ with voxel spacing $(1mm, 1mm, 1mm)$. Along the axial direction of each kidney data, we totally sample 16050 2D images with 1mm interval distance (training: 9000, validating: 3000, testing: 4050). Patients randomly used in dataset 2 and 3 are independent from others. All the images are resized to 128×128, and their pixel intensity is linearly normalized in [0, 1].

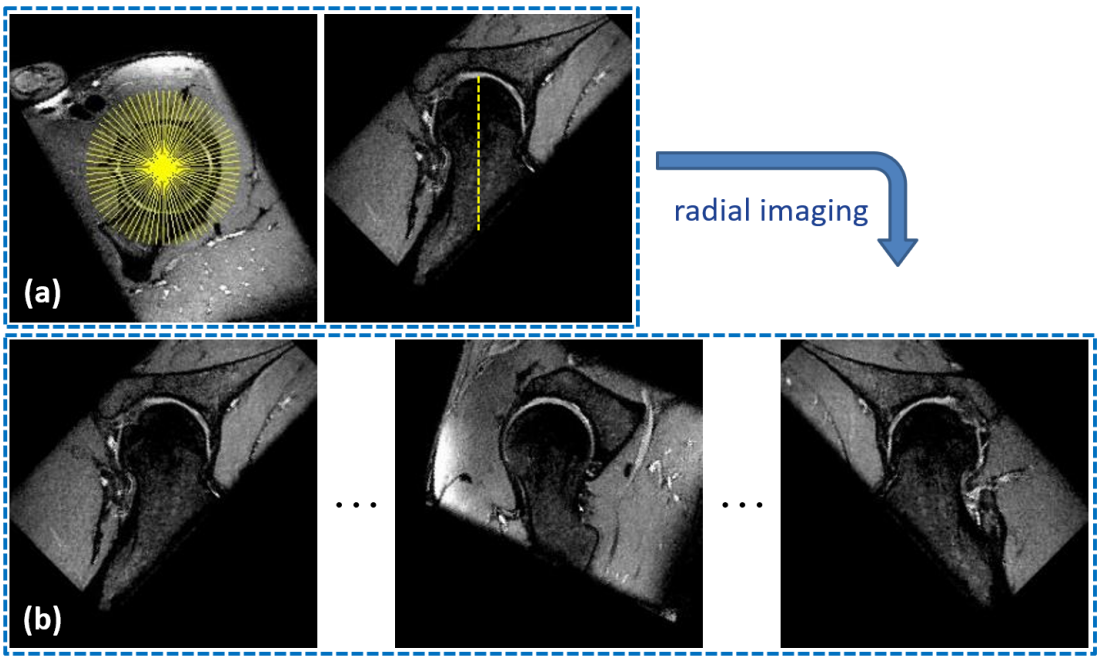


Figure 3.5: Radial imaging of femoral series. The left of (a) shows the rotated sampling in cross-section axis, and the right half displays the axis of rotation (yellow dashed line) in longitudinal-section axis. (b) are the radially acquired images from a 3D data.

Two state-of-the-art medical segmentation approaches are evaluated with our method. One is the UNet [105], and the second is the deep contour-aware networks for accurate gland segmentation (DCAN [24]). For validation, dice similarity coefficient ($DSC = \frac{2TP}{2TP+FP+FN}$) and relative error ($RE = \frac{FP+FN}{TP+FN}$) between the ground truth (GT) labels and segmentation results are reported. TP, FP and FN are the number of pixels correctly identified, incorrectly identified and incorrectly rejected respectively. The mini-batch is employed in the training phase, and its size is set to around 80 for each training of the compared methods. We use a momentum of 0.9 and a learning rate initially set as 0.001 (multiplied by a factor of 0.95 every 10,000 iterations).

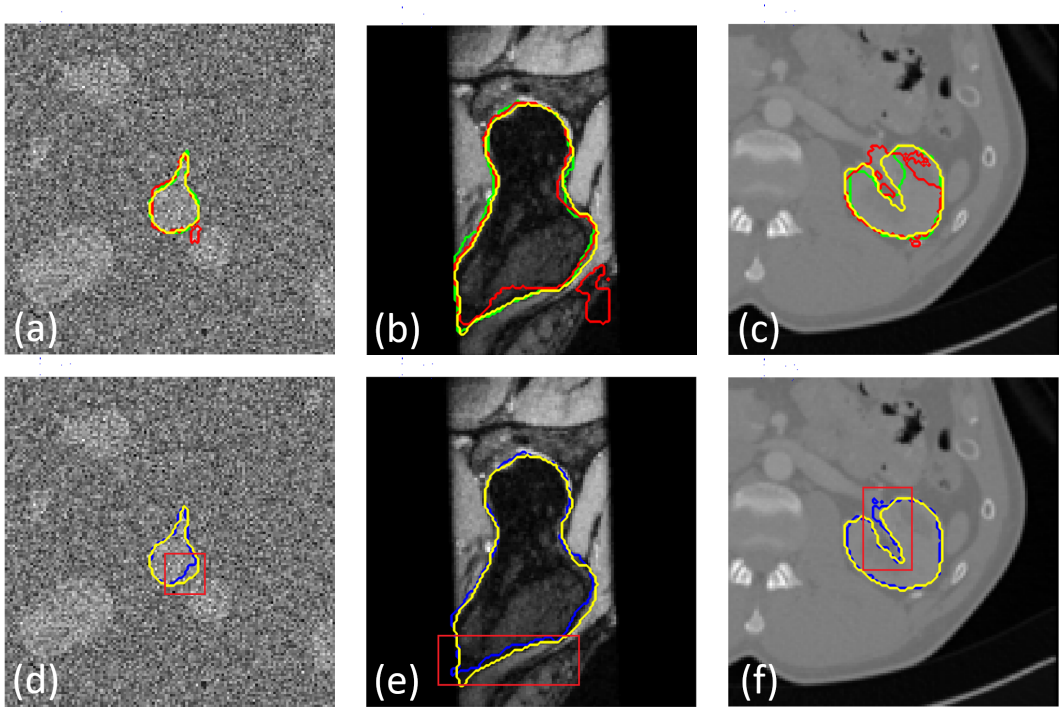


Figure 3.6: 2D visual comparisons for simulated data, femur and kidney. Green, red, blue and yellow lines are for the GT, UNet, DCAN and proposed method, respectively.

3.3.2 Experimental results

Fig. 3.6 (a)-(c) show visual comparisons between the proposed method and the UNet model for three cases. Shown by these, the UNet model can locate the position of organ (or target) correctly. Its segmented regions capture most of the correct tissue areas, and get good quantitative measurements. Yet its results suffer from leakages to the surrounding areas, where have similar pixel intensity to the targets. On the other hand, the classification branch of our method is the same as the UNet, but with the additional regression branch and the joint training, the proposed approach can prevent the leakage issue, and thus obtain better total segmentation performance. Since the few visual comparisons may not reflect the overall performance clearly, quantitative comparisons of overlapping accuracies are shown in Table 3.1.

Table 3.1: Quantitative comparisons.

Method	Simulation		Femur Bone		Kidney	
	DSC	*RE*	*DSC*	*RE*	*DSC*	*RE*
UNet	0.90	0.19	0.90	0.19	0.70	0.62
DCAN	0.92	0.15	0.91	0.18	0.83	0.31
Ours	**0.96**	**0.09**	**0.93**	**0.14**	**0.90**	**0.19**

After showing the effectiveness of the proposed model to prevent leakages, we also visually compare the boundary smoothness of results by the DCAN and proposed method. In order to have a better view, only the segmented contours by the two approaches are plotted for the same cases in Fig. 3.6 (d)-(f). The DCAN method does not show significant leakage problem, and obtains higher total segmentation accuracy comparing with the UNet in Table 3.1. However, by considering the boundary smoothness shown in the red boxes, the proposed method achieves better performance. The two methods both utilize multi-task strategy to preserve shape, but in the DCAN, its boundary classification task offers discrete boundary labels which may cause non-smooth boundary. In our regression branch, the boundary distance regression could provide continuously numerical constraints of spatial information during optimizing the regression loss. Hence the proposed method could produce higher smoothness on boundary. Besides the visual comparisons, the overall quantitative measurements between the two approaches are shown in Table 3.1.

3.3.3 Extensive experiments of femoral head-neck morphology

With the help of the proposed method, researchers could efficiently and accurately complete the measurements of femoral head-neck morphology visually demonstrated in Fig. 3.7. These demonstration results are from an experiment on 50 3D MR femur-neck data. In the future, the detected morphological characteristics (e.g., concavity

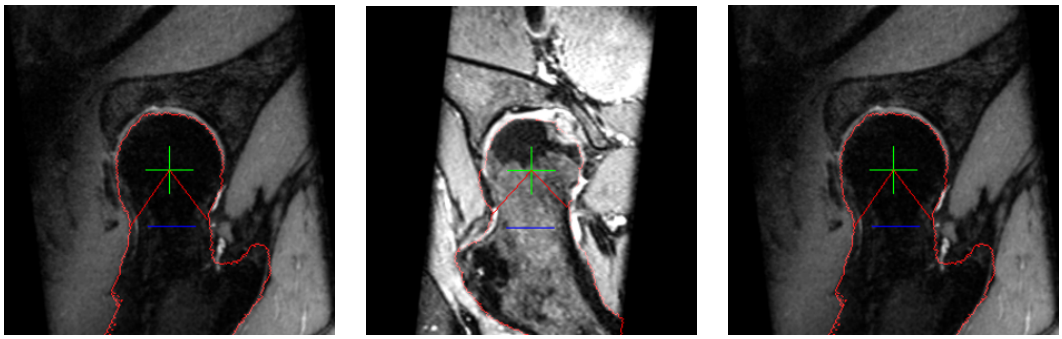


Figure 3.7: Measurements of the femoral head-neck morphology. In each demonstration image, the green cross shows the center of fitted circle on femoral head. The red curve is segmented by the proposed method, and two straight lines intersect with the curve at the two concavities of femoral neck, respectively. The blue line is used to help visually indicate the horizontal direction

of femoral neck) could be employed to extend some FAI abnormality analyses (e.g., in [32]) for large-scale dataset, to investigate the longitudinal correlations of hip OA.

3.4 Conclusion

In the present work, we propose a deep multi-task network for robust shape preserved organ segmentation. The network has a unified architecture to formulate organ segmentation as multi-task learning that combines both region and boundary identification. This multi-task learning with the novel boundary distance regression can alleviate spatially isolated segmentation errors as well as ensure the smoothness of segmented contours. The proposed deep network is designed as a "Y" shape, bifurcated at the end of the encoding path. Hence the shared encoding and non-shared decoding paths have balanced layers and parameters for each task branch.

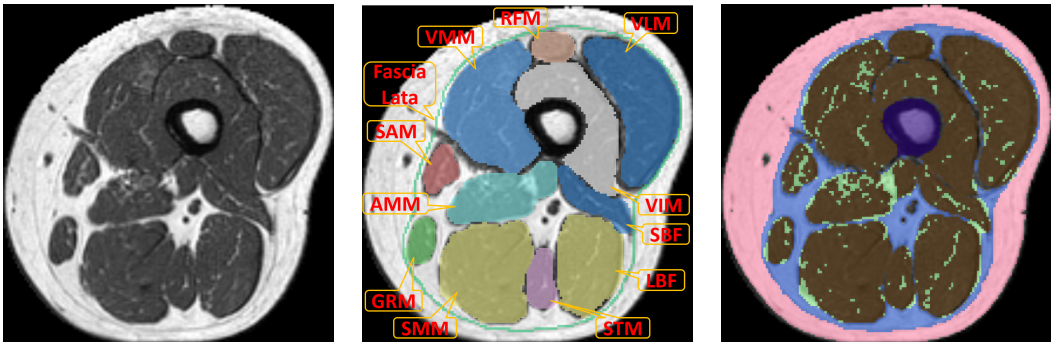
Chapter 4

Towards Large-scale MR Thigh Image Analysis via An Integrated Quantification Framework

4.1 Introduction

In previous chapters, deep learning based segmentation methods have been discussed. However, deformable models and machine learning based algorithms are still practical and effective techniques for medical image segmentation and tissue quantitative analysis. In this chapter, an integrated framework is discussed for accurately quantifying major tissue composition in MR thigh images. In the integrated framework, a data-driven and sparsity-constrained deformable segmentation is proposed, and a joint label fusion based multi-atlas labeling is also utilize, to enable robust quantification. Quantification of all the major thigh tissue plays a critical role in various medical data analysis tasks, e.g., the analysis of physical performance and the progression/diagnosis of knee osteoarthritis.

In Fig. 4.1, the magnetic resonance (MR) image illustrates the tissues in thigh. Thigh intra-muscular adipose tissue (IAMAT) is defined as the AT visible between muscle fibers. Corresponding to the IAMAT, thigh inter-muscular adipose tissue (IMAT) lies within the fascia lata (a fibrous membrane giving off sheath to the thigh muscles) surrounding the leg musculature. Different from subcutaneous adipose tissue (SAT) in the thigh, IMAT encompasses and permeates skeletal muscles, with which it shares a



(a) (b) (c)

Figure 4.1: Schematic diagram of thigh tissue.

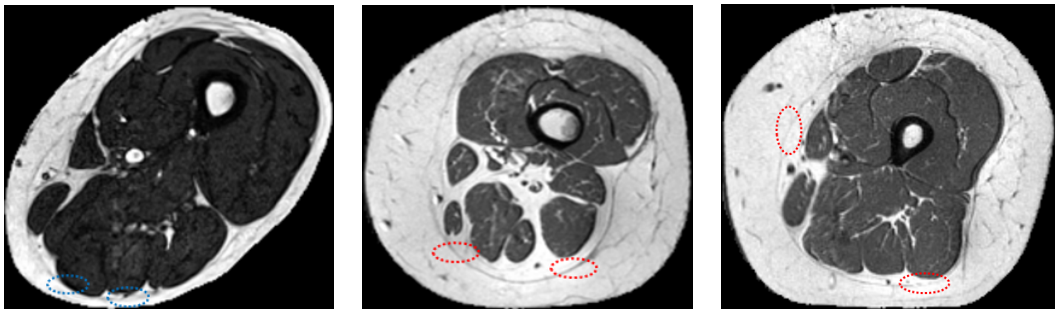
direct vascular connection [33]. The curve of fascia lata is demonstrated as the closed dark-green contour in Fig. 4.1 (b), which is a fibrous membrane giving off sheaths to the thigh muscles. Encircled by the fascia lata, the colored labels demonstrate the 11 individual muscle sub-regions: vastus lateralis (VLM), rectus femoris (RFM), vastus intermedius (VIM), vastus medialis (VMM), sartorius (SAM), gracilis (GRM), adductor magnus (AMM), semimembranosus (SMM), semitendinosus (STM), long biceps femoris (LBF), short biceps femoris (SBF). The boundaries of these sub-regions provide explicit divide of IMAT and IAMAT. Fig. 4.1 (c) further shows different labels of the major thigh tissue composition: skeletal muscle (brown), SAT (pink), IMAT (blue), IAMAT (green), bone and marrow (purple). Bone surrounds marrow, having low and high pixel intensity in MR images, respectively. With high pixel intensity, SAT, also called superficial fascia, is the adipose layer between the dermis and the deep fascia around the thigh muscles. Skeletal muscle, or leg musculature, has intermediate intensity. IMAT and IAMAT are defined as the adipose tissue (AT) visible between muscle groups and muscle fibers, respectively. Their pixel intensity are both high.

Research in [78] showed that thigh IMAT and IAMAT appear to blunt the adaptive muscle quality. Kumar et al. [62] indicated people with higher IAMAT is more related

to symptomatic and structural severity of knee OA. While Maly et al. [77, 76] revealed the relationship between IMAT volume and knee extensor strength and physical performance among women with or at risk of knee OA using magnetic resonance imaging (MRI). Moreover the further study [29] displayed increase of IMAT content is associated with structural and pathological progression of knee OA in women. Additionally, analysis on quadriceps in [67] found that subjects with knee OA have significantly less quadriceps strength relative to body mass index (BMI). Thus, an accurate quantitative assessment of thigh tissue is crucial for the purpose of clinical OA research.

However, with respect to knee OA, the relevance of skeletal muscle and adipose tissue parameters such as absolute or relative volume is quantitatively indistinct. A key reason is the lack of effective approaches to quantify these kinds of thigh tissue, and sufficient experiments with considerable data. Manual delineation and measurement may work well for a single MR slice, yet it is time-consuming, operator-dependent, and inconsistent for numerous multi-slices or 3D data. On the other hand, in recent years, several techniques have been proposed to perform automatic thigh tissue assessments. A 2D segmentation by a gradient vector flow (GVF) based active contour could assess IMAT and other tissues in thigh [96]. Makrogiannis et al. [75] incorporated parametric deformable model and unsupervised tissue clustering for an improved IMAT extraction. [119] used a model based on k-means clustering and mathematical morphology to classify and delineate IMAT and SAT. By utilizing a variational Bayesian Gaussian mixture model, Tan et al. [112] proposed a more reliable framework to obtain IAMAT, IMAT and other thigh tissues from multi-type imaging sources.

So far, all the above approaches still have space for improvement: (I) The first point is the distinction of IMAT and SAT. Most of the above-mentioned existing methods



(a) A lean case (b) An obese case (c) Weak fascia lata edges

Figure 4.2: (a) In a lean case, fascia lata (blue marked regions) is close to muscle regions. (b) The red marked regions show fascia lata is away from muscle boundary in a obese case. (c) Weak fascia lata edges are shown in the red marked regions.

assumed that the contour of the general muscular region is the boundary separating SAT and IMAT, if we call the enclosed area including all the skeletal muscles as the general muscular region. Since the fascia lata is close to the muscle groups in normal subjects (as shown in Fig. 4.2 (a)), such an assumption is acceptable. However, this assumption does not hold for pathological subjects, because the fascia lata may not be close to muscle regions (as shown in Fig. 4.2 (b) and (c)). Although Orgiu et al. [89] revised the separation of IMAT and SAT by snake-based muscle fascia segmentation, this approach still has the parameter sensitivity issue in setting the 2D active contours, and may merely work for a small dataset. (II) The second is how to explicitly discriminate IMAT and IAMAT. Prescott et al. [99] integrated thresholding and morphological skeletonization to locate IAMAT. While Tan et al. [112] utilized level-set-based segmentation to estimate the connectivity of the adipose tissue to muscle, and then implicitly separate these two adipose. Due to the unclear appearance cues or high shape variation of thigh tissue, these unsupervised strategies might still require manual intervention when processing large dataset, which causes serious inconvenience. (III) The last but not least one is the experiment setting. Some tissue analyses of existing frameworks were only based on very few test samples. So adequate subjects with

different pathological and temporal conditions should be included into the analysis to show a clear relevance of different thigh tissue for knee OA.

Therefore, in this work, we focus on a novel integrated framework to perform 3D segmentation and classification of the major thigh tissue composition using volumetric mid-thigh axial T1-weighted MR images. In particular, we present supervised tissue discrimination techniques utilizing small amount of supervised information in learning to locate the borders that separate IMAT from SAT and IAMAT. Specifically, in preprocessing, a boosted cascade with the Haar-like is trained to locate a sub-window of femur, and the negative effect of femur for the quantification can be eliminated. According to the boundary of the general muscular region acquired by an efficient unsupervised tissue initialization, we formulate a novel detection-driven and sparsity-constrained deformable model to obtain accurate labeling of fascia lata, as well as integrate an individual skeletal muscles segmentation with joint label fusion based multi-atlas labeling. The segmented fascia lata and muscle sub-regions instruct explicit divide of all the thigh tissue. The proposed method permits finer control over adaptivity allowing the 4 primary tissue to be reasonably segmented rather than just unsupervised tissue distinction approaches as in some existing papers.

The main contribution of this work is twofold. (1) A comprehensive framework is designed to discriminate thigh tissue in a supervised way. Especially, to our knowledge, the proposed method is the first one to integrate both the fascia lata and individual muscles segmentations, and hence give a practical distinguishment of the three different adipose tissue: IMAT, IAMAT and SAT. (2) The framework is applied to test a large dataset with hundreds of temporal related subjects, to investigate the knee OA related changes of thigh muscle and adipose tissue volume.

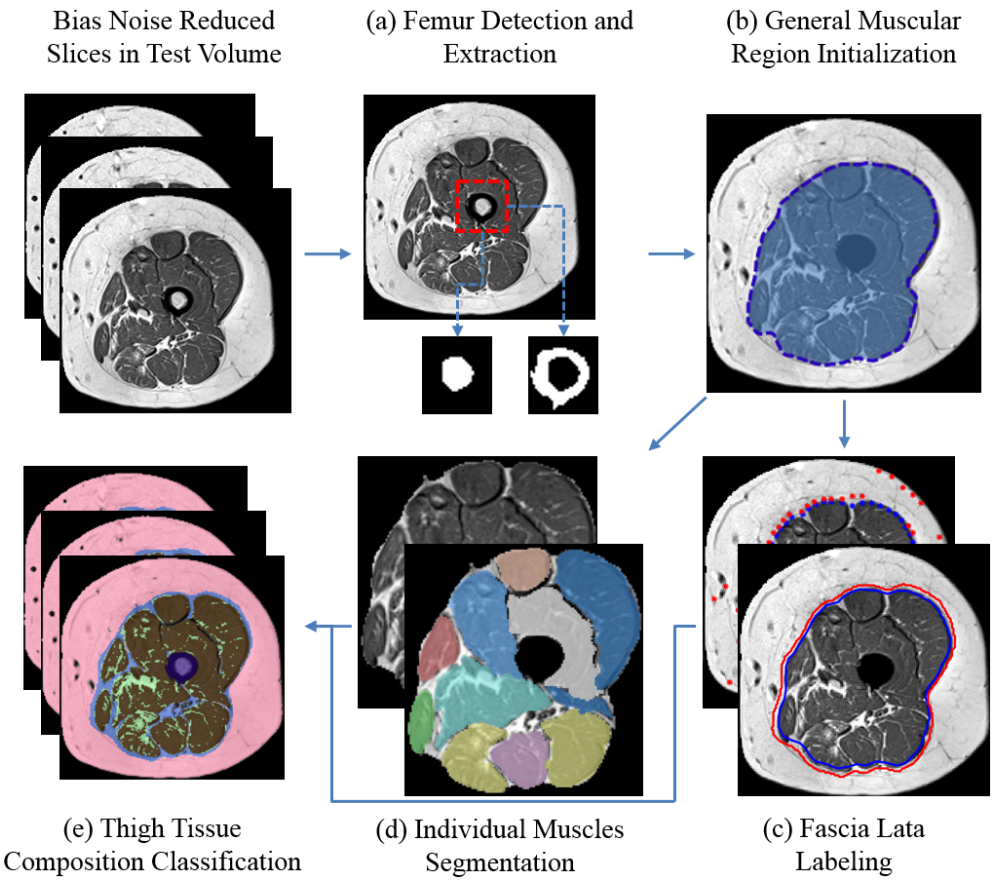


Figure 4.3: Flowchart of the thigh tissue quantification framework.

4.2 METHODS

Accurate labeling of the skeletal muscle, IMAT and IAMAT in thigh benefits the quantification of thigh tissue composition, and even provides quantitative basis for the diagnosis of knee OA. Thus, we formulate this problem as a series of detection, segmentation and classification work, and present a novel integrated framework. Start from noise reduced data, we robustly extract femur by a boosted cascade with Haar-like features, and hence the negative effects of femur, which may mislead the follow-up quantification, are eliminated. Then the general muscular region is efficiently obtained through an unsupervised tissue segmentation integrating global image context and shape refinement. Meanwhile, based on the remained image areas, we segment fascia lata as well as

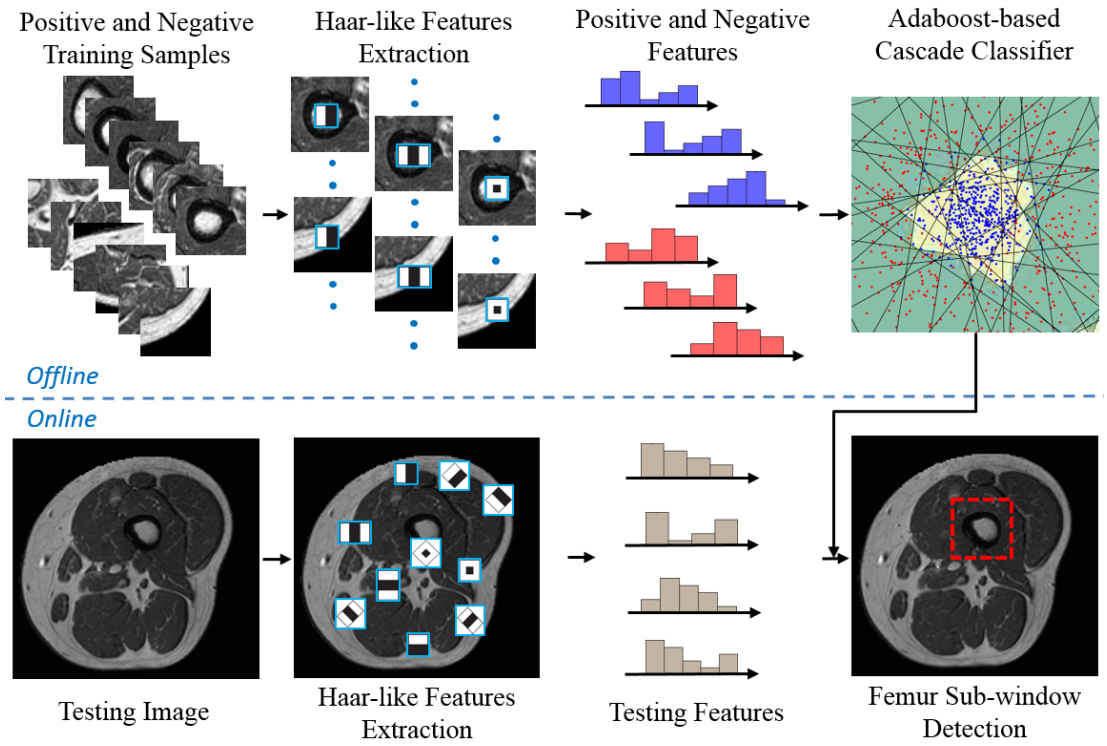


Figure 4.4: Overview of the femur sub-window detection.

individual skeletal muscles by a data-driven and sparsity-constrained deformable model and label fusion based multi-atlas technique, respectively. After delineating the boundaries between individual muscle regions and IMAT, we utilize a statistical classifier to separate IAMAT from skeletal muscles. The demonstration of the algorithm flowchart is shown in Fig. 4.3.

4.2.1 Femur Extraction

Femur, shown as the purple label in Fig. 4.1 (c), includes marrow encompassed by bone (or called cortical bone). We propose a coarse-to-fine strategy to extract femur and benefit the AT, muscle quantification.

In the first step, we train a detector to locate the sub-window of femur. Fig. 4.4 shows the overview of the femur sub-window detection. In the offline or online stage,

the Haar-like features are extracted on each positive/negative patch or testing image. The small boxes consisting of white and dark rectangles represent the edge, line and center-surround feature prototypes (total 14 types) stated in [69]. In the map of the trained Adaboost-based cascade classifier [56], the red and blue points represent the positive and negative samples, respectively, and the black line is the decision border of each cascade classifier. Accordingly, the yellow area is the acceptance field, while the green area represents the denial field in the feature space. Finally, a local femur sub-window is detected shown as the red-dashed box.

In the second step, we separate marrow and bone in the detected sub-window by an intensity-based Gaussian mixture model (GMM) with K Gaussian components. This model describes the probability distributions of K classes of tissue components in whole thigh or part of thigh. In this case, the K is 3, representing cortical bone, skeletal muscle and marrow, which have low, middle-level, and high pixel value, respectively. The intensity-based classification is merely carried out on the local femur region without strong interference from other tissue compositions, so we can very robustly obtain the femur area and exclude it from the follow-up steps. The separated marrow and bone labels are demonstrated in Fig. 4.3 (a).

4.2.2 Unsupervised General Muscular Region Initialization

Reliable general muscular region initialization is a key step to provide baselines for both the fascia lata labeling and the individual muscles segmentation. In the T1-weighted MR thigh images, the pixel intensity of the primary anatomical tissues (adipose, muscle) have significant differences. The adipose has high pixel value, while the value of the muscle is at intermediate level. Additionally, SAT and muscles have considerably

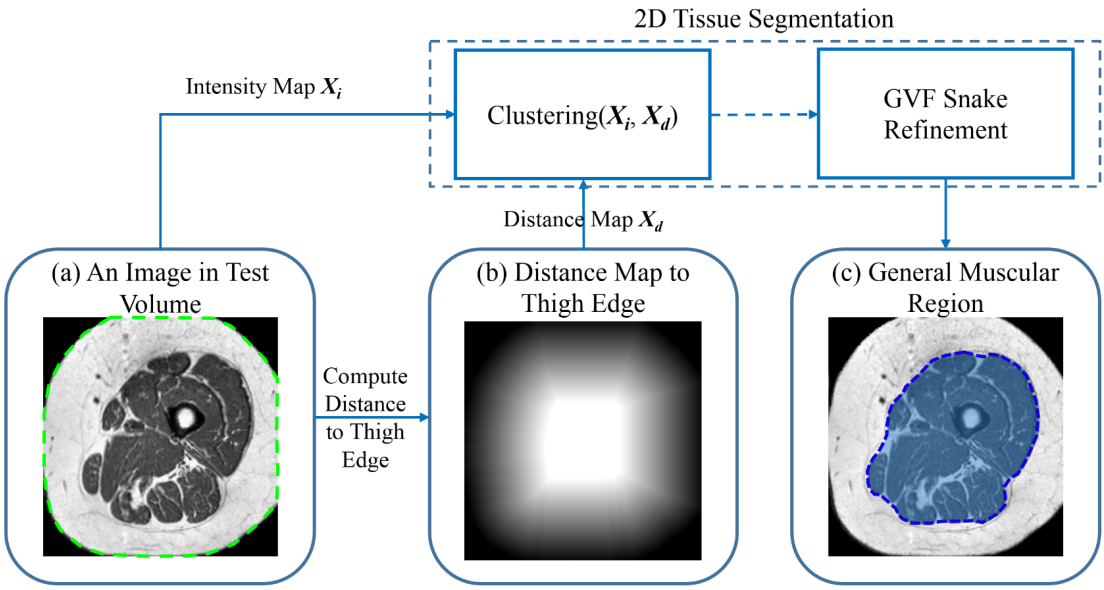


Figure 4.5: Unsupervised general muscular region extraction.

consistent spatial distributions. SAT, shown as the pink label in Fig. 4.1 (c), is a kind of adipose tissue between the skin (green-dashed contour in Fig. 4.5 (a)) and muscles, and firmly encircles the muscle region (brown label in Fig. 4.1 (c)). Thus, we employ intensity map $\boldsymbol{X}_i$ and distance map $\boldsymbol{X}_d$ (for each pixel in the thigh, calculate its closest distance to the skin) as features. Then utilize a common clustering method (e.g., fuzzy c-means [13]) to roughly label the general muscular region. These two features constrain the contour of the intermediate labels to be connected. At last, we use this contour to set a GVF-snake [131] to evolve and obtain a contour-refined general muscular region. The initialized model V_0 is defined as a set of meshless nodes on the muscle region contours across all axial images. The initialization is shown in Fig. 4.5.

4.2.3 Fascia Lata Labeling

As the pathological changes, the initialized boundary of the general muscular region can not accurately distinguish the IMAT and SAT. Thus, capture the fascia lata and

reconstruct its surface is an essential step in the thigh tissue quantification. The provided fascia lata segmentation approach includes two sub-steps: (1). A random forest based detection using narrow-band and sparse strategy to search a set of points as fascia lata prior; (2). A sparsity-constrained deformable model to effectively suppress the detection outliers in the fascia lata prior, and obtain the surface.

Fascia Lata Prior Detection

We define the fascia lata prior as a set of control points obtained from the learned detector. The dataset used in the paper consists of volumetric mid-thigh axial MR scans. The axial scanning makes fascia lata in adjacent slices strongly correlated, so we learn a 2D fascia lata detector by a random forest [16]. We use points on the manually labeled fascia lata as positive samples. From each positive point, we collect f points forward and b points backward along its normal direction at equal intervals as the negative samples. To detect fascia lata, both local and global features at each point are extracted. We extract steerable features [150] as the local feature. To compute the steerable features of a image pixel, a few points are sampled from the image patch centered at the pixel using a regular sampling pattern. Then local (intensity-based and gradient-based) features are extracted for each sampling point. The steerable features embed orientation and scale information into the distribution of sampling points, while each individual feature set is locally defined. In this way, the transformation of image patch can be well captured by the transformation of the sampling pattern. For the global feature, we put the origin of the global coordinate at the thigh mass center $\mathbf{p}_o = (x_o, y_o)$. The x-axis $\mathbf{p}_x$ and y-axis $\mathbf{p}_y$ are parallel to the image's horizontal and vertical axes, respectively. The global feature of pixel at $\mathbf{p}_s$ is represented as a relative

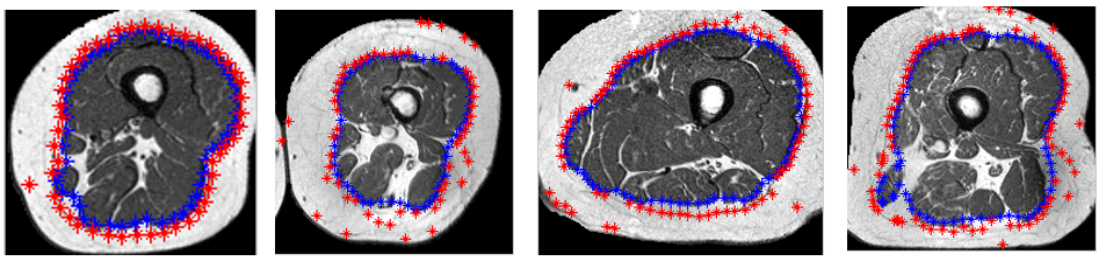


Figure 4.6: Detection results. The blue star symbols are the initial points (blue) in C_0, and the red ones are the detected control points C_e.

angle θ between $\mathbf{p}_s - \mathbf{p}_o$ and $\mathbf{p}_x$.

In detecting, we consider two optimizing strategies. The first is narrow-band processing, which restricts most computations into a thin band, obviously improves the searching efficiency. Because the fascia lata, as shown in Fig. 4.2, is near or attached to the outer boundary of muscles, we define the narrow band as the range surrounding the initialized model. Another practical strategy is to detect a sparse set of control points. Instead of detecting the control points for all the n vertices in the initialized model V_0, we only detect for a subset $C_0 \subset V_0$ which contains $m(\ll n)$ vertices. The detected result is denoted by C_e. To catch the control point $q^e \in C_e$ of a certain vertex $q \in C_0$, the 2D detector searches along q's normal direction in the 2D image for the position with highest probability to be the fascia lata. So there is a one-to-one correspondence between the points of the two sets, C_0 and C_e. Fig. 4.6 gives some detection results.

Fascia Lata Contour Reconstruction

For the fascia lata surface reconstruction, we utilize a deformable model and formulate it as an energy equation problem as:

$$E_{total} = E_{data}\left(V_d\right) + E_{prior}\left(V_d, V_e\right) \tag{4.1}$$

Here, V_e denotes a meshless model inferred by C_e. Starting from the initialized model V_0, we compute the deformed model V_d, which is driven by C_e to best approximate V_e. Thus, the data energy term E_{data} aims to preserve the local shape structure in V_d by using the Laplacian coordinate. The prior energy term E_{prior} is used to minimize the distance between the current model and the control model V_e.

A vertex $i \in V_d$ is represented by the 3D homogeneous coordinate $\mathbf{v}_i = (x_i, y_i, z_i, 1)^\top$, and $\top$ stands for transpose. A control point $k \in C_e$ is represented by the 3D cartesian coordinate $\mathbf{v}_k^e = (x_k, y_k, z_k)^\top$.

A. Data Term

The proposed model deforms each meshless vertex $i \in V_d$ by a transformation matrix $T_i \in R^{3\times4}$. The T_i is a special linear matrix defined in [144]. As formulated by Eq. 4.2, the matrix includes a scalar a_i for isotropic 3D scaling, a translation vector (p_i^x, p_i^y, p_i^z), and parameters (h_i^1, h_i^2, h_i^3) for linear approximation of rotations with small angles.

$$T_i = \begin{pmatrix} a_i & -h_i^1 & h_i^2 & p_i^x \\ h_i^1 & a_i & -h_i^3 & p_i^y \\ -h_i^2 & h_i^3 & a_i & p_i^z \end{pmatrix} \tag{4.2}$$

In our meshless model, each vertex has a Laplacian representation defined by the distances to its neighbors. Since the transformation matrices (e.g., T_i and T_j) of two neighboring vertices (i and j) should be similar for a local shape-preserving deformation, we regularize the difference between the deformed neighboring vertices. Thus, the energy function of the data term is defined as:

$$E_{data} = \sum_{i \in V_d} \sum_{j \in \mathcal{N}(i)} \|T_i \mathbf{v}_i - T_j \mathbf{v}_j\|_2^2 \tag{4.3}$$

where $\mathcal{N}(i) \subset V_d$ is the neighboring node set of vertex i, with s closest nodes. As a result of minimizing the summation of the deforming difference, the energy term ensures the local similarity and thus preserve shape details.

B. Prior Term

With the detected control points C_e, the proposed model can be robustly guided to the estimated location. To obtain more accurate fascia lata reconstruction, we define the energy function of the prior term as:

$$E_{prior} = \sum_{k \in C_e} \|T_k^m \mathbf{v}_k^m - \mathbf{v}_k^e\|_2^2 \tag{4.4}$$

Here, each control point $k \in C_e$ with 3D cartesian coordinate $\mathbf{v}_k^e$ has an one-to-one-matched vertex in V_d with 3D homogeneous coordinate $\mathbf{v}_k^m$. T_k^m is the transformation matrix of this vertex (in V_d) corresponding to control point k. The prior term measures the distance from C_e to V_d. So by minimizing this term, the deformable model evolves to the detected fascia lata.

C. Energy Formulation

Integrating both ℓ_1 and ℓ_2 regularization as [147, 120] can balance the model deformation and the suppression of gross outliers from the control points detection. Thus, we combine the energy term (4.3) and (4.4), and sovle the deformable model by minimizing the total energy (Eq. 4.1) as:

$$\underset{\forall i, T_i; \mathbf{e}^C}{\arg\min} \left\{ \sum_{i \in V_d} \sum_{j \in \mathcal{N}(i)} \|T_i \mathbf{v}_i - T_j \mathbf{v}_j\|_2^2 + \alpha \sum_{k \in C_e} \|T_k^m \mathbf{v}_k^m - \mathbf{v}_k^e - \mathbf{e}_k\|_2^2 + \beta \left\| \mathbf{e}^C \right\|_1 \right\} \quad (4.5)$$

where α and β are balance weights. $\mathbf{e}_k$ represents the detection gross error of a control point k, and $\mathbf{e}^C = ((\mathbf{e}_1)^\top, ..., (\mathbf{e}_m)^\top)^\top$ is a restructured error vector.

4.2.4 Individual Muscles Segmentation

With the detection and reconstruction of fascia lata, we can capture the important anatomic trail and distinguish between IMAT and SAT in thigh. However, the thigh tissue quantification still requires another key step, the discrimination of IMAT and IAMAT, and hence provides an exact assessment of the two different adipose tissue. IMAT lies within the fascia lata and encompasses skeletal muscles, while IAMAT is the adipose located between muscle fibers. If we could get all the 11 individual skeletal muscles, the segmented sub-regions would not only guide the skeletal muscle assessment, their boundaries would also give explicit divide of these two muscular adipose.

In order to segment a target image I_t, we define an atlas-based segmentation with r pre-registered atlases: $A_1 = (I_1', L_1'), ..., A_r = (I_r', L_r')$, I_i' denotes the ith on-line training image nonlinearly registered to I_t, and L_i' is the corresponding warped manual labels. The result is an estimated label $\hat{L}_t$ for a pixel (indexed by x) in I_t is usually obtained by weighted summary as:

$$\hat{L}_t(x) = \sum_{i=1}^{r} w_i(x) L_i(x) \quad (4.6)$$

here $w_i(x)$ is a weight of pixel x assigned to ith atlas, and $\sum_{i=1}^{r} w_i(x) = 1$.

However, this weight-determination way might have limitation, for it independently assigns weight to each atlas. Wang et al. [121] stated that labeling errors produced by

(a) (b) (c)

Figure 4.7: A demonstration on individual muscles segmentation. (a) A slice from a femur and SAT removed input data. (b) 11 individual muscle regions are extracted by the joint label fusion based atlas method. (c) Skeletal muscle (brown) and adipose tissue (IAMAT) (green) are separated based on the segmented sub-regions.

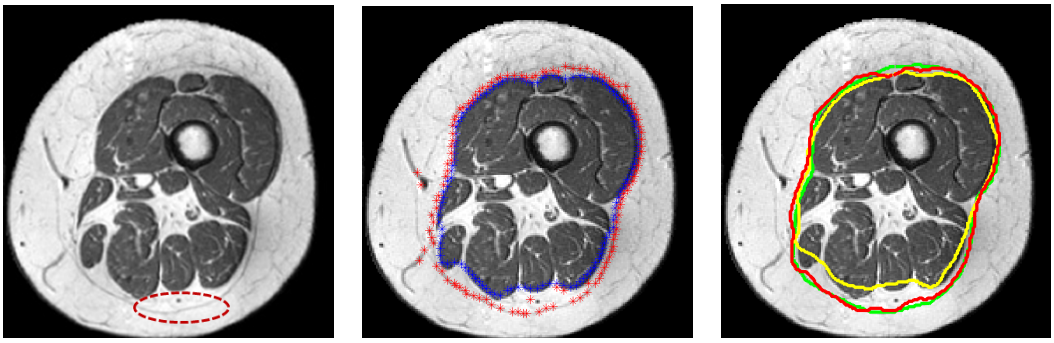
different atlases may be correlated, and hence they modeled a novel atlas-based segmentation by minimizing the total expected error between $\hat{L}_t$ and the true segmentation L_t. The joint label fusion can effectively correct labeling errors from the individual muscles segmentation. This strategy obtains the fusing weights by minimizing a total expected error which takes mutual dependency of each atlas into consideration, so it can process thigh muscles with more variants of shape and appearance. Moreover, because of this advantage, this muscle segmentation merely require a small training set for on-line registration and fusion. Secondly, the fusing weights can be very fast solved in a closed form, which is a key acceleration for 3D thigh data processing. After removing some segmentation errors through morphological operators and connected components, in the muscle regions, we employ the pre-defined VBGMM classifier from subsection 4.2.1 to discriminate between IAMAT and muscle. In this case, we set $K = 2$, representing these two types of tissue. Fig. 4.7 demonstrates the muscle and adipose labeling results.

4.2.5 Thigh Tissue Composition Classification

Now, based on all the intermediate results from the above steps, we can classify all the compositions of thigh tissue. (i) From the noise-reduced input data U_i, the femur U_f is detected and can be excluded from the tissue quantification. (ii) The fascia lata border segmentation can adjust the rough SAT labeling results from subsection 4.2.2, and acquires exact SAT partition U_s. (iii) Based on the remained volume $U_r = U_i - U_f - U_s$, the individual muscles segmentation explicitly discriminates between IMAT and IAMAT. Thus we can obtain obtain the accurate labeling of muscle tissue as U_m, and that of IAMAT as U_{ia}. (iv) The partition of IMAT is computed as $U_{ima} = U_r - U_m - U_{ia}$. Last, the integrated framework quantifies skeletal muscle and adipose tissue (IMAT, IAMAT, SAT) in thigh.

4.3 Experiment

In this part, we first introduce the validation of fascia lata reconstruction and individual muscles segmentation, including visualization of segmentation, quantitative and statistical accuracy comparisons. Then, hundreds of subjects with different pathological and temporal conditions are included to build a radiographic OA related quantitative analysis of thigh tissue composition in detail. All the data using in the experiment part are mid-thigh MR volumetric scans from the public Osteoarthritis Initiative (OAI) database. All the MR scans are axial T1-weighted, aged from 45 to 82. Images have 0.98mm or 0.78mm in-plane resolution and 5.0mm between-slice resolution. The in-plane sizes range from 145 $\times$ 125 to 274 $\times$ 249, and the slice numbers are all 15. Due to strong bias errors on most right legs, we only extracted and processed left legs. All the modules in this section are implemented in Matlab and C++, tested on a computer



(a) (b) (c)

Figure 4.8: 2D visual results on fascia lata segmentation. (a) A slice from original thigh volume. (b) Blue and red star symbols are the initialized and detected control points by the proposed method. (c) Red, yellow and green contours are from our model, MDM and GT, respectively.

with 4.0 GHz Intel Core i7 4790K and 32Gb RAM and 4G-RAM GPU.

4.3.1 Training and Validation of Fascia Lata Reconstruction

A dataset containing 40 subjects is used for the training of the 2D fascia lata detector and the validation of the fascia lata reconstruction. The training set of the detector consists of 90 2D thigh images with ground truth labeled by experts. These images were extracted from 30 volumetric scans. For each volumetric scan, we selected 3 cross-sectional slices from the top, middle, and bottom of the middle thigh, respectively. We set parameters $f = 4$ and $b = 3$ for the negative points collection. The random forest has 500 trees. The number of predictors sampled for splitting at each tree node is square root of the feature number. The rest 10 labeled volumes are used for the reconstruction validating. For evaluation, we fill the interior of fascia lata, and report the mean and standard deviation of DSC and RE. Besides, the distance errors between the reconstructed surface and the ground truth surface were also reported.

Fig. 4.8 shows some results of a subject (female, 63 years old). A comparative method, Metamorphs based deformable model (MDM) from [112] is involved in the

Table 4.1: Quantitative comparisons of fascia lata reconstruction.

Method	DSC		RE		Dist Errs	
	μ	σ	μ	σ	μ(mm)	σ
Initial	0.9503	0.0079	0.0951	0.0150	3.49	0.51
MDM	0.9574	0.0082	0.0851	0.0174	3.45	0.38
Ours	0.9784	0.0137	0.0438	0.0292	2.14	0.35

test. This method extracts the boundary between SAT and IMAT by a modified parametric deformable model, integrating a region-of-interest-based balloon term of the Metamorphs model [52]. The MDM dynamically estimates tissue's intensity probability map in unsupervised way to generate a region based force. This force can enhance external force and facilitate the model convergence. When the MDM runs on the regions without loose fascia lata, it can obtain a strong region force and work well because the fascia lata is located at muscle's edges. However, in the red circled region of Fig. 4.8(a), the fascia lata is loose from the muscle boundary. In this case, the loose fascia lata is excluded from the probability map, so the MDM can not obtain an effective region force to capture the whole fascia lata contour. Moreover, segmentation using insufficient meshes may make the MDM capture imprecise boundaries of muscles. Thus, the MDM has under-segmentation issue (yellow contour in Fig. 4.8(c)). On the other hand, although the testing case has the loose symptom, our method can detect most of the control points on the fascia lata (Fig. 4.8(b)), which provides accurate cues for the driving forces. Furthermore, because of the sparsity constraints, outliers do not have much negative impacts on the model deformation.

Quantitative comparisons are shown in Table 4.1. We report *DSC* and *RE* of the initialized model in subsection 4.2.2, the MDM, and the proposed approach. For the distance errors (Dist Errs), we first calculate the closest distance to the GT at each point on the resulting surfaces, and then compute the means and the standard deviations.

Table 4.2: Quantitative comparisons of VLM, RFM and VIM segmentation.

Method	VLM		RFM		VIM	
	DSC	*RE*	*DSC*	*RE*	*DSC*	*RE*
MV	0.8376	0.3281	0.3967	1.0896	0.8239	0.3450
MA-EASA	0.8459	0.2988	0.2268	0.9994	0.8216	0.3841
Ours	0.9240	0.1585	0.6324	0.8280	0.8983	0.2062

4.3.2 Validation of Individual Muscles Segmentation

On the other hand, for the muscles segmentation, another data set with 8 expert-labeled atlas pairs is utilized for Atlas on-line training. The subjects selected for this training set are verified as representative of the while experimental patients with respect to age, gender, and appearance of thigh muscles. In the validation of the segmentation approach, we conduct leave-one-out cross validation on the 8 training subjects, and evaluate the mean of *DSC* and *RE* on the segmented regions of 11 individual thigh muscles and their total volume (TotM). The comparisons include the common majority voting (MV) [48] and a state-of-the-art Atlas segmentation [135] which is a hybrid of multi-atlas and extended adaptive statistical atlas (MA-EASA).

Quantitative results are reported from Table 4.2 to 4.5. The proposed muscles segmentation achieves the best performance, comparing with the MV and MA-EASA. Especially for the TotM estimation, its DSC score is larger than 0.95, which provides an important foundation to assure the correctness of IMAT and IAMAT analysis for the knee OA diagnosis. However, for the analysis of separate muscle sub-region, e.g. RFM, STM and SBF, the proposed approach still has space for improvement. The area, shape and location of RFM, STM and SBF have very high variation in different pathological cases, so the situation causes the registration of these muscle sub-regions inaccurate in atlas and hence reduces the robustness of fusion.

Visualization results of the three segmentations are also demonstrated in Fig. 4.9. In

Table 4.3: Quantitative comparisons of VMM, SAM and GRM segmentation.

Method	VMM		SAM		GRM	
	DSC	*RE*	*DSC*	*RE*	*DSC*	*RE*
MV	0.8063	0.3937	0.7068	0.5229	0.7878	0.4203
MA-EASA	0.8093	0.4054	0.5395	0.6160	0.5623	0.6716
Ours	0.8657	0.2773	0.8807	0.2438	0.9011	0.2083

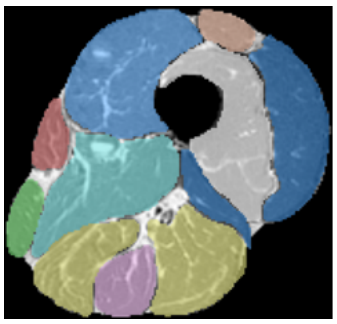
Table 4.4: Quantitative comparisons of AMM, SMM and STM segmentation.

Method	AMM		SMM		STM	
	DSC	*RE*	*DSC*	*RE*	*DSC*	*RE*
MV	0.8265	0.3879	0.7480	0.5606	0.5770	0.8749
MA-EASA	0.8020	0.4949	0.6798	0.6931	0.5383	0.8928
Ours	0.8876	0.2396	0.8104	0.4602	0.7883	0.5305

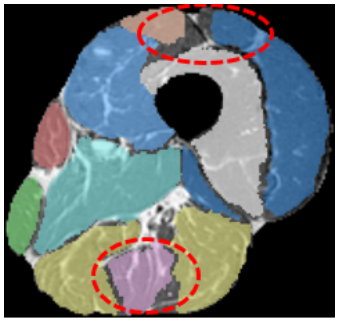
Fig. 4.9 (b), the MV for this case can correctly capture most of the muscle sub-regions. However in the red-dashed circles, this fusion strategy has obvious segmentation error that wrongly labels RF muscle as VL muscle, or it can not label the entire area of a muscle. On one point of the target image, the MV directly selects the major label from the warped manual labels. If the on-line training set does not contain atlases having high similarity to the target, registration error might cause unlabeled area or labeling errors. In Fig. 4.9 (c), the MA-EASA fills the unlabeled muscles by involving a label-entropy-based relaxation map and muscular spatial priors. But the red-dashed circles of this figure show that the MA-EASA goes through the under-segmentation issue of IAMAT. Furthermore, the MA-EASA's segmented boundaries are serrated, for the high variance of muscle shape makes the relaxation map coarse. In Fig. 4.9 (d), by minimizing the atlas-dependent expected error, the joint label fusion handles most of the issues of the previous methods. Although this approach still has over-segmentation

Table 4.5: Quantitative comparisons of LBF, SBF and TotM segmentation.

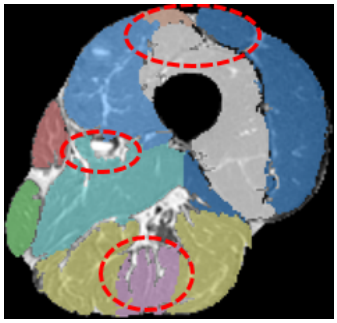
Method	LBF		SBF		TotM	
	DSC	*RE*	*DSC*	*RE*	*DSC*	*RE*
MV	0.8097	0.3797	0.5339	0.8363	0.8871	0.2199
MA-EASA	0.8362	0.3273	0.4762	0.8112	0.9201	0.1551
Ours	0.9053	0.1925	0.7395	0.6105	**0.9566**	**0.0881**



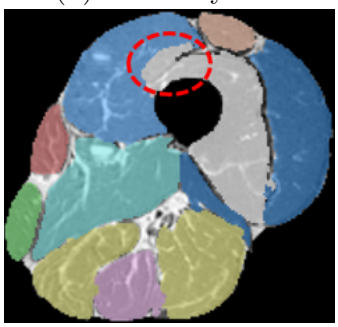
(a) Ground truth of 11 muscles



(b) Labels by MV



(c) Labels by MA-EASA



(d) Labels by the proposed

Figure 4.9: 2D visual results on individual muscles segmentation. (a) A slice from a ground truth volume. (b), (c), (d) Segmented individual muscles by the MV, the MA-EASA and the proposed approach, respectively.

Table 4.6: p-values of tissue volume of G_1 and G_2.

Group	smVol	matVol	imaVol	iaVol
G_1	0.1397	0.2240	**0.0084**	**0.0014**
G_2	0.8291	0.7663	0.2163	0.0855

problem, the application described in this paper only requires the TotM estimation, and hence the joint label fusion can work well for the thigh tissue quantification.

Table 4.7: p-values of tissue volume normalized by BMI of G_1 and G_2.

Group	smVol/BMI	matVol/BMI	imaVol/BMI	iaVol/BMI
G_1	0.2089	0.2550	**0.0047**	**0.0020**
G_2	0.8709	0.7010	0.2120	0.1158

Table 4.8: p-values of tissue volume normalized by smVol of G_1 and G_2.

Group	matVol/smVol	imaVol/smVol	iaVol/smVol
G_1	0.4931	0.0346	**0.0012**
G_2	0.8731	0.3084	0.1180

Table 4.9: p-values of tissue volume normalized by totVol of G_1 and G_2.

Group	smVol/totVol	matVol/totVol	imaVol/totVol	iaVol/totVol
G_1	0.9364	0.7381	**0.0095**	**0.00006**
G_2	0.4431	0.3689	0.6043	0.0159

4.3.3 Radiographic OA Related Analysis of Thigh Tissue Composition

With respect to knee OA, the relevance of muscle and adipose tissue parameters such as absolute or relative volume are still unclear. In this subsection, we investigate radiographic OA (ROA) related changes of muscle and adipose tissue volume in two subcohorts (G_1 and G_2) from the OAI study. The OAI database is queried for subjects with bilateral KLG (Kellgren and Lawrence system) scores at baseline (BL) and 48 months (M48). Subjects with bilateral KLG scores 0 or 1 are labeled as nROA (non-radiographic OA), while those with KLG scores 2, 3 or 4 are labeled as ROA. G_1 has 85 patients (29 male, 56 female, age: 45-76), and G_2 includes 61 patients (29 male, 32 female, age: 45-78). All subjects in the two sets are nROA at baseline. At M48, G_1 subjects are diagnosed with ROA, while G_2 subjects stay nROA. Thus, entire 292 MR subjects are processed in this test.

All results are visually inspected for segmentation errors, for we do not yet have manual ground truth for these two subcohorts. Volume of skeletal muscle (smVol), muscular adipose tissue (matVol), and separately of IMAT (imaVol) and IAMAT (iaVol) are assessed. The total leg volume (totVol) is the sum of SAT, skeletal muscle, MAT, and femur. So far paired t-tests are used for each variable to investigate whether changes over 4 years in volume or in volume normalized by subject's BMI (body mass index),

smVol or totVol are significant.

Table 4.6 to 4.9 show the p-values of different variables in G_1 and G_2, which are conducted by two-tailed hypothesis. In this medical test, the level of significance is set as 0.01, so two paired samples have no evident difference when p-value$>$ 0.01. The p-values of G_2 over 4 years are all larger than 0.01. Since subjects in G_2 at BL and M48 are without radiographic OA, age progress may have no effort to the volume of these 4 tissue components or their normalized values. On the other hand, in G_1 over 4 years, significant changes exist in the volume of IMAT and IAMAT, for the p-values are less than 0.01. Even after normalizing to BMI, smVol and totVol, changes remain significant (except for IMAT by smVol). Based on the experimental results, patients with knee OA may have obvious changes of the adipose tissue factors, e.g. volume of IMAT and IAMAT. These factors may be also helpful in studies of relevance between thigh tissue composition and knee ROA.

4.4 Summary

The main advantage of this work is that the framework comprehensively discriminates thigh tissue in a supervised way. Comparing with the snake-based muscle fascia segmentation in Orgiu et al. [89], the random-forest-based detection can handle more complex cases for large-scale analysis, and the follow-up fascia reconstruction using sparsity constraint can make up for the deficiencies from detection errors. Thus, the proposed framework can provide a more robust distinguishing way of the IMAT and SAT. The second is how to explicitly discriminate IMAT and IAMAT. Due to the highly variational appearance cues of thigh muscles, existing unsupervised strategies have to require manual intervention, which is an serious shortcoming when processing

large data set. Yet the proposed joint-label-fused and morphological-refined muscles extraction can work well under such situation. Firstly, the atlas-based segmentation does not require large on-line training data. In addition, the loss function in the joint label fusion considers mutual error of each atlas, which can effectively correct labeling mistakes. Lastly, the formulation to determine the fusing weights is a closed form, and hence it can be very fast solved. Thus, all of these merits make the work in this paper as a robust and scalable thigh tissue quantification.

Quantitative assessment of thigh tissue composition is essential in medical and clinical analyses. In this chapter, we present an effective integrated framework to tackle this application. The key contribution of this work is to exploit supervised tissue discrimination techniques utilizing small amount of supervised information in learning to automatically extract different thigh tissue: femur, skeletal muscle, SAT, IAMAT and IMAT. Moreover, since only few manual annotations are required to label thigh fascia and muscles, our method becomes scalable to thigh tissue analysis with different input data format. We validate the novelties of the proposed methods, fascia lata reconstruction and individual muscles segmentation, with state-of-the-art thigh tissue classification and atlas-based segmentation, respectively. The quantitative comparisons show that the accuracy and effectiveness of the proposed framework. Importantly, another major contribution is to apply the framework to conduct extensive evaluations on a data set including 292 3D MR thigh subjects in temporal and pathological progression. The experimental results confirm that the volume of IMAT and IAMAT, and their normalized values by BMI, smVol and totVol are changeable factors under knee OA conditions. Thus, the proposed framework on large MR thigh data set could be helpful in studies of relevance between thigh tissue composition and knee ROA.

Chapter 5

Conclusions

In this dissertation, we focus on three challenging but essential tissue extraction problems. We propose three machine learning based segmentation of knee cartilage, femoral head-neck junction and thigh muscular/adipose tissue in MR data. Specifically, the new segmentation paradigms include collaborative multi-agent learning, multi-task learning with boundary distance regression, and integrated scheme with data-driven and sparsity-constrained deformable segmentation. Furthermore, we discuss the feasibility of large-scale knee/hip OA analysis using the proposed segmentation methods. The contributions are summarized as follows.

1. First, we develop a novel framework to effectively segment knee cartilages in fine details from high-resolution and large-sized 3D MR data. The key contribution is an end-to-end framework which utilizes adversarial learning to enable a joint learning of shape and spatial relations between different cartilages from segmentation agents. The 3D morphology and quantitative assessment of knee cartilages is of great importance for knee radiographic OA diagnostic decision making. The proposed scheme is effective for cartilage biomarkers (e.g., surface area and volume) estimation in large-scale quantitative tests. We have applied the scheme on a 36-month subset (1000 knees) of OAI for OA-MOAKS analysis.

2. Second, we propose a deep multi-task network for robust shape preserved femoral

head-neck segmentation. The novelty of this work is a unified end-to-end network to combine classification and regression tasks into one "Y" shape net for accurate bone segmentation. The auxiliary branch with the novel boundary distance regression can help the segmented outcomes to alleviate spatially isolated segmentation errors as well as ensure the smoothness of segmented contours. At last, the designed Y-shape network structure, which bifurcates at the end of the encoding path, could balance the training between the shared encoding and non-shared decoding paths. The proposed approach not only obtains the highest performance in the testings, but only has been used for measurements of the femoral head-neck morphology in the extensive experiment for hip OA analysis.

3. Besides the above deep learning based segmentation, we also explore the correlation analysis between knee OA and thigh tissue composition by an integrated framework for MR thigh tissue quantifying. The designed framework is able to distinguish muscular tissue and different types of adipose tissues, i.e., SAT, IMAT and IAMAT, efficiently. Deformable models and learning based techniques are integrated in the framework to enable robust quantification. Importantly, extensive evaluations are conducted on a large set of 3D MR thigh volumes from longitudinal studies of hundreds of subjects to investigate knee OA related changes of muscular and adipose tissue volumes. The analysis is constructed by two subcohorts (G_1 and G_2). G_2 has 61 patients which keep healthy at baseline (BL) and 48 months (M48), while G_1's 85 patients are healthy at BL but have knee OA at M48. Paired t-tests are used to investigate the changes of these tissue size over time passing with/without pathological progression. The experimental results show that, in G_1, patients' IMAT and IAMAT are statistically significant

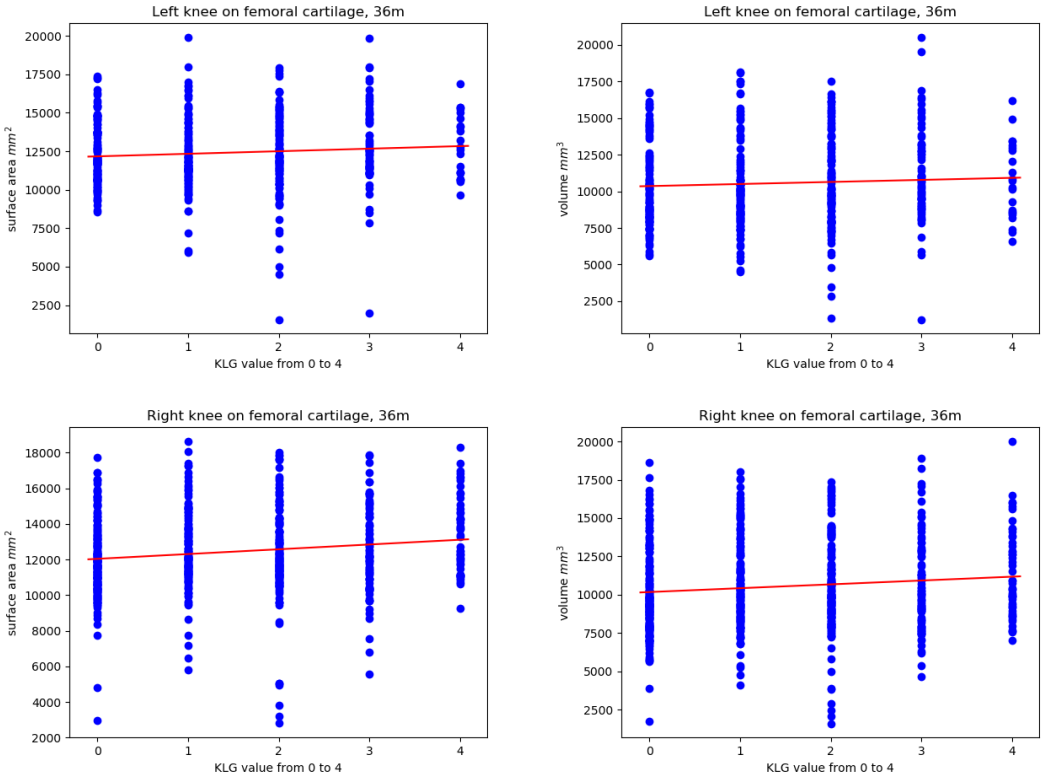


Figure 5.1: Correlation between KLG scores and mean volume/surface area of femoral cartilage on left/right leg. In (a) to (d), the x-axis represents the KLG scores, and y-axis denotes the value of surface area and volume.

respectively, yet G_2 has no such variation in the same tissue type. Thus we conclude from the statistical analysis that age may not directly affect thigh tissues, but IMAT and IAMAT may have obvious changes in patients with knee OA.

This thesis work proposes effective knee cartilage, femoral head-neck junction and thigh muscular/adipose tissue segmentation paradigms, and attempt to apply the segmented results for quantitative computation of OA-related tissual analysis on large-scale dataset. The encouraged potential future work includes the following several directions.

1. On the one hand, besides the OA-MOAKS analysis of patellar cartilage on the 36-month OAI subset, we also do an OA-KLG analysis of femoral cartilage using the same biomarkers and subset. The correlation between the KLG scores and mean

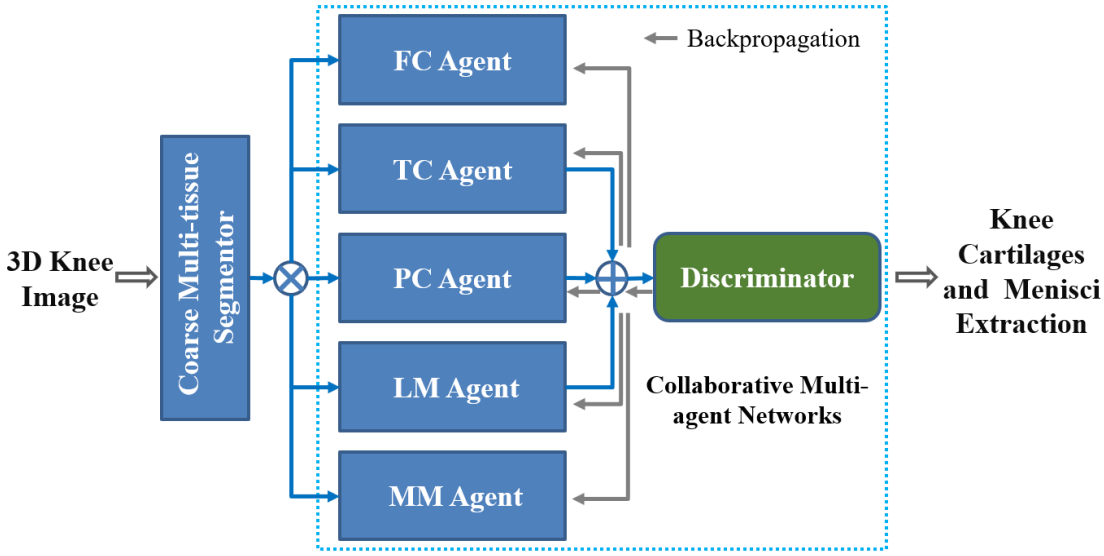


Figure 5.2: Flowchart of the extended multi-agent segmentation with meniscal agents.

volume/surface area of femoral cartilage is displayed in Fig. 5.1. However, as the plots in this figure shown, the biomarkers do not reveal significant correlation. Using the whole femoral cartilage structure may be the main reason causing the correlation searching failed. Only a portion of the femoral cartilage tissue is in contact with the menisci and tibial cartilage, and the most relevant area to the KLG scores may exist in the place where two bones touch/move with each other. Thus, the cartilage segmentation framework should be extended and other indicative knee tissue should be included into the task.

On the another hand, the menisci (i.e., lateral and medial meniscus, shown in Fig. 2.1) of the knee are two pieces of cartilaginous tissue found between the lower leg (tibia) and the thigh (femur) [125]. The menisci disperse friction, absorb shock produced by activities from the two bones meeting in the knee joint space. The meniscus is another important type of tissue in the knee joint for the estimation of knee OA progression. For instance, Paproki et al. [92] developed a scheme using 3D active shape model (ASM) to fit the shape of menisci for computing the

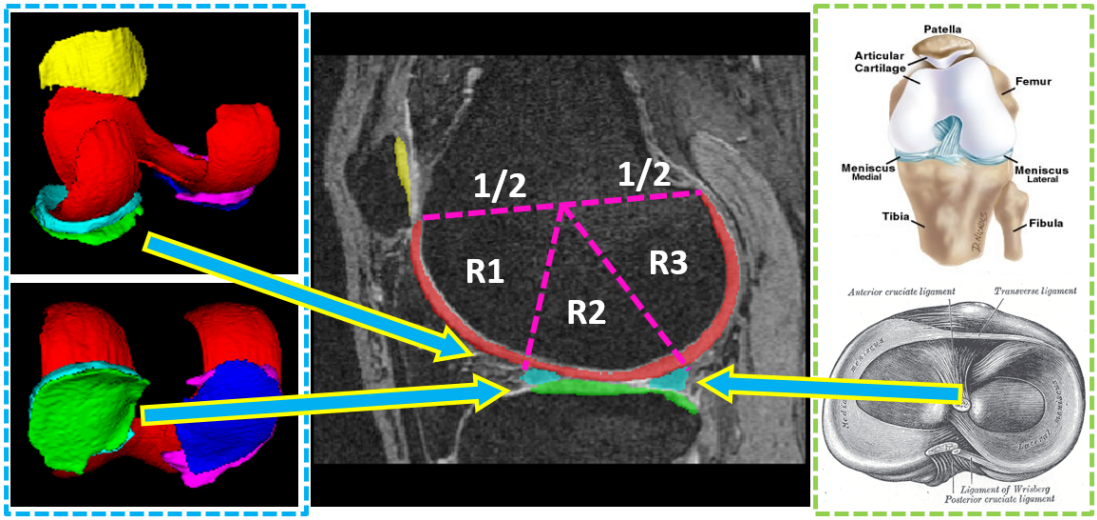


Figure 5.3: Anatomical delineation of FC mask into three subregions (R1-3) located by the boundary of LM mask in the light blue dashed box. The anatomical structure of menisci is shown in the green dashed box [54, 125].

OA-related quantitative parameters including volume, sub-luxation and tibial-coverage, and Kawahara et al. [57] estimated the OA progression from the KLG scale by manually measuring several meniscal parameters (e.g., anterior/posterior wedge thickness, anterior wedge angle). Thus, we could jointly segment the knee cartilages and menisci under the collaborative multi-agent processing. Because the menisci exist in the gap between the FC and TC, adding the meniscal agents (i.e., LM and MM) into the proposed adversarial learning could enhance the distinguishment of targets' boundaries. The extended segmentation is illustrated in Fig. 5.2. Next, the segments of menisci could help build the subregion-based OA progression assessment [53]. As Fig. 5.3 shows, with the masks of LM and FC, an automated subregion delineation could be set up on the FC to detect the highest correlated subregion to the OA conditions.

2. Besides the femoral head-neck bone segmentation on 2D MR data, its 3D version on CT data [66, 117] is a potential direction of development. 3D segmentation

with deep learning modules have been comprehensively developed. From the basic 3D DCNN based end-to-end baseline [83], Dou et al. developed the 3D deeply supervised network [31]. Then Yang et al. integrated the GAN's discriminating module into the 3D semantic end-to-end segmentation [138]. Recently, an end-to-end adversarial shape learning [18] also shows improved capture ability of organ shape. Yet, the performance on 3D DCNN based segmentation is significantly affected by number of training data. It is extremely challenging to build a sufficiently large dataset of 3D CT femoral head-neck bone due to difficulty of data acquisition and annotation in 3D data. Inspired some transfer learning work, like [26], start by training a baseline 3D network with our current MR femur scans, the trained model could be transferred to solve the CT femoral head-neck junction problem.

3. With the help of the integrated extraction framework of thigh tissue, we could not only carry out the correlation analysis between knee OA and thigh tissue composition, but also effectively prepare a larger set of segmented data for training. Start from this labeled set, some manual quality control and annotation refinement would be involved to obtain a reliable training set. In this way, the difficulty of data annotation starting from scratch could be avoided or obviously declined. However, on the other hand, the modules implemented by traditional supervised learning in the thigh tissue analysis still have some issues, and could be replaced and improved by deep learning based approaches. (I) From our observations, in the fascia lata reconstruction, the main segmentation errors are due to the inaccurate control points from the fascia lata detection. The reasons

causing these fallacious detections come from the blurred boundaries in some low-quality areas or the heavy bias noise existing in the testing MR images. Hence, in future, to improve the fascia lata reconstruction, we could combine the interior area segmentation encircled the fascia lata and the contour prediction using deep shape priors [2, 133]. (II) For the muscles segmentation, although Wang et al. [121] utilized the unsupervised search to optimize the pairwise Atlas dependency matrix M_x, the joint label fusion might still not precisely locate the muscular boundaries, when these border areas have indistinct neighboring image patches. In recent years, depending on the techniques from other fields, for example, the nuclei and glands segmented in the histopathology image analysis have the similar morphological appearances to the muscular tissue. Thus, the instance segmentation with attention mechanism in the histopathology problems [142, 102] could enhance the robustness and accuracy on the thigh muscle segmentation, especially on the border areas.

References

[1] AAOS: Femoroacetabular impingement. https://orthoinfo.aaos.org/en/diseases--conditions/femoroacetabular-impingement/ (2016)

[2] Al Arif, S.M.R., Knapp, K., Slabaugh, G.: Spnet: Shape prediction using a fully convolutional neural network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 430–439. Springer (2018)

[3] Amanatullah, D.F., Antkowiak, T., Pillay, K., Patel, J., Refaat, M., Toupadakis, C.A., Jamali, A.A.: Femoroacetabular impingement: current concepts in diagnosis and treatment. Orthopedics **38**(3), 185–199 (2015)

[4] Ambellan, F., Tack, A., Ehlke, M., Zachow, S.: Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the osteoarthritis initiative. Medical Image Analysis (2018)

[5] Anderson, L.A., Peters, C.L., Park, B.B., Stoddard, G.J., Erickson, J.A., Crim, J.R.: Acetabular cartilage delamination in femoroacetabular impingement: risk factors and magnetic resonance imaging diagnosis. JBJS **91**(2), 305–313 (2009)

[6] Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence **33**(5), 898–916 (2010)

[7] Banerjee, P., Mclean, C.R.: Femoroacetabular impingement: a review of diagnosis and management. Current reviews in musculoskeletal medicine **4**(1), 23 (2011)

[8] Barrientos, C., Barahona, M., Diaz, J., Branes, J., Chaparro, F., Hinzpeter, J.: Is there a pathological alpha angle for hip impingement? a diagnostic test study. Journal of hip preservation surgery **3**(3), 223–228 (2016)

[9] Beattie, K.A., MacIntyre, N.J., Ramadan, K., Inglis, D., Maly, M.R.: Longitudinal changes in intermuscular fat volume and quadriceps muscle volume in the thighs of women with knee osteoarthritis. Arthritis care & research **64**(1), 22–29 (2012)

[10] Beaulé, P.E., Lee, J.L., Le Duff, M.J., Amstutz, H.C., Ebramzadeh, E.: Orientation of the femoral component in surface arthroplasty of the hip: a biomechanical and clinical analysis. JBJS **86**(9), 2015–2021 (2004)

[11] Beck, M., Kalhor, M., Leunig, M., Ganz, R.: Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. The Journal of bone and joint surgery. British volume **87**(7), 1012–1018 (2005)

[12] Beck, M., Leunig, M., Parvizi, J., Boutier, V., Wyss, D., Ganz, R.: Anterior femoroacetabular impingement: part ii. midterm results of surgical treatment. Clinical Orthopaedics and Related Research® **418**, 67–73 (2004)

[13] Bezdek, J.C., Ehrlich, R., Full, W.: Fcm: The fuzzy c-means clustering algorithm. Computers & Geosciences **10**(2), 191–203 (1984)

[14] Borges, P.D.N., Vincent, T.L., Marenzana, M.: Automated assessment of bone changes in cross-sectional micro-ct studies of murine experimental osteoarthritis. PloS one **12**(3), e0174294 (2017)

[15] Botser, I.B., Ozoude, G.C., Martin, D.E., Siddiqi, A.J., Kuppuswami, S., Domb, B.G.: Femoral anteversion in the hip: comparison of measurement by computed tomography, magnetic resonance imaging, and physical examination. Arthroscopy: The Journal of Arthroscopic & Related Surgery **28**(5), 619–627 (2012)

[16] Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)

[17] Bricca, A., Juhl, C.B., Steultjens, M., Wirth, W., Roos, E.M.: Impact of exercise on articular cartilage in people at risk of, or with established, knee osteoarthritis: a systematic review of randomised controlled trials. Br J Sports Med pp. bjsports–2017 (2018)

[18] Cai, J., Xia, Y., Yang, D., Xu, D., Yang, L., Roth, H.: End-to-end adversarial shape learning for abdomen organ deep segmentation. In: International Workshop on Machine Learning in Medical Imaging. pp. 124–132. Springer (2019)

[19] Cao, Q., Thawait, G., Gang, G.J., Zbijewski, W., Reigel, T., Brown, T., Corner, B., Demehri, S., Siewerdsen, J.H.: Characterization of 3d joint space morphology using an electrostatic model (with application to osteoarthritis). Physics in medicine and biology **60**(3), 947 (2015)

[20] Carballido-Gamio, J., Bauer, J.S., Stahl, R., Lee, K.Y., Krause, S., Link, T.M., Majumdar, S.: Inter-subject comparison of mri knee cartilage thickness. Medical image analysis **12**(2), 120–135 (2008)

[21] CDC: Osteoarthritis (oa). `https://www.cdc.gov/arthritis/basics/osteoarthritis.htm` (2019)

[22] Chan, E.F., Farnsworth, C.L., Koziol, J.A., Hosalkar, H.S., Sah, R.L.: Statistical shape modeling of proximal femoral shape deformities in legg–calvé–perthes disease and slipped capital femoral epiphysis. Osteoarthritis and cartilage **21**(3), 443–449 (2013)

[23] Chan, T.F., Vese, L.A.: Active contours without edges. Image Processing, IEEE Transactions on **10**(2), 266–277 (2001)

[24] Chen, H., Qi, X., Yu, L., Heng, P.A.: Dcan: Deep contour-aware networks for accurate gland segmentation. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 2487–2496 (2016)

[25] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv:1412.7062 (2014)

[26] Chen, S., Ma, K., Zheng, Y.: Med3d: Transfer learning for 3d medical image analysis. arXiv preprint arXiv:1904.00625 (2019)

[27] Cheng, J.Z., Ni, D., Chou, Y.H., Qin, J., Tiu, C.M., Chang, Y.C., Huang, C.S., Shen, D., Chen, C.M.: Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in ct scans. Scientific reports **6**, 24454 (2016)

[28] Dai, W., Dong, N., Wang, Z., Liang, X., Zhang, H., Xing, E.P.: Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 263–273. Springer (2018)

[29] Dannhauer, T., Ruhdorfer, A., Wirth, W., Eckstein, F.: Quantitative relationship of thigh adipose tissue with pain, radiographic status, and progression of knee osteoarthritis. Investigative Radiology (2014)

[30] Doi, K.: Current status and future potential of computer-aided diagnosis in medical imaging. The British journal of radiology **78**(suppl_1), s3–s19 (2005)

[31] Dou, Q., Yu, L., Chen, H., Jin, Y., Yang, X., Qin, J., Heng, P.A.: 3d deeply supervised network for automated segmentation of volumetric medical images. Medical image analysis **41**, 40–54 (2017)

[32] Dudda, M., Albers, C., Mamisch, T.C., Werlen, S., Beck, M.: Do normal radiographs exclude asphericity of the femoral head-neck junction? Clinical orthopaedics and related research **467**(3), 651–659 (2009)

[33] Durheim, M.T., Slentz, C.A., Bateman, L.A., Mabe, S.K., Kraus, W.E.: Relationships between exercise-induced reductions in thigh intermuscular adipose tissue, changes in lipoprotein particle size, and visceral adiposity. American Journal of Physiology-Endocrinology and Metabolism **295**(2), E407–E412 (2008)

[34] Eckstein, F., Maschek, S., Wirth, W., Hudelmaier, M., Hitzl, W., Wyman, B., Nevitt, M., Le Graverand, M.H., Group, O.I., et al.: One year change of knee cartilage morphology in the first release of participants from the osteoarthritis initiative progression subcohort: association with sex, body mass index, symptoms and radiographic osteoarthritis status. Annals of the rheumatic diseases **68**(5), 674–679 (2009)

[35] Eckstein, F., Wirth, W.: Quantitative cartilage imaging in knee osteoarthritis. Arthritis **2011** (2010)

[36] Foreman, S.C., Zhang, A.L., Neumann, J., von Schacky, C.E., Souza, R.B., Majumdar, S., Link, T.M.: Postoperative mri findings and associated pain changes after arthroscopic surgery for femoroacetabular impingement. American Journal of Roentgenology pp. 1–8 (2019)

[37] Fu, G.S., Levin-Schwartz, Y., Lin, Q.H., Zhang, D.: Machine learning for medical imaging. Journal of healthcare engineering **2019** (2019)

[38] Ganz, R., Parvizi, J., Beck, M., Leunig, M., Nötzli, H., Siebenrock, K.A.: Femoroacetabular impingement: a cause for osteoarthritis of the hip. Clinical Orthopaedics and Related Research® **417**, 112–120 (2003)

[39] Gao, Y., Shao, Y., Lian, J., Wang, A.Z., Chen, R.C., Shen, D.: Accurate segmentation of ct male pelvic organs via regression-based deformable models and multi-task random forests. IEEE transactions on medical imaging **35**(6), 1532–1543 (2016)

[40] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J.: A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857 (2017)

[41] Ghosh, A., Kulharia, V., Namboodiri, V.P., Torr, P.H., Dokania, P.K.: Multi-agent diverse generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8513–8521 (2018)

[42] Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 1440–1448 (2015)

[43] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in neural information processing systems. pp. 2672–2680 (2014)

[44] Graichen, H., Eisenhart-Rothe, R.v., Vogl, T., Englmeier, K.H., Eckstein, F.: Quantitative assessment of cartilage status in osteoarthritis by quantitative magnetic resonance imaging: technical validation for use in analysis of cartilage volume and further morphologic parameters. Arthritis & Rheumatism **50**(3), 811–816 (2004)

[45] Haider, I., Speirs, A., Alnabelseya, A., Beaulé, P., Frei, H.: Femoral subchondral bone properties of patients with cam-type femoroacetabular impingement. Osteoarthritis and cartilage **24**(6), 1000–1006 (2016)

[46] He, K., Cao, X., Shi, Y., Nie, D., Gao, Y., Shen, D.: Pelvic organ segmentation using distinctive curve guided fully convolutional networks. IEEE transactions on medical imaging **38**(2), 585–595 (2019)

[47] He, X., Tan, C., Qiao, Y., Tan, V., Metaxas, D., Li, K.: Effective 3d humerus and scapula extraction using low-contrast and high-shape-variability mr data. In: Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging. vol. 10953, p. 109530O. International Society for Optics and Photonics (2019)

[48] Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A.: Automatic anatomical brain mri segmentation combining label propagation and decision fusion. NeuroImage **33**(1), 115–126 (2006)

[49] Heijboer, M.P., Bierma-Zeinstra, S.M., Verhaar, J.A., Weinans, H., Waarsing, J.H., et al.: Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (check). Annals of the rheumatic diseases **72**(6), 918–923 (2013)

[50] Heimann, T., Meinzer, H.P.: Statistical shape models for 3d medical image segmentation: a review. Medical image analysis **13**(4), 543–563 (2009)

[51] Hong, Y., Hwang, U., Yoo, J., Yoon, S.: How generative adversarial networks and their variants work: An overview. ACM Computing Surveys (CSUR) **52**(1), 10 (2019)

[52] Huang, X., Metaxas, D.N.: Metamorphs: deformable shape and appearance models. Pattern Analysis and Machine Intelligence, IEEE Transactions on **30**(8), 1444–1459 (2008)

[53] Hunter, D.J., Guermazi, A., Lo, G.H., Grainger, A.J., Conaghan, P.G., Boudreau, R.M., Roemer, F.W.: Evolution of semi-quantitative whole joint assessment of knee oa: Moaks (mri osteoarthritis knee score). Osteoarthritis and Cartilage **19**(8), 990–1002 (2011)

[54] Institute, P.: Anatomy of the knee joint. `https://paleyinstitute.org/centers-of-excellence/cartilage-repair/anatomy-of-the-knee-joint/` (2018)

[55] Jetley, S., Lord, N.A., Lee, N., Torr, P.H.: Learn to pay attention. arXiv preprint arXiv:1804.02391 (2018)

[56] Jiri Matas, J.S.: Adaboost. Center for Machine Perception, Czech Technical University (2007)

[57] Kawahara, T., Sasho, T., Ohnishi, T., Haneishi, H.: Stage-specific meniscal features predict progression of osteoarthritis of the knee: a retrospective cohort study using data from the osteoarthritis initiative. BMC musculoskeletal disorders **20**(1), 33 (2019)

[58] Kemnitz, J., Wirth, W., Eckstein, F., Culvenor, A.: Thigh muscle and adipose tissue changes during symptomatic and radiographic knee osteoarthritis progression–data from the osteoarthritis initiative. Osteoarthritis and Cartilage **26**, S406 (2018)

[59] Kemp, J., Waarsing, J., Weinans, H., Verhaar, J., Runhaar, J., Bierma-Zeinstra, S., et al.: Femoroacetabular impingement syndrome is associated with development of hip osteoarthritis within 10-years follow-up: Data from the check cohort. Osteoarthritis and Cartilage **27**, S57 (2019)

[60] Kokkinos, I.: Ubernet: Training auniversal'convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory. arXiv preprint arXiv:1609.02132 (2016)

[61] Kornaat, P.R., Bloem, J.L., Ceulemans, R.Y., Riyazi, N., Rosendaal, F.R., Nelissen, R.G., Carter, W.O., Hellio Le Graverand, M.P., Kloppenburg, M.: Osteoarthritis of the knee: association between clinical features and mr imaging findings. Radiology **239**(3), 811–817 (2006)

[62] Kumar, D., Karampinos, D., Macleod, T., Lin, W., Nardo, L., Li, X., Link, T., Majumdar, S., Souza, R.: Quadriceps intramuscular fat fraction rather than muscle size is associated with knee osteoarthritis. Osteoarthritis and Cartilage **22**(2), 226–234 (2014)

[63] Larson, C.M., Stone, R.M., Birmingham, P.M.: Arthroscopic treatment of combined cam-and pincer-type femoroacetabular impingement. In: Hip Joint Restoration, pp. 489–501. Springer (2017)

[64] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. nature **521**(7553), 436–444 (2015)

[65] Lee, S., Nardo, L., Kumar, D., Wyatt, C.R., Souza, R.B., Lynch, J., McCulloch, C.E., Majumdar, S., Lane, N.E., Link, T.M.: Scoring hip osteoarthritis with mri (shomri): a whole joint osteoarthritis evaluation system. Journal of Magnetic Resonance Imaging **41**(6), 1549–1557 (2015)

[66] Lerch, T.D., Degonda, C., Schmaranzer, F., Todorski, I., Cullmann-Bastian, J., Zheng, G., Siebenrock, K.A., Tannast, M.: Patient-specific 3-d magnetic resonance imaging–based dynamic simulation of hip impingement and range of motion can replace 3-d computed tomography–based simulation for patients with femoroacetabular impingement: Implications for planning open hip preservation surgery and hip arthroscopy. The American journal of sports medicine **47**(12), 2966–2977 (2019)

[67] Lewek, M.D., Rudolph, K.S., Snyder-Mackler, L.: Quadriceps femoris muscle weakness and activation failure in patients with symptomatic knee osteoarthritis. Journal of Orthopaedic Research **22**(1), 110–115 (2004)

[68] Li, X., Hu, X., Yu, L., Zhu, L., Fu, C.W., Heng, P.A.: Canet: Cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. IEEE transactions on medical imaging (2019)

[69] Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. In: Image Processing. 2002. Proceedings. 2002 International Conference on. vol. 1, pp. I–900. IEEE (2002)

[70] Liu, F., Zhou, Z., Jang, H., Samsonov, A., Zhao, G., Kijowski, R.: Deep convolutional neural network and 3d deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magnetic resonance in medicine (2017)

[71] Loeser, R.F., Goldring, S.R., Scanzello, C.R., Goldring, M.B.: Osteoarthritis: a disease of the joint as an organ. Arthritis & Rheumatism **64**(6), 1697–1707 (2012)

[72] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015)

[73] Lu, L., Zheng, Y., Carneiro, G., Yang, L.: Deep learning and convolutional neural networks for medical image computing. Advances in Computer Vision and Pattern Recognition; Springer: New York, NY, USA (2017)

[74] Luc, P., Couprie, C., Chintala, S., Verbeek, J.: Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408 (2016)

[75] Makrogiannis, S., Serai, S., Fishbein, K.W., Schreiber, C., Ferrucci, L., Spencer, R.G.: Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed mr images. Journal of Magnetic Resonance Imaging **35**(5), 1152–1161 (2012)

[76] Maly, M.R., Calder, K.M., MacIntyre, N.J., Beattie, K.A.: Relationship of intermuscular fat volume in the thigh with knee extensor strength and physical performance in women at risk of or with knee osteoarthritis. Arthritis care & research **65**(1), 44–52 (2013)

[77] Maly, M., MacIntyre, N., Beattie, K.: Intermuscular fat volume in the thigh relates to knee strength and physical performance among women at risk for or with knee osteoarthritis. Osteoarthritis and Cartilage **19**, S181–S181 (2011)

[78] Marcus, R.L., Addison, O., LaStayo, P.C.: Intramuscular adipose tissue attenuates gains in muscle quality in older adults at high risk for falling. a brief report. The journal of nutrition, health & aging **17**(3), 215–218 (2013)

[79] mayoclinic: Osteoarthritis overview. `https://www.mayoclinic.org/diseases-conditions/osteoarthritis/symptoms-causes/syc-20351925` (2016)

[80] Mendy, A., Park, J., Vieira, E.R.: Osteoarthritis and risk of mortality in the usa: a population-based cohort study. International journal of epidemiology **47**(6), 1821–1829 (2018)

[81] Messier, S.P., Beavers, D.P., Loeser, R.F., Carr, J.J., Khajanchi, S., Legault, C., Nicklas, B.J., Hunter, D.J., DeVita, P.: Knee-joint loading in knee osteoarthritis: influence of abdominal and thigh fat. Medicine and science in sports and exercise **46**(9), 1677 (2014)

[82] Meybodi, M.K.E., Rasouli, H.R., Ghafouri, M., Bafghi, S.A., Ghafouri, M.: The correlation between hip morphology and hip osteoarthritis: A cross-sectional study. Iranian Red Crescent Medical Journal **20**(6) (2018)

[83] Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 3D Vision (3DV), 2016 Fourth International Conference on. pp. 565–571. IEEE (2016)

[84] Nemtala, F., Mardones, R.M., Tomic, A.: Anterior and posterior femoral head-neck offset ratio in the cam impingement. Cartilage **1**(3), 238–241 (2010)

[85] Nepple, J.J., Brophy, R.H., Matava, M.J., Wright, R.W., Clohisy, J.C.: Radiographic findings of femoroacetabular impingement in national football league combine athletes undergoing radiographs for previous hip or groin pain. Arthroscopy: The Journal of Arthroscopic & Related Surgery **28**(10), 1396–1403 (2012)

[86] NIH: Fact sheet of osteoarthritis. National Institutes of Health (2010)

[87] Nötzli, H., Wyss, T., Stoecklin, C., Schmid, M., Treiber, K., Hodler, J.: The contour of the femoral head-neck junction as a predictor for the risk of anterior impingement. The Journal of bone and joint surgery. British volume **84**(4), 556–560 (2002)

[88] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)

[89] Orgiu, S., Lafortuna, C.L., Rastelli, F., Cadioli, M., Falini, A., Rizzo, G.: Automatic muscle and fat segmentation in the thigh from t1-weighted mri. Journal of Magnetic Resonance Imaging (2015)

[90] Pal, N.R., Pal, S.K.: A review on image segmentation techniques. Pattern recognition **26**(9), 1277–1294 (1993)

[91] Panzer, S., Augat, P., Esch, U.: Ct assessment of herniation pits: prevalence, characteristics, and potential association with morphological predictors of femoroacetabular impingement. European radiology **18**(9), 1869 (2008)

[92] Paproki, A., Engstrom, C., Chandra, S.S., Neubert, A., Fripp, J., Crozier, S.: Automated segmentation and analysis of normal and osteoarthritic knee menisci from magnetic resonance images–data from the osteoarthritis initiative. Osteoarthritis and cartilage **22**(9), 1259–1270 (2014)

[93] Parvizi, J., Leunig, M., Ganz, R.: Femoroacetabular impingement. JAAOS-Journal of the American Academy of Orthopaedic Surgeons **15**(9), 561–570 (2007)

[94] Pfirrmann, C.W., Mengiardi, B., Dora, C., Kalberer, F., Zanetti, M., Hodler, J.: Cam and pincer femoroacetabular impingement: characteristic mr arthrographic findings in 50 patients. Radiology **240**(3), 778–785 (2006)

[95] Pham, D.L., Xu, C., Prince, J.L.: Current methods in medical image segmentation. Annual review of biomedical engineering **2**(1), 315–337 (2000)

[96] Positano, V., Christiansen, T., Santarelli, M.F., Ringgaard, S., Landini, L., Gastaldelli, A.: Accurate segmentation of subcutaneous and intermuscular adipose tissue from mr images of the thigh. Journal of Magnetic Resonance Imaging **29**(3), 677–684 (2009)

[97] Potter, H.G., Koff, M.F.: Mr imaging tools to assess cartilage and joint structures. HSS journal **8**(1), 29–32 (2012)

[98] Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M.: Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: International conference on medical image computing and computer-assisted intervention. pp. 246–253. Springer (2013)

[99] Prescott, J.W., Best, T.M., Swanson, M.S., Haq, F., Jackson, R.D., Gurcan, M.N.: Anatomically anchored template-based level set segmentation: application

to quadriceps muscles in mr images from the osteoarthritis initiative. Journal of digital imaging **24**(1), 28–43 (2011)

[100] Pun, S., Kumar, D., Lane, N.E.: Femoroacetabular impingement. Arthritis & rheumatology **67**(1), 17–27 (2015)

[101] Qin, C., Yao, D., Shi, Y., Song, Z.: Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomedical engineering online **17**(1), 113 (2018)

[102] Qu, H., Yan, Z., Riedlinger, G.M., De, S., Metaxas, D.N.: Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 378–386. Springer (2019)

[103] Rakhra, K.S., Sheikh, A.M., Allen, D., Beaulé, P.E.: Comparison of mri alpha angle measurement planes in femoroacetabular impingement. Clinical orthopaedics and related research **467**(3), 660–665 (2009)

[104] Rhee, C., Le Francois, T., Byrd, J.T., Glazebrook, M., Wong, I.: Radiographic diagnosis of pincer-type femoroacetabular impingement: a systematic review. Orthopaedic journal of sports medicine **5**(5), 2325967117708307 (2017)

[105] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)

[106] Segal, N.A., Findlay, C., Wang, K., Torner, J.C., Nevitt, M.C.: The longitudinal relationship between thigh muscle mass and the development of knee osteoarthritis. Osteoarthritis and cartilage **20**(12), 1534–1540 (2012)

[107] Shan, L., Zach, C., Charles, C., Niethammer, M.: Automatic atlas-based three-label cartilage segmentation from mr knee images. Medical image analysis **18**(7), 1233–1246 (2014)

[108] Shi, J., Malik, J.: Normalized cuts and image segmentation. Departmental Papers (CIS) p. 107 (2000)

[109] Song, Y., Greve, J., Carter, D., Koo, S., Giori, N.: Articular cartilage mr imaging and thickness mapping of a loaded knee joint before and after meniscectomy. Osteoarthritis and cartilage **14**(8), 728–737 (2006)

[110] Tan, C., Li, K., Yan, Z., Yang, D., Zhang, S., Yu, H.J., Engelke, K., Miller, C., Metaxas, D.: A detection-driven and sparsity-constrained deformable model for fascia lata labeling and thigh inter-muscular adipose quantification. Computer Vision and Image Understanding **151**, 80–89 (2016)

[111] Tan, C., Li, K., Yan, Z., Yi, J., Wu, P., Yu, H.J., Engelke, K., Metaxas, D.N.: Towards large-scale mr thigh image analysis via an integrated quantification framework. Neurocomputing **229**, 63–76 (2017)

[112] Tan, C., Yan, Z., Zhang, S., Belaroussi, B., Yu, H.J., Miller, C., Metaxas, D.N.: An automated and robust framework for quantification of muscle and fat in the thigh. In: 2014 22nd International Conference on Pattern Recognition. pp. 3173–3178. IEEE (2014)

[113] Tan, C., Yan, Z., Zhang, S., Li, K., Metaxas, D.N.: Collaborative multi-agent learning for mr knee articular cartilage segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 282–290. Springer (2019)

[114] Tan, C., Zhao, L., Yan, Z., Li, K., Metaxas, D., Zhan, Y.: Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 1221–1224. IEEE (2018)

[115] Tannast, M., Siebenrock, K.A., Anderson, S.E.: Femoroacetabular impingement: radiographic diagnosis—what the radiologist should know. American Journal of Roentgenology **188**(6), 1540–1552 (2007)

[116] Tannenbaum, E.P., Zhang, P., Maratt, J.D., Gombera, M.M., Holcombe, S.A., Wang, S.C., Bedi, A., Goulet, J.A.: A computed tomography study of gender differences in acetabular version and morphology: implications for femoroacetabular impingement. Arthroscopy: The Journal of Arthroscopic & Related Surgery **31**(7), 1247–1254 (2015)

[117] Turmezei, T.D., Treece, G.M., Gee, A.H., Fotiadou, A.F., Poole, K.E.: Quantitative 3d analysis of bone in hip osteoarthritis using clinical computed tomography. European radiology **26**(7), 2047–2054 (2016)

[118] Uzunbaş, M.G., Chen, C., Zhang, S., Pohl, K.M., Li, K., Metaxas, D.: Collaborative multi organ segmentation by integrating deformable and graphical models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 157–164. Springer (2013)

[119] Valentinitsch, A., C Karampinos, D., Alizai, H., Subburaj, K., Kumar, D., M Link, T., Majumdar, S.: Automated unsupervised multi-parametric classification of adipose tissue depots in skeletal muscle. Journal of Magnetic Resonance Imaging **37**(4), 917–927 (2013)

[120] Wang, G., Zhang, S., Xie, H., Metaxas, D.N., Gu, L.: A homotopy-based sparse representation for fast and accurate shape prior modeling in liver surgical planning. Medical image analysis **19**(1), 176–186 (2015)

[121] Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on **35**(3), 611–623 (2013)

[122] Wikipedia: Computed tomography scan. `https://en.wikipedia.org/wiki/CT_scan` (2019)

[123] Wikipedia: Magnetic resonance imaging. `https://en.wikipedia.org/wiki/Magnetic_resonance_imaging` (2019)

[124] Wikipedia: Medical imaging. https://en.wikipedia.org/wiki/Medical_imaging (2019)

[125] Wikipedia: Menisci. https://en.wikipedia.org/wiki/Meniscus_(anatomy) (2019)

[126] Wikipedia: Thigh. https://en.wikipedia.org/wiki/Thigh (2019)

[127] Wluka, A.E., Forbes, A., Wang, Y., Hanna, F., Jones, G., Cicuttini, F.M.: Knee cartilage loss in symptomatic knee osteoarthritis over 4.5 years. Arthritis research & therapy **8**(4), R90 (2006)

[128] Xia, Y., Fripp, J., Chandra, S.S., Schwarz, R., Engstrom, C., Crozier, S.: Automated bone segmentation from large field of view 3d mr images of the hip joint. Physics in Medicine & Biology **58**(20), 7375 (2013)

[129] Xia, Y., Fripp, J., Chandra, S.S., Walker, D., Crozier, S., Engstrom, C.: Automated 3d quantitative assessment and measurement of alpha angles from the femoral head-neck junction using mr imaging. Physics in Medicine & Biology **60**(19), 7601 (2015)

[130] Xu, C., Xu, L., Brahm, G., Zhang, H., Li, S.: Mutgan: Simultaneous segmentation and quantification of myocardial infarction without contrast agents via joint adversarial learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 525–534. Springer (2018)

[131] Xu, C., Prince, J.L.: Gradient vector flow: A new external force for snakes. In: cvpr. vol. 97, p. 66. Citeseer (1997)

[132] Xu, Z., Shen, Z., Niethammer, M.: Contextual additive networks to efficiently boost 3d image segmentations. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 92–100. Springer (2018)

[133] Xue, Y., Tang, H., Qiao, Z., Gong, G., Yin, Y., Qian, Z., Huang, C., Fan, W., Huang, X.: Shape-aware organ segmentation by predicting signed distance maps. arXiv preprint arXiv:1912.03849 (2019)

[134] Yan, Z., Zhan, Y., Peng, Z., Liao, S., Shinagawa, Y., Zhang, S., Metaxas, D.N., Zhou, X.S.: Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition. IEEE transactions on medical imaging **35**(5), 1332–1343 (2016)

[135] Yan, Z., Zhang, S., Liu, X., Metaxas, D.N., Montillo, A., et al.: Accurate whole-brain segmentation for alzheimer's disease combining an adaptive statistical atlas and multi-atlas. In: International MICCAI Workshop on Medical Computer Vision. pp. 65–73. Springer (2013)

[136] Yan, Z., Zhang, S., Tan, C., Qin, H., Belaroussi, B., Yu, H.J., Miller, C., Metaxas, D.N.: Atlas-based liver segmentation and hepatic fat-fraction assessment for clinical trials. Computerized Medical Imaging and Graphics **41**, 80–92 (2015)

[137] Yang, D., Wu, P., Tan, C., Pohl, K.M., Axel, L., Metaxas, D.: 3d motion modeling and reconstruction of left ventricle wall in cardiac mri. In: International Conference on Functional Imaging and Modeling of the Heart. pp. 481–492. Springer (2017)

[138] Yang, D., Xu, D., Zhou, S.K., Georgescu, B., Chen, M., Grbic, S., Metaxas, D., Comaniciu, D.: Automatic liver segmentation using an adversarial image-to-image network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 507–515. Springer (2017)

[139] Yang, D., Zhang, S., Yan, Z., Tan, C., Li, K., Metaxas, D.: Automated anatomical landmark detection ondistal femur surface using convolutional neural network. In: 2015 IEEE 12th international symposium on biomedical imaging (ISBI). pp. 17–21. IEEE (2015)

[140] Yang, M., Yu, K., Zhang, C., Li, Z., Yang, K.: Denseaspp for semantic segmentation in street scenes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3684–3692 (2018)

[141] Yi, J., Wu, P., Huang, Q., Qu, H., Liu, B., Hoeppner, D.J., Metaxas, D.N.: Multi-scale cell instance segmentation with keypoint graph based bounding boxes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 369–377. Springer (2019)

[142] Yi, J., Wu, P., Jiang, M., Huang, Q., Hoeppner, D.J., Metaxas, D.N.: Attentive neural cell instance segmentation. Medical image analysis **55**, 228–240 (2019)

[143] Yu, H., Tan, C., Yan, Z., Miller, C., Fuerst, T., Engelke, K., Metaxas, D.: Age-related changes of thigh muscle fat infiltration in subjects of the oai mri cohort with and without radiographic osteoarthritis. In: OSTEOPOROSIS INTERNATIONAL. vol. 26, pp. S325–S326. SPRINGER LONDON LTD 236 GRAYS INN RD, 6TH FLOOR, LONDON WC1X 8HL, ENGLAND (2015)

[144] Yu, Y., Zhang, S., Li, K., Metaxas, D., Axel, L.: Deformable models with sparsity constraints for cardiac motion analysis. Medical image analysis **18**(6), 927–937 (2014)

[145] Zhang, C., Li, L., Forster, B.B., Kopec, J.A., Ratzlaff, C., Halai, L., Cibere, J., Esdaile, J.M.: Femoroacetabular impingement and osteoarthritis of the hip. Canadian Family Physician **61**(12), 1055–1060 (2015)

[146] Zhang, D.Q., Chen, S.C.: A novel kernelized fuzzy c-means algorithm with application in medical image segmentation. Artificial intelligence in medicine **32**(1), 37–50 (2004)

[147] Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S.: Deformable segmentation via sparse shape representation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011, pp. 451–458. Springer (2011)

[148] Zhang, S., Zhan, Y., Metaxas, D.N.: Deformable segmentation via sparse representation and dictionary learning. Medical Image Analysis **16**(7), 1385–1396 (2012)

[149] Zhang, Z., Luo, P., Loy, C.C., Tang, X.: Learning and transferring multi-task deep representation for face alignment. CoRR (2014)

[150] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. Medical Imaging, IEEE Transactions on **27**(11), 1668–1681 (2008)