Fall 2016

Quantitative analysis and segmentation of knee MRI using layered optimal graph segmentation of multiple objects and surfaces

Satyananda Kashyap
University of Iowa

QUANTITATIVE ANALYSIS AND SEGMENTATION OF KNEE MRI USING LAYERED OPTIMAL GRAPH SEGMENTATION OF MULTIPLE OBJECTS AND SURFACES

by

Satyananda Kashyap

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy degree
in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

December 2016

Thesis Supervisor: Professor Milan Sonka

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Satyananda Kashyap

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the December 2016 graduation.

Thesis Committee: _____
Milan Sonka, Thesis Supervisor

_____
Punam K. Saha

_____
Donald D. Anderson

_____
Reinhard Beichel

_____
Xiaodong Wu

ACKNOWLEDGEMENTS

This journey of my Ph.D. has been a long and grueling one with several high and low points throughout the period. I would not have been able to complete it without the wonderful support of the people all around me. First I would like to thank my advisor Dr. Milan Sonka from whom I have learned a lot over the years. His expertise, vision, and patient guidance helped shape my thinking and develop the project into what you see in this thesis. His advice on how to think critically, formulate a research problem and ask the right kind of questions has improved me as a researcher and will hold me in good stead throughout my career. But most of all, I want to thank him for believing in my abilities at times when I did not think I could do it myself.

I would like to thank Dr. Ipek Oguz who helped me immensely especially sitting down with me and working through solving several of the implementation and algorithmic issues in the project. It was really fun brainstorming and debugging several of the algorithms used in the thesis. Thank you for your patience, guidance and ever willing nature to help me.

I acknowledge Dr. Honghai Zhang for helping me with developing the interactive aspects of the methods. His knowledge on GUI design really provided a platform for me to create the interactive algorithms for knee MRI. Without Honghai's basic template, there is no way I would have been able to create all that has been shown in this work from scratch. Thank you Dr. Andreas Wahle and the rest of the IIBI group for your guidance and support during the course of my Ph.D.

To my committee members, Drs. Punam Saha, Donald Anderson, Reinhard Beichel and Xiaodong Wu, thank you for taking pains to understand my research and offer valuable insights and corrections to improve the quality of this work.

I would like to thank Matt, Dan, Zach and the rest of the team at CSS who helped me with every one of my computing needs. Cathy, Dina & Kim for helping me with all the administrative aspects of grad school.

My life in Iowa City during my Ph.D. journey was made infinitely better thanks to all my wonderful friends. Thank you Ipshita, Vishwa, Sajan, Sampada, Sujit, Nirmal, Oishik, Phatak and the others for a lifetime of awesome memories and friendship.

My parents who I would like to thank immensely for encouraging me to pursue my Ph.D. and for being understanding throughout my studies. Whatever I am today is all due to their blessings and good wishes. I affectionately dedicate my thesis to them. I would also like to thank my late Ajji and Tata both of whom passed away earlier this year. I am sure they would have been very happy to see me complete. Finally my guru for all his blessings and spiritual guidance for helping me successfully complete this work.

ABSTRACT

Knee osteoarthritis is one of the most debilitating aging diseases as it causes loss of cartilage of the knee joint. Knee osteoarthritis affects the quality of life and increases the burden on health care costs. With no disease-modifying osteoarthritis drug currently available there is an immediate need to understand the factors triggering the onset and progression of the disease. Developing robust segmentation techniques and quantitative analysis helps identify potential imaging-based biomarkers that indicate the onset and progression of osteoarthritis. This thesis work developed layered optimal graph image segmentation of multiple objects and surfaces (LOGISMOS) framework based knee MRI segmentation algorithms in 3D and longitudinal 3D (4D). A hierarchical random forest classifier algorithm was developed to improve cartilage costs functions for the LOGISMOS framework. The new cost function design significantly improved the segmentation accuracy over the existing state of the art methods. Disease progression results in more artifacts appearing similar to cartilage in MRI. 4D LOGISMOS segmentation was developed to simultaneously segment multiple time-points of a single patient by incorporating information from earlier time points with a relatively healthier knee in the early stage of the disease. Our experiments showed consistently higher segmentation accuracy across all the time-points over 3D LOGISMOS segmentation of each time-point. Fully automated segmentation algorithms proposed are not 100 % accurate especially for patient MRI's having severe osteoarthritis and require interactive correction. An interactive technique called just-enough interaction (JEI) was developed which added a fast correction step to the automated LOGISMOS, speeding up the interactions substantially over the current slice-by-slice manual editing while maintaining high accuracy. JEI editing modifies the graph nodes instead of the boundary surfaces of the bones and cartilages providing globally optimally corrected results. 3D JEI was extended to 4D JEI allowing for simultaneous visualization and interaction of multiple time points of the same

patients.

Further quantitative analysis tools were developed to study the thickness losses. Nomenclature compliant sub-plate detection algorithm was developed to quantify thickness in the smaller load bearing regions of the knee to help understand the varying rates of thickness losses in the different regions. Regression models were developed to predict the thickness accuracy on a patient MRI at a later follow-up using the available thickness information from the LOGISMOS segmentation of the current set of MRI scans of the patient. Further non-cartilage based imaging biomarker quantification was developed to analyze bone shape changes between progressing and non-progressing osteoarthritic populations. The algorithm quantified statistically significant local shape changes between the two populations. Overall this work improved the state of the art in the segmentation of the bones and cartilage of the femur and tibia. Interactive 3D and 4D JEI were developed allowing for fast corrections of the segmentations and thus significantly improving the accuracy while performing many times faster. Further, the quantitative analysis tools developed robustly analyzed the segmentation providing measurable metrics of osteoarthritis progression.

PUBLIC ABSTRACT

Knee osteoarthritis is one of the most debilitating aging diseases as it causes loss of cartilage of the knee joint with nearly 30.8 million adults affected in the US between 2008-2011. Further the disease is expected to affect nearly 78 million people in the US by 2040. Knee osteoarthritis affects the quality of life and increases the burden on health care costs. With no disease-modifying osteoarthritis drug currently available there is an immediate need to understand the factors triggering the onset and progression of the disease. Developing robust segmentation techniques and quantitative analysis helps identify potential biomarkers causing the initial onset of the disease and its progression.

This thesis work developed state of the art graph based knee MRI segmentation algorithms in 3D. The work also developed the world's first longitudinal 3D (4D) graph based segmentation algorithm which simultaneously segmented multiple follow-up visits of the patient. This technique was demonstrated to be highly beneficial in tackling the hard to segment follow-up MRI's with severe disease by using information from healthy MRI scans taken from earlier visits. Machine learning algorithms were used to improve the accuracy of the segmentation task further. The limitation of any fully automated technique is that they are not 100% accurate every single time. An interactive technique called just-enough interaction (JEI) was developed which added a fast correction step to the automated method, providing fast interactive correction while maintaining high accuracy.

To quantify osteoarthritis, several automated tools were developed to study the thickness losses using the segmented bones and cartilages of the knee joint. Algorithms were developed to quantify thickness losses in the smaller load bearing regions of the knee joint which are the most affected by osteoarthritis. Regression models were designed to predict the thickness accuracy on a patient MRI at a later follow-up using the available thickness information of the current set of MRI scans of the patient.

Further bone shape changes were studied to understand differences between progressing and non-progressing osteoarthritic populations. Overall this work improved the state of the art in the segmentation of the bones and cartilage of the femur and tibia. Interactive 3D and 4D JEI were developed allowing for fast corrections of the segmentations and thus significantly improving the accuracy while performing many times faster. Further, the quantitative analysis tools developed robustly analyzed the segmentation providing measurable metrics of osteoarthritis progression.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTRODUCTION

1.1 Anatomy of The Knee Joint

The knee is one of the most complex joints in the human body. It is made up of the thigh bone (femur), shin bone (tibia), the knee cap (patella) and a smaller bone running alongside the tibia (fibula), (Fig. 1.1). Soft connective tissues such as tendons connect the knee bones to the muscles in the leg which help in movement. Articular cartilage is found covering the ends of the femur, tibia and the patella bone in the regions where they come in contact with each other. They are adapted to resist and dampen the compressive and tensile forces which occur during the movement of the knee joint. There are four ligaments in the knee joint responsible for providing stability to the knee: 1. Anterior cruciate ligament that prevents the femur from sliding backward on the tibia, 2. Posterior cruciate ligament which prevents the femur from sliding forward on the tibia and 3. Medial and Lateral collateral ligaments that protect the femur from sliding side to side. In order to provide a smooth contact surface between the femur and tibia cartilage two C-shaped pieces of cartilage called medial and lateral menisci occur in between the femur and tibia. These cartilages along with the fluid-filled sacs help in the smooth movement of the knee joint.

1.2 Significance of Osteoarthritis as a Disease

Osteoarthritis (OA) is a disease which causes loss of articular cartilage in the joints over time. OA of the knee is one of the leading causes of functional disability among the aging. Estimates project that the disease prevalence will double by 2020 [1]. The disability risk attributed to knee OA is greater than that caused by any other medical conditions amongst the aging population [23].

Currently, there are no drugs that can halt the progress of OA let alone reverse the degradation. The only treatments available are therapeutic which helps to manage the progressively worsening pain using oral pain medications, intravenous steroid

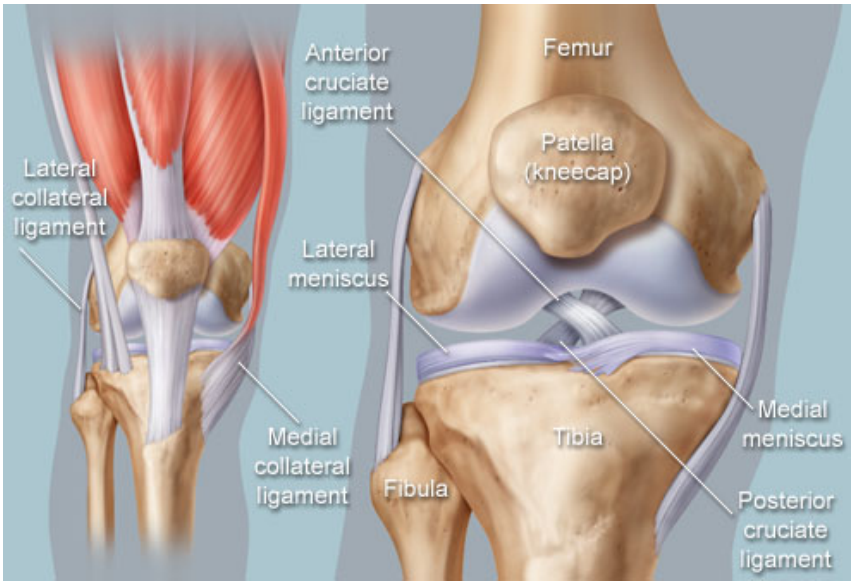


Figure 1.1: The anatomy of the knee. *Source: WebMD*

injections, and/or weight management. Other methods to alleviate pain includes wearing support braces which help ease the pressure in the joint where there is no cartilage. The placebo effect can also be a power tool for therapeutic relief. The sham surgery reported by Moseley et al. [44] is a famous example of the placebo effect. The patients were divided into three groups out of which two of them received actual arthroscopic lavage (Group 1) and arthroscopic débridement[1] (Group 2). The placebo group (Group 3) only had similar incisions and underwent a simulated débridement without arthroscope insertion. It turned out that at no point after the surgery did the outcome of the control groups (Groups 1 & 2) show greater improvement than the placebo group in terms of pain reduction and improved functionality.

Upon onset of OA, patient's knee joint gradually worsens loosing functionality and causing increasing pain. Eventually, the loss of cartilage will completely denude the joint causing bone-on-bone interaction which can be extremely painful. Further research has shown that the loss of cartilage accelerates towards the end of the disease

[1]The trimming and removal of loose debris of osteophytes, extruded meniscus, synovium or cruciate ligament .

[17] leading to complete loss of joint functionality and eventually forcing the patients to undergo total knee replacement (TKR). Therefore it is crucial to understand what triggers the onset of OA.

1.3 Role of Imaging in Knee Osteoarthritis

Several large-scale clinical studies for developing a Disease Modifying Osteoarthritis Drug (DMOAD) are underway. Imaging is crucial in these studies to determine the effect of such drugs on disease progression. For DMOAD, the guidelines recommend that the benefits of reduction of joint tissue pathology should be accompanied with clinical outcome benefits, i.e., reduction of joint pain and improvement of function [18]. Therefore it is especially important to study the relationship between imaging biomarkers and clinical outcomes. The two most commonly used techniques to image the knees are radiographs and Magnetic Resonance Imaging (MRI).

Radiography is an inexpensive and simple technique which can be used to detect OA associated features such as osteophytes and subchondral cysts. Commonly the knee is imaged in its weight bearing, anteroposterior (AP) view, fixed-flexion radiography with 10^0 caudal beam angle on both the knee joints [35,36]. In order to obtain reproducible radiographs, a specially designed frame by SynaFlexer is used that helps standardize the positioning of the knee during imaging. An indirect measure of the cartilage thickness and meniscal positioning can be obtained by determining the joint space width (JSW). The JSW calculation from a projection of a 3D object onto a 2D plane has several limitations with studies showing that the lack of sensitivity and specificity of radiographs in detecting OA associated features and changes in the follow-up radiographs inherently limiting their use [26].

MRI has several advantages over radiographs. Unlike radiographs which project a volume onto a 2D plane, MRI can obtain detailed 3D images of the full knee joint and can be used to analyze more accurate quantitative indices. MRI can directly visualize different components of the knee besides articular cartilage such as liga-

ments, synovial fluid, meniscus, bone marrow lesions (BML's) and cysts. Studies have shown that quantitative measurements from MRI are more sensitive to cartilage losses than qualitative methods [57]. In [53] the study showed an increased risk of radiographic OA from examining the structural damage on patient MRI from 2 years prior. Different measurements on MRI can be enhanced further by using the appropriate pulse sequences. Therefore good pulse sequence designs are essential for good quality images of the specific tissue to be studied. Several sequences have been developed to study the physiological content of articular cartilage. The onset of OA reduces the collagen and proteoglycan content thereby disrupting the collagen network. As a result, the water content increases and further matrix degradation occurs [13]. Transverse mapping (T2) is shown to be useful in analyzing cartilage composition with its spatial distribution linked with OA severity [15]. $T_{1\rho}$ imaging was found to be sensitive to collagen and correlated with the proteoglycan concentrations in the cartilage. $T_{1\rho}$ mapping was also found to have increased values in knee cartilage with OA in comparison with normal cartilage [52].

Contrast based techniques such as delayed Gadolinium-enhanced MRI (dGEMRIC) use a negatively charged contrast agent such as Gd(DTPA) (Magnevist; Berlex-Imaging, Wayne, NJ, USA) which is injected intravenously and imaged after around 90 minutes. This contrast agent accumulates in low concentration in areas rich in negatively charged glycosaminoglycan (GAG) found in abundance in healthy cartilage. Whereas there is high accumulation in regions lacking GAG. Imaging this gives us a dGEMRIC index distribution where low index indicates high Gd(DTPA) accumulation and therefore greater GAG depletion indicating damaged cartilage [41].

Gradient-recalled echo (GRE) based sequences such as 3D spoiled gradient echo at steady state (SPGR), fast low-angle shot (FLASH) and double echo steady state (DESS) provide excellent cartilage to subchondral bone contrast and nearly isotropic spatial resolution which helps to accurately measure changes in the bone and the

cartilage structures. They are most commonly used in research for providing detailed quantitative analysis of several structures. Over longitudinal studies, these sequences can visualize the loss of thickness in cartilage and the growth of several pathologies and lesions with worsening of the disease. Overall they provide a comprehensive view of the structure of the bones and cartilages changes in the knee joint along with the other pathologies that are found to be useful indicators.

1.4 Quantitative and Semi-quantitative Imaging Biomarkers

For an imaging biomarker to be considered useful it should lie on the disease pathway, i.e., the biomarker should detect changes in the disease progression and report them with sufficient accuracy and precision. OA is measured either quantitatively or as semi-quantitative grading schemes in radiographs. Quantitative JSW measurements are done either manually or semi-automatically/automatically using software.

The most popular semi-quantitative measure is the Kellgren and Lawrence (KL) [32] classification system. The presence of osteophytes and progression of the joint space narrowing (JSN) are the main criteria used to assess the progression of OA. Lack of JSW indicated by bone to bone interaction is an indicator for TKR. Image repositories of radiographs of different disease states are used as references to assigning KL grades. These grades range from 0 to 4 with 0 indicating a healthy knee and 4 indicating a severely diseased knee. A KL grade of 0 indicates a healthy knee with no evidence of OA. Some possible JSN and presence of osteophytes are assigned a KL grade of 1. KL grade of 2 is assigned to radiographs with definite osteophytes and possible JSN in the weight bearing regions. As the disease progresses multiple osteophytes and bone deformations with definite JSN and sclerosis can be seen in KL grade of 3. Finally, KL grade of 4 indicates the presence of significant JSN, severe bone deformities, and sclerosis. The KL grading scheme has several limitations. For example KL grade of 3 can include all different degrees of JSN irrespective of the

actual extent [18]. More recently, the Osteoarthritis Research Society International (OARSI) atlas was developed which is graded based on individual compartments of the knee. The grades are based on the extent of osteophytes and JSW in the tibiofemoral regions of each compartment. This method has been shown to be more sensitive to longitudinal changes than KL grading [3].

Several semi-quantitative scoring systems exist for scoring MRI's [61]. They all primarily grade the focal losses of cartilage on a five-point scale. More recently introduced scores provide a more complex analysis of the cartilage and also scoring other tissues involved in OA such as bone marrow, and meniscus. The two most commonly used ones are whole-organ magnetic resonance imaging score (WORMS) and Boston-Leeds Osteoarthritis Knee score (BLOKS). The real advantage of quantitative measurements over semi-quantitative scoring methods is that they are less prone to observer biases. They are more objective and relatively small changes in the cartilage thickness or volume can be detected longitudinally [19].

As mentioned in the earlier section, GRE based sequences such as DESS/FLASH are very good in providing quantitative analysis. Common quantitative measures computed are thickness, the total area of the subchondral bone, cartilage area, thickness over the total area of the bone. A consensus compliant nomenclature for subplate definitions and structural features were defined by experts [16]. For each sub-region defined the above-mentioned quantitative indices were computed. In order to compute these accurate segmentation needs to be performed either manually or using automated/semi-automated algorithms.

The potential link between cartilage morphology as a surrogate measure of disease progression and clinical outcomes (such as TKR) have been studied. In [11], a four-year study conducted found out that the rate of change in cartilage volume from baseline (BL) to 2 years was an independent predictor of TKR at 4 years. Further, the study concluded that for every 1% increase in the cartilage loss rate, the risk

factor for undergoing TKR increased to 20%. In contrast, the radiographic scores of OA did not predict TKR in the same study. In [17], a 4-year trajectory of femorotibial cartilage thickness loss was examined in a case-control study. The study found that in cases with subsequent knee replacement the losses were found to accelerate 2 years prior to surgery. The most responsive sub-region were found to be the central medial tibia (cMT) in terms of identifying the accelerated losses in patients with TKR. Several similar such studies have shown sensitivity to change of cartilage thickness and volume loss. Studies report factors which accelerate the loss of cartilage seen useful for enriching clinical trials have been reported. These include misalignment (varus/valgus), advanced KL grade at BL, pain, and high body mass index (BMI). In [9], it was reported that patients with early radiographic OA displayed thickening of the peripheral sub-regions.

Besides the cartilage, other structures such as bone and meniscus have been studied. Bone shape vector has been shown to be a promising marker which indicates changes in the structure early [45]. In the study, they concluded that bone shape vector is an important biomarker which could detect changes 12 months before the incident change in Kellgren - Lawrance (KL) grade (KL $\geq$ 2). OA knees were found to have a larger increase in bone area longitudinally than those without OA [6]. Denuded area of subchondral bone (dAB) has been associated with knee pain [43]. Meniscus extrusion was able to predict progression but was suggestive of a lack of increase in sensitivity to change [51].

Trochlear dysplasia has been identified as a potential non-cartilage biomarker of OA, characterized by the abnormal shape and depth of the trochlear groove. The work by Pfirrmann et al. [50] was the first attempt in a MR-based identification of the trochlear dysplasia. Jungmann et al. [29] found that shallow trochlear depth was associated with reduced T2 relaxation times and lowered cartilage volume in comparison with the healthy controls. In both studies, the trochlear groove was

manually segmented in 2D and abnormalities were identified based on thresholding of a 1D depth measurement. Pedoia et al. semi-automatically segmented the femur and tibia to create statistical shape models and reported a reduction of surface area around the trochlear groove region [48]. In [45], the authors used an active shape model to automatically segment the femoral bones and used a linear discriminant analysis classifier to distinguish between healthy and OA knees and reported a narrowing of the trochlear groove. All the above methods quantify either individual 2D slices or report high-level shape changes for the entire structure.

All these studies point towards the importance of quantitative indices as biomarkers in the study of OA highlighting the promises and pitfalls of current biomarkers and the need to identify novel biomarkers to aid the understanding of disease progression, especially in intervention drug trials.

1.5 Background of Automated Segmentation Techniques used in Analyzing Knee MRI

Several of the quantitative indices mentioned in the previous section requires accurate segmentation of the knee as the first step. Segmentation is performed either manually by the reader or using semi-automated/automated algorithms. Here we focus on a survey of existing automated algorithms available in the literature. A survey of semi-automated algorithms is presented in the next section.

Automated segmentation of knee MRI allows more objective and more efficient quantification of cartilage than manual segmentation by trained radiologists, which can take several hours of effort and is prone to inter/intraobserver variability and operator-induced biases. However, the complex anatomy of the knee joint, the challenges from the MR appearance of the cartilage make fully automated segmentation a challenging task. In particular, the closely positioned cartilage and bone surfaces at the interface of the femur and tibia are highly difficult to distinguish from each other. As a result, methods segmenting the bones and cartilages of the femur and

tibia objects without considering the contextual information between them frequently fail.

Several segmentation algorithms have been previously proposed for this task. Approximate binary k-NN classifiers was proposed by Folkesson et al. [20]. To correct for variations in placement of test subjects in the scanner a corrective position adjustment was applied iteratively. They validated their algorithm on both healthy and OA knees acquired on low field scanners. Dodin et al. proposed using a ray casting technique to segment the bones of the femur and tibia [14]. They formulated the problem by decomposing the MR images into multi-surface layers localizing the boundaries of bones and several partial segmentation objects which are merged to obtain the final complete bone segmentation. In [22], a fully automated segmentation method was proposed for non-pathological knees. A 3D active shape model was used as initialization to extract the bone-cartilage interface; the cartilage was then segmented using a deformable model with patient-specific tissue estimation. Markov random fields were used to construct and optimize local image patches for region and boundary probabilities from local shape and appearance information in [39]. Lee et al. demonstrated a fully automated segmentation of 3D MR images using a multi-atlas model and local structural scheme [38]. Their methodology built a multi-label atlas which was merged using a locally weighted voting followed by region adjustment. A hierarchical two-stage random forest classifier system was used for cartilage classification in [66]. The algorithm used multi-label graph cuts to optimize the background (combining bone regions with the true background as a single label) and the cartilage classifier outputs.

The surveyed segmentation techniques make use of locally optimal strategies to solve the segmentation leading to sub-optimal solutions especially for pathological cases. For example, Wang et al. [66] used multi-label graph cuts to optimize the background combining bone regions with true background and the cartilage classifier

outputs. The proposed framework in this thesis guarantees globally optimal solutions with respect to the cost functions provided. Our work also focuses on designing highly accurate machine learning based costs functions to further improve the segmentation accuracy.

1.6 Interactive Segmentation Techniques

State-of-the-art automated segmentation algorithms are not 100% accurate, especially when segmenting difficult to interpret datasets like those with severe osteoarthritis (OA). With disease progression, automated algorithms face a challenging problem of delineating the bones and cartilages in presence of bone marrow lesions, cartilage surface thinning, meniscal extrusion and synovial fluid leakage. Many of these artifacts are symptoms of the disease and appear similar in texture and intensity to the cartilage on MR volumes. Therefore interactive correction methods are designed to help ease the post-processing needed.

Several such techniques have been proposed in the literature. Ross et al. [54] demonstrated an interactive segmentation of the pulmonary fissure surface from user inputted seeds and thin plate spline interpolation. Mortensen et al. [4] demonstrated the 2D live-wire based interaction by placing user inputted control points to denote the object border that was solved using the graph optimization. In [55], the live-wires were extended to 3D by using 2D live wires on a few slices to generate the contours and using prior shape knowledge to interpolate for the remaining slices. Live-wires have embedded user interactions that require a seed based initialization which is user inputted or based on some initialization algorithm. However, they are inherently 2D mechanisms which were later extended to 3D with a drawback of being unable to maintain global optimality for multiple surfaces and objects. Schwarz et al. [56] used an active shape model for interactive editing by utilizing the user inputs to mark the expected boundary location. This was corrected by determining the new surface points near the clicked points followed by the shape model updating. In Boykov's

graph cuts [7] the user marked the approximate foreground and background regions of the image which were used to construct the appropriate cost functions for the underlying graph. The graph was solved using a max-flow optimization algorithm. Several of the above-mentioned methods correct for segmentation inaccuracies by directly matching the object boundaries with the interaction which after several repetitions results in the final surfaces having local topological errors.

1.7 Problem Statement and Thesis Aims

The thesis aims can be divided into the following main categories :

- Aim 1: Improve and validate a LOGISMOS based method for fully automated segmentation of knee-joint bones and cartilages in 3D MRI
- Aim 2: Develop and validate a 4D LOGISMOS method utilizing temporal contextual information to improve the segmentation accuracy in longitudinal studies.
- Aim 3: Develop just-enough-interaction (JEI) approaches to refine 3D and 4D segmentations of the knee.
- Aim 4: Develop automated tools to better quantify changes in the bone and cartilage due to osteoarthritis.

The rest of the thesis is organized as follows: Chapter 2 introduces the basic graph search and the LOGISMOS segmentation pipeline for knee MRI. We describe the gradient based costs and the learning based hierarchical cost function designed for knee MRI. Chapter 3 introduces the automated sub-plate thickness detection algorithm to quantify changes in cartilage thickness in the load bearing regions of the knee. In Chapter 4 we extend the 3D LOGISMOS graph algorithm to handle multiple time points of the same patient detailing the non-trivial graph extension to incorporate the inter-time point constraints. Chapter 5 introduces the just-enough

interaction (JEI) techniques that were extended to handle knee MRI detailing the 3D as well as the 4D JEI algorithms. Chapter 6 uses the segmentations to develop different automated quantitative analyses demonstrating the powerful capabilities of the different methods in understanding the changes to bone and cartilage due to osteoarthritis. Finally, we provide a summary of the contributions, limitations of the algorithms from this thesis and a glimpse of possible future directions this research can be directed towards.

CHAPTER 2
3D LOGISMOS ALGORITHM FOR KNEE MRI AND COST FUNCTION DESIGN

2.1 Motivation

Several of the surveyed segmentation algorithms make use of local optimization methods to solve the segmentation problem which can give sub-optimal solutions especially in challenging to segment disease cases. The LOGISMOS framework based knee MRI segmentation guarantees globally optimal solution with respect to the cost functions provided. This framework was developed in [30] and extended to segment multiple objects and surfaces simultaneously [70].

2.2 Basics of the knee segmentation algorithm

The LOGISMOS algorithm models the n-D simultaneous detection of multiple interrelated objects and surfaces as a $(n+1)$-D geometric graph with the surface interrelations modeled as graph arcs between the desired surfaces. The combinatorial optimization problem is solved by computing the minimum s-t cut on an appropriately transformed geometric graph as described in previous works [30, 70].

In the following paragraphs, detailed explanations of each of the steps in the LOGISMOS algorithm pipeline are provided.

1. **Pre-segmentation & Mesh Generation:**

   Pre-segmentation yields approximate mesh surfaces of the target objects. The approximate surface detection is done using various methods such as Active Shape Model (ASM), level-sets, atlas-based registration [60] and single-surface graph segmentation. Generally, if the geometry is simple (terrain like, spherical or cylindrical) the first step is a trivial task of fitting corresponding simple shapes.

   For the LOGISMOS applied to knee MRI, the pre-segmentation consists of three main steps : volume of interest (VOI) detection, affine fitting of a mean

shape model followed by a single surface graph segmentation of the bone. VOI detection reduces the computation by isolating volumes to include the femur and tibia bone and corresponding articular cartilages. The VOI detection is done using an AdaBoost classifier [21] trained on manually isolated VOI's. Nine different Haar-like features at different scale responses are applied on the input volume converted to an integral image. The difference in the pattern based on the subtraction and addition of the areas under the filter were used as features to train the AdaBoost classifier and to isolate VOI on the test MRI volume.

Mean shape meshes $\bar{\mathcal{S}}_0$ were created from an earlier training set for the femur and tibia shapes respectively. Using the VOI bounds, $\bar{\mathcal{S}}_0$ were fitted using an affine transform. Ensuring well-defined bone structure is important for the final simultaneous LOGISMOS segmentation of bones and cartilages of the knee joint. Therefore after fitting $\bar{\mathcal{S}}_0$, single surface LOGISMOS segmentation is performed independently for the bones of the femur and tibia producing accurate patient-specific meshes $\mathcal{S}$ (Fig. 2.1). The graph column construction is similar to the multi-surface multi-object version explained below.

2. **Image Re-sampling for Graph Column Construction:** The graph column construction is an important step in mesh based graph optimization. Graph columns drawn as simple normals to the mesh surface results in more than one surface intersection per graph column. This violates the graph search constraint of exactly one intersection per graph column. We instead borrow the idea from electric field theory that multiple charged particles exhibit non-intersecting electric lines of force. We mimic this behavior by simulating mesh surface points as holding like charges. Using Coulomb's law the graph columns constructed follow the electric lines of force growing outwards from a charged body and therefore do not intersect.

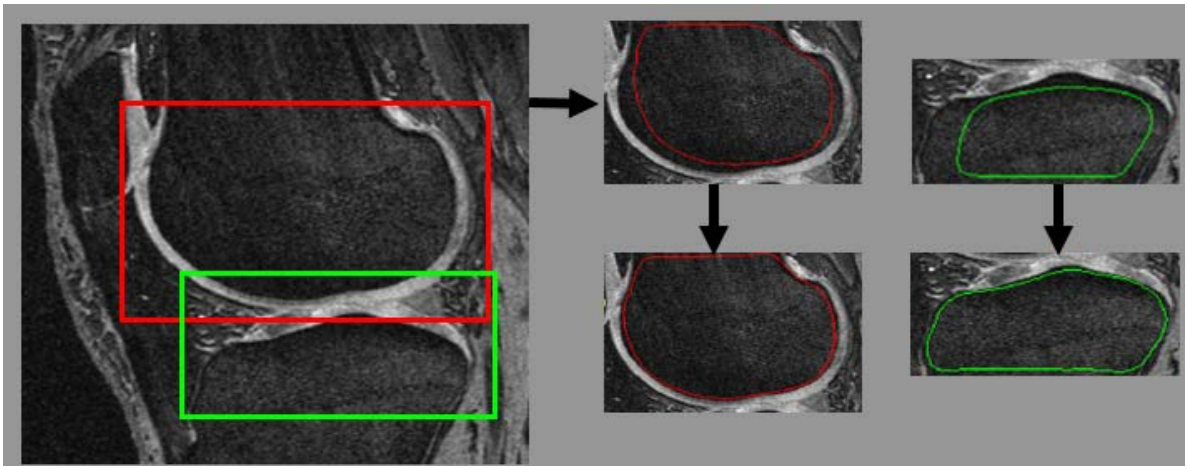


Figure 2.1: 2D representative slice view of the pre-segmentation of the knee. *Left*, the VOI extraction for the femur (red) and tibia (green). *Top row*, Affine fitted mean shapes $\bar{\mathcal{S}}_0$. *Bottom row*, accurate patient-specific bone structure $\mathcal{S}$ resulting from a single-surface LOGISMOS segmentation.

Since the ELF computation is done on a discrete mesh we modify it as follows:

(a) To cope with the non-uniform vertex distribution on the mesh, each vertex (v_i) on the mesh surface is assigned a positive charge computed as the sum of associated surface areas with the triangles t_j that are adjacent to v_i.

(b) The radius in the denominator follows the inverse square law. In order to reduce the influence of mesh points further away from the point of interest, the radius is modified from r^2 to r^m ($m > 2$) which decreases the influence of more distant vertices and increases robustness.

This results in the following formulation of the ELF field:

$$\hat{E} = \sum_i \frac{\sum_j AREA(t_j)}{{r_i}^m} \hat{\mathbf{r}}_i \ , \tag{2.1}$$

where $v_i \in t_j$ and $m > 2$.

3. **Graph Search Setup and Segmentation** An important innovation of the

method is the non-trivial graph construction which transforms the surface segmentation problem into a combinatorial optimization of computing the minimum closed set in a node-weighted digraph. A closed set in the digraph Z is a subset of nodes such that all the outgoing arcs from any of the nodes in Z are also contained with it. The total cost of the closed set is the sum of the costs of all the nodes belonging to the set. The optimization searches for the closed set that has the minimum cost by computing the s–t cut in a derived arc-weighted digraph [25] in polynomial time.

Here we first explain the construction of the graph for a single surface (used for presegmentation) with the basic graph constraints to find a feasible surface.

(a) **Single Surface Graph Segmentation:** A node weighted digraph $G = (V, E)$ is built with each ELF vector corresponding to the nodes in a single graph column. Every node ($v_i(k) \in V$) per column i with $k \in K$ nodes/-column represents a single re-sampled voxel in the image ($I(x, y, z)$) with a corresponding cost $w(x, y, z)$ associated to it where $x \in \mathbf{x} = \{0, \ldots, X-1\}$, $y \in \mathbf{y} = \{0, \ldots, Y-1\}$ and $z \in \mathbf{z} = \{0, \ldots, Z-1\}$ are the VOI bounds. The node costs can encode both region-based and edge-based costs. They represent the inverse likelihood that the desired surface contains the voxel. Neighborhood relations between voxels are represented by defining the adjacency relations among the columns of G. Every column on the graph intersects the sought after surface with exactly one node. To ensure that the surface is considered feasible, graph arcs are introduced to enforce application specific smoothness constraints ($\delta_{smoothness}$). To ensure surface feasibility *intra-column* and *inter-column arcs* are defined:

- *Intra-column arcs* E^{intra_column}: $\forall k$ on each column i, the intra-column

arcs are defined by

$$E^{intra\_column} = \{\langle V_i(k), V_i(k-1)\rangle\}$$

- *Inter-column arcs* E^{inter_column}: Between any two adjacent columns i and j the arcs enforce the smoothness constraints ensuring that the surface cannot jump by more than the desired number of columns above or below.

$$E^{inter\_column} = \{\langle V_i(k), V_j(k-\delta_{smoothness})\rangle \cup \{\langle V_i(k-\delta_{smoothness}), V_j(k)\rangle\}$$

Information on the constraints and cost functions of a target segmentation problem needs to be provided beforehand. An optimal surface is one which has the minimum cost among all feasible surfaces. An example single surface segmentation scenario is shown in Fig. 2.2. For the simultaneous multiple surfaces and objects segmentation, we enforce additional arcs to encode the mutual relationships between these entities.

(b) **Multiple Surfaces/ Multiple Objects Graph Segmentation:**

For segmenting the bones and cartilages of the femur and tibia, multiple sub-graphs as described above are joined together by graph arcs which specify pairwise topological relationships between the surfaces and/or objects that can also encode multi-surface/multi-object constraints. By linking these surface-specific sub-graphs together, a single graph results. In order to enforce minimum and maximum surface separation constraints between the bone and cartilage sub-graphs of each of the femur and tibia

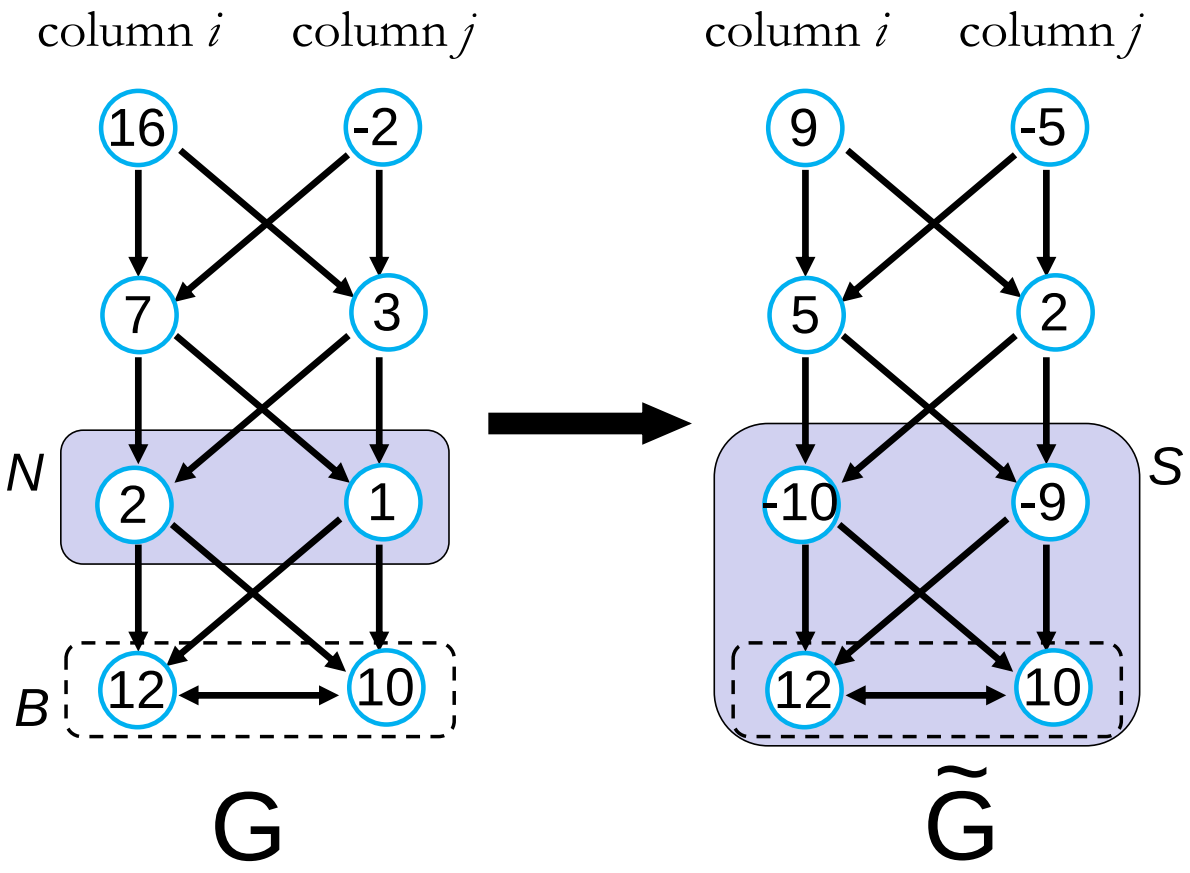


Figure 2.2: Basics of LOGISMOS graph search. Every surface to be segmented is represented as a sub-graph in our problem. For every sub-graph, we have directed intra-column arcs going from the topmost node to the base graph B as shown for columns i and j. We also have directed inter-column arcs which control the smoothness jumps between columns. These two arcs are essential for the graph search to work. The individual nodes have node unlikeliness costs assigned to them in G. The V-weight net problem in G is converted to finding a non-empty closed set on an equivalent graph $\tilde{G}$. Here the column length is $K = 4$. Note that the costs are transformed between the left and right graphs so that the minimum-cost closed set approach can be applied to the graph on the right; see [30, 60] for details of this cost transform.

objects, we introduce pairwise inter-surface arcs. We use apriori knowledge about the anatomically feasible locations of the cartilage for each knee-joint to set the minimum and maximum pairwise distances ($\delta_{minDist},\ \delta_{maxDist}$). In regions with no cartilage, the $\delta_{minDist}$ & $\delta_{maxDist}$ were enforced to zero forcing the cartilage to be of zero thickness in these regions.

- *Inter-surface arcs* $E^{inter\_surface}$: For every $n \in \{cart, bone\}$ the pairwise inter-surface arcs are defined as

$$E^{inter\_surface} = \{\langle V_i(n_{bone}, k), V_j(n_{cart}, k - \delta_{minDist})\rangle \cup \{\langle V_i(n_{bone}, k + \delta_{maxDist}), V_j(n_{cart}, k)\rangle\}$$

The knee anatomy is such that portions of the tibia and cartilage are in close apposition. Therefore inter-object constraints are established so that they do not overlap since it is not anatomically feasible. Using ELF based constraint point mapping [70], inter-object arcs with minimum and maximum inter-object distances ($\delta_{objectMin}$ & $\delta_{objectMin}$) are drawn between the femur and tibia cartilage to ensure no overlap.

- *Inter-object arcs* E^{inter_object}: For every $m \in \{femur, tibia\}$ being the two interacting objects and d being the column offset between objects, the pairwise inter-object arcs are defined as

$$E^{inter\_object} = \{\langle V_i(m_{femur}, n_{cart}, k), V_j(m_{tibia}, n_{cart}, k - d + \delta_{objectMin})\rangle | \forall k :$$
$$\max(d - \delta_{objectMin}, 0) \leq k \leq \min(K - 1 + d - \delta_{objectMin}, K - 1)\}$$
$$\cup\{\langle V_j(m_{tibia}, n_{cart}, k), V_i(m_{femur}, n_{cart}, k + d - \delta_{objectMax})\rangle | \forall k :$$
$$\max(\delta_{objectMax} - d, 0) \leq k \leq \min(K - 1 - d + \delta_{objectMax}, K - 1)\}$$

The graph construction ensures that the desired optimal surfaces correspond to an *optimal closed set* in the weighted directed graph G [30, 69]. Searching for an optimal closed set in G can be solved in polynomial time by computing a minimum s–t cut in a derived arc-weighted digraph [25].

2.3 Cost function design

LOGISMOS algorithm is guaranteed global optimality based on the cost functions provided. Therefore it is crucial to design good application specific cost function.

The costs associated with the bone detection exploits the property of finding a strong dark to bright edge when traversing the graph column from inside of the bone surface out. We designed directional 1D derivative operators which give us the corresponding costs based on the edges encountered along the search lines. For the cartilage, a weighted combination of the first and second derivative operators given by $w_1 * \nabla(x, y, z) + (1 - w_1) * \nabla^2(x, y, z)$ was used, where w_1 was determined experimentally. This helps prevent interpreting cartilage inhomogeneities as edges.

Although human expert designed cost functions are very effective at capturing the desired features there are several challenges : **1)** Choosing the correct weighting combination is challenging. **2)** Same costs may not work for all parts of the anatomy. **3)** The same anatomical objects (eg. bone, cartilage) will appear differently in pathological cases because of the loss of structure and/or the appearance of lesions. Therefore we designed random forest classifiers [8] to provide the cost functions. Random forest

(RF) classifiers use the concept of bagging where for each decision tree, a random subset of features are chosen thereby reducing correlation between the decision trees which improves accuracy.

2.3.1 Classifier System Design

The previously used classifier [70] was a single stage RF classifier. The collected features on the volumetric images were interpolated onto the nodes of the ELF geometric graph. The major drawback of this method was that all the information used were localized along the graph columns. Because of this locality, features failed to capture a neighborhood of information that appears larger than a few nodes along the column. There are several anatomical features that appear locally like a cartilage however when examined in a global neighborhood will reveal themselves as a pathology (e.g., synovial fluid). Therefore there was a need to use a combination of global and local contextual features. Further, the single stage RF used a single classifier to train all the columns of the graph. This did not account for the regionally-specific appearance of the surrounding menisci, muscle, bone and other anatomies. The result of which was that certain intensity profiles appearing in a specific region of the knee that indicated a normal cartilage would only occur for a pathological case in another local region. This discrepancy resulted in the improper training of the classifier which caused an increase in the segmentation inaccuracy.

In this work [31] two RF based classifiers in hierarchy were used to train cartilage regions (Fig. 2.4). The first stage used a neighborhood approximation forests (NAF) [34] trained on example image patches. The second RF classifier [8] collected features along the ELF based geometric graph columns. For the second RF classifier, the local regions of the knee were spatially clustered using k-means clustering to account for the highly variable local anatomies surrounding the cartilage. Each of the spatially clustered regions was trained using a different random forest classifier for the second stage. The output probability maps of the NAF were used with other

image-based features for training each of the regionally specific second RF classifiers. The advantage of this approach is that the NAF classifier gathers contextual and textual information from a larger global neighborhood of 3D image patches while the RF classifier collects local feature information along the geometric search columns of the graph. Disjoint training sets were used to help build a more realistic RF model based on actual NAF performance on unseen images.

2.3.2 Neighborhood Approximation Forests

NAF is a supervised learning method which approximates the neighbors of an unseen test image based on a similarity criterion. The definition of similarity is based on a distance metric that is application specific. NAF uses a training set to learn which clusters of images have the most similar neighborhood structures based on these distances. NAF uses a random forest framework which consists of a collection of several binary decision trees where each tree independently learns to predict the closest neighborhood. Another advantage of NAF is the ability to use a wide variety of distance-based metrics without modifying the core underlying approach.

For each tree in the forest, the training phase starts with the root node and continues to add new branches. Each node branch is trained to learn a set of binary tests which progressively partition the image into subsets with respect to the user-defined distance metric. The node split is optimized such that the user-defined distance function yields the most compact partitioning. For every tree in the forest, a subset of the entire feature set is chosen for training thereby improving generalization and producing an independent prediction.

For our application, the NAF was trained on image patches where the pairwise distance function $\rho(I, J)$ captured the similarity of the image patches based on their segmentation labels. The training image patches was defined as $\rho(I, J) = \| seg(I) - seg(J) \|_{l_0}$ where $seg(.)$ is the segmentation label map for the corresponding image patch. The algorithm learned to group image patches that appeared similar to

each other based on the segmentation similarity.

The unseen test image was partitioned into several smaller image patches. Each patch was passed through a trained NAF classifier. The probability output of all the image patches was combined to produce the final probability map (Fig. 2.3). This was used as one of the inputs for the second clustered RF classifier.

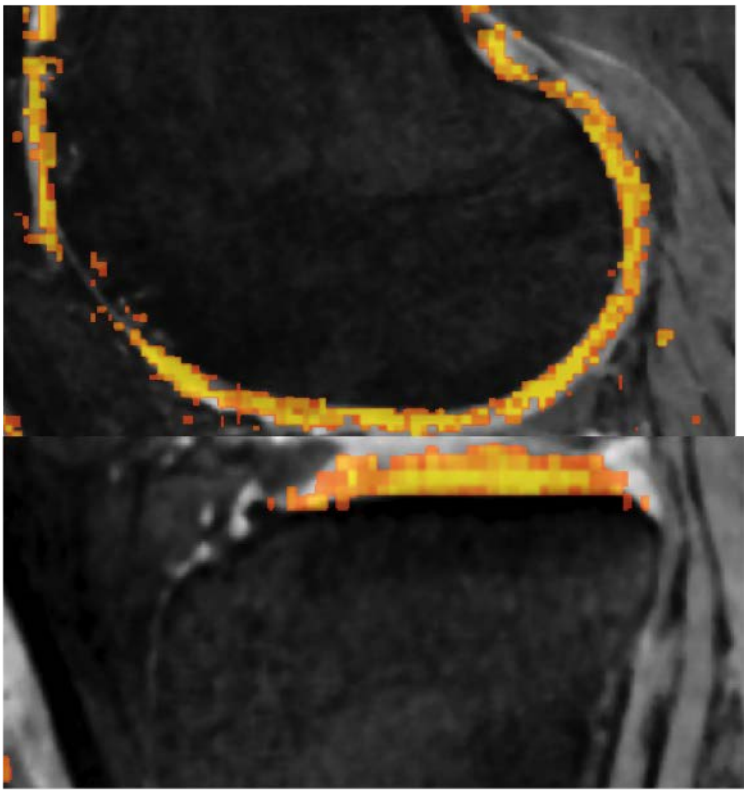


Figure 2.3: The output probability map of the NAF for an unseen image overlaid on the image volume. The color map indicates the probability output values with brighter color indicating higher probability of the voxel being a cartilage region.

2.3.3 Clustered Random Forest Classifier

The second RF was trained on features collected at each node of the geometric graph. Pre-segmented bone mesh surfaces that were corrected using just-enough in-

teraction (See Chapter 5) were used for geometric graph construction during training. Positive example labels corresponded to the nearest cartilage mesh intersection along each graph column. The different features collected at each node point are shown in Table 2.1 with feature values interpolated to the search path points from corresponding feature volumes. To handle the large variability of cartilage intensities in the volume, a k-means clustering algorithm was applied to the S_0 mesh of femur and tibia respectively resulting in spatial parcellation of the pre-segmented mesh surfaces into 40 clusters each (total 80). Each of the clusters was trained using a separate RF classifier that learned the regionally-specific appearance accounting for the surrounding menisci, muscle, bone and other anatomies. The probability response to the features along the search nodes in the testing datasets provided the node costs for graph optimization.

Table 2.1: A list of features used to train the second RF classifier.

Index	**Description**
1–9	3 eigenvalues of Hessian matrices on intensity image at $\sigma = 0.5, 1.0, 2.0$ mm
10–15	1st Gaussian gradient on intensity and NAF probability volumes at $\sigma = 0.36, 0.7, 1.4$ mm
16–18	Intensity, Gaussian smoothed intensity, and NAF probability volumes
19–20	Laplacian derivative of intensity volume at $\sigma = 0.36, 0.7$ mm
21	Gabor texture feature
22–25	Intensity statistics: mean, variance, skewness and kurtosis of a $2\ mm^3$ region centered around each graph node
26–28	Haar features (1.5mm kernel) along horizontal, vertical & diagonal directions
29–30	1D directional gradient along the search column direction on NAF probability and intensity volume

2.4 Experimental Validation

The data used to evaluate our method was provided by the Osteoarthritis Initiative (OAI) which is available for public access[1]. All MR acquisitions were done using a Double Echo Steady State (DESS) pulse sequence with an in-plane resolution of 0.36mm × 0.36mm and a slice thickness of 0.7mm, resulting in 384 × 384 × 160 voxels. 88 datasets with independent standards at baseline (BL) and 12-months follow-up (12M) scans were used (176 3D MRI in total) to validate the accuracy of LOGISMOS 3D segmentation. All MR volumes were from diseased subjects.

Fig. 2.4 shows the learning based segmentation workflow. For bone surface segmentation, the initially employed gradient-based costs were very robust and remained unchanged. For the cartilage classifier, the patient volumes at baseline were divided into two training sets with 15 and 19 patients which were used to train the NAF and the clustered RF classifier respectively. The datasets used for training the clustered RF classifier were first inspected and just-enough interaction edited (See Chapter 5 for details).

Although the training datasets consisted of BL patient volumes only, we eliminated any and all 12M volumes of the same patients from the testing set given the similarity in appearance. The resulting testing set, therefore, consisted of 108 MR volumes (54 MR volumes BL and 12M respectively). In other words, the training and testing sets were distinguished at the patient level. All image volumes were first LOGISMOS segmented using gradient costs. The geometric graph had 8006 and 8002 graph columns for the femur and tibia objects, respectively. The graph parameters are listed in Table 2.4.

The NAF features consisted of image patches sampled over 15 datasets with 1521 sample points per patch. Because of the highly imbalanced ratio between the negative and positive labels, we considered a neighborhood around the cartilage labels and

[1] `https://oai.epi-ucsf.org/datarelease/`

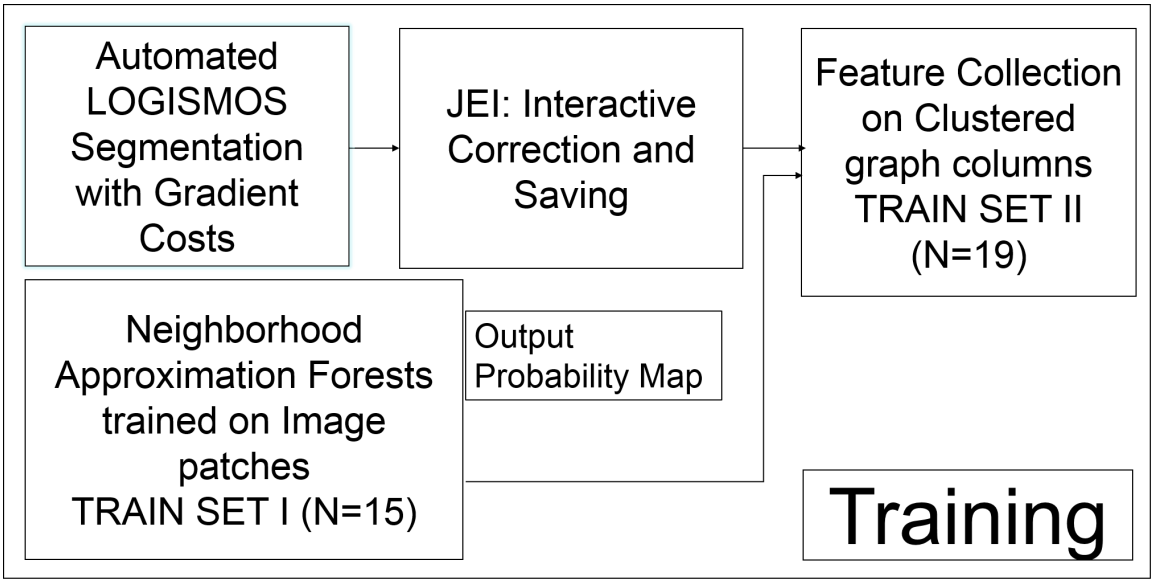


Figure 2.4: The learning-based segmentation algorithm work-flow for training the hierarchical random forests classifiers.

marked them as negative examples. The collected image patches consisted of all the positive and the surrounding negative labels. We trained a set of 200 trees with 40,000 images patches as inputs to each tree. The second RF classifier was trained on 19 JEI-corrected datasets with 30 features (see Table 2.1) along with the ELF search path for each node. 80 (40×2) RF classifiers were trained with each one representing the given cluster with 800 trees per forest.

We also trained a second RF classifier without any features from the NAF while keeping the same graph parameters as the hierarchical classifier system. The classifier was trained on the same 19 patient datasets. The motivation behind the single stage classifier was to evaluate the benefits of adding the NAF stage to the classifier system. The validation workflow is shown in Fig. 2.4.

Table 2.2: Parameters used for graph construction. Minimum inter-surface and inter-object separations are zero.

	Inter-surface max (nodes)	Inter-object max (nodes)	Smoothness (nodes)	Column size (nodes)
Learned cost	40	120	4	121
Gradient cost	20	60	2	61

2.5 Results

Surface positioning errors (compared against independent standard) achieved by the hierarchical classifier, gradient cost, and single stage RF classifier are listed in Tables 2.5, 2.5. Each resulting surface from the above methods were sub-plate analyzed for studying the thickness of specific sub-regions where the effects of the disease were more pronounced (see Chapter 3 for sub-plate analysis algorithm details). Table 2.5 shows a significant reduction in signed and unsigned errors for all the sub-plates of the femur ($p \ll 0.001$) over the single stage RF classifier and the gradient based costs. The tibial sub-plates in Table 2.5 showed a significant reduction in unsigned errors over the gradient costs for all the sub-plates. The signed errors showed a significant reduction over the gradient based costs on all sub-plates except the central Medial tibia (cMT). Comparison with the single stage RF classifier saw significant improvement in unsigned errors over the overall tibia and central medial tibia (cMT) plate while the errors in the central lateral tibia (cLT) were not significant.

Fig. 2.6 qualitatively compares the segmentation accuracies between the gradient based costs and the hierarchal classifier with respect to the independent standard. Both the femur and tibia are shown with their respective bone and cartilage segmentations showing good agreement between learning-based segmentation and the independent standard.

Table 2.3: Cartilage border positioning errors (in mm) for femoral sub-plates achieved by hierarchical classifier, gradient cost and single stage RF classifier.

	NAF+RF	Gradient	p-value	RF only	p-value
Femur signed	**-0.01±0.18**	-0.31±0.24	$\ll 0.001$	-0.10±0.17	$\ll 0.001$
Femur unsigned	**0.55±0.11**	0.69±0.13	$\ll 0.001$	0.56±0.10	$\ll 0.001$
cMF signed	**-0.04±0.29**	-0.38±0.58	$\ll 0.001$	-0.11±0.27	$\ll 0.001$
cMF unsigned	**0.52±0.16**	0.78±0.35	$\ll 0.001$	0.55±0.17	$\ll 0.001$
cLF signed	**-0.26±0.24**	-0.52±0.35	$\ll 0.001$	-0.36±0.20	$\ll 0.001$
cLF unsigned	**0.42±0.12**	0.65±0.20	$\ll 0.001$	0.47±0.11	$\ll 0.001$

Table 2.4: Cartilage border positioning errors (in mm) for tibial sub-plates achieved by hierarchical classifier, gradient cost and single stage RF classifier.

	NAF+RF	Gradient	p-value	RF only	p-value
Tibia signed	**0.06±0.17**	-0.11±0.35	$\ll 0.001$	0.11±0.22	$\ll 0.001$
Tibia unsigned	**0.60±0.14**	0.79±0.20	$\ll 0.001$	0.62±0.18	$\ll 0.001$
cMT signed	-0.15±0.31	-0.25±0.77	0.193	**-0.09±0.34**	0.003
cMT unsigned	**0.52±0.20**	0.92±0.41	$\ll 0.001$	0.58±0.22	$\ll 0.001$
cLT signed	**-0.03±0.32**	0.36±1.12	$\ll 0.001$	-0.01±0.32	0.16
cLT unsigned	**0.46±0.17**	0.79±0.99	$\ll 0.001$	0.47±0.18	0.31

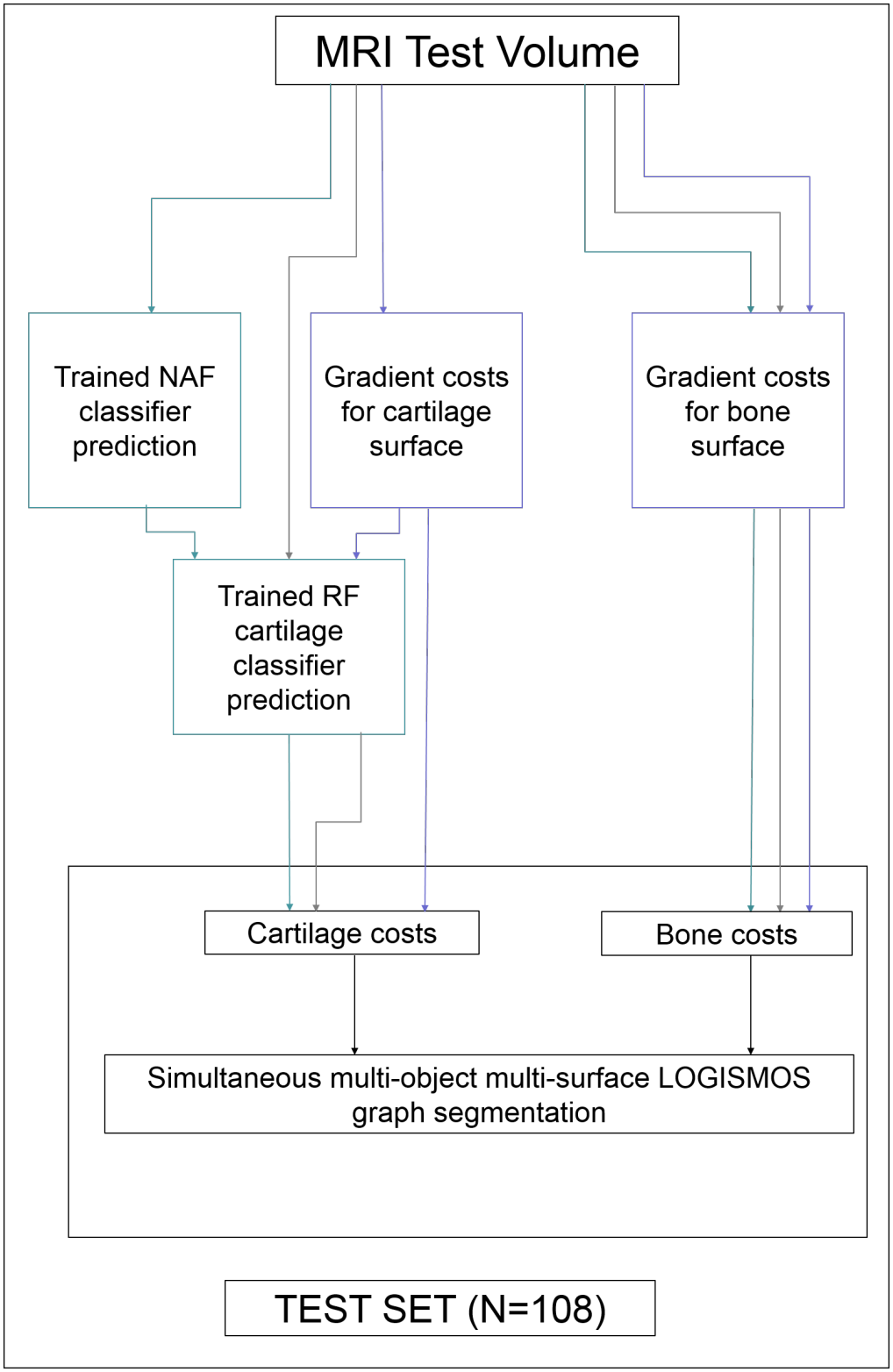


Figure 2.5: The testing workflow to compare the clustered random forest classifiers with the existing methods on a set of 108 patients.

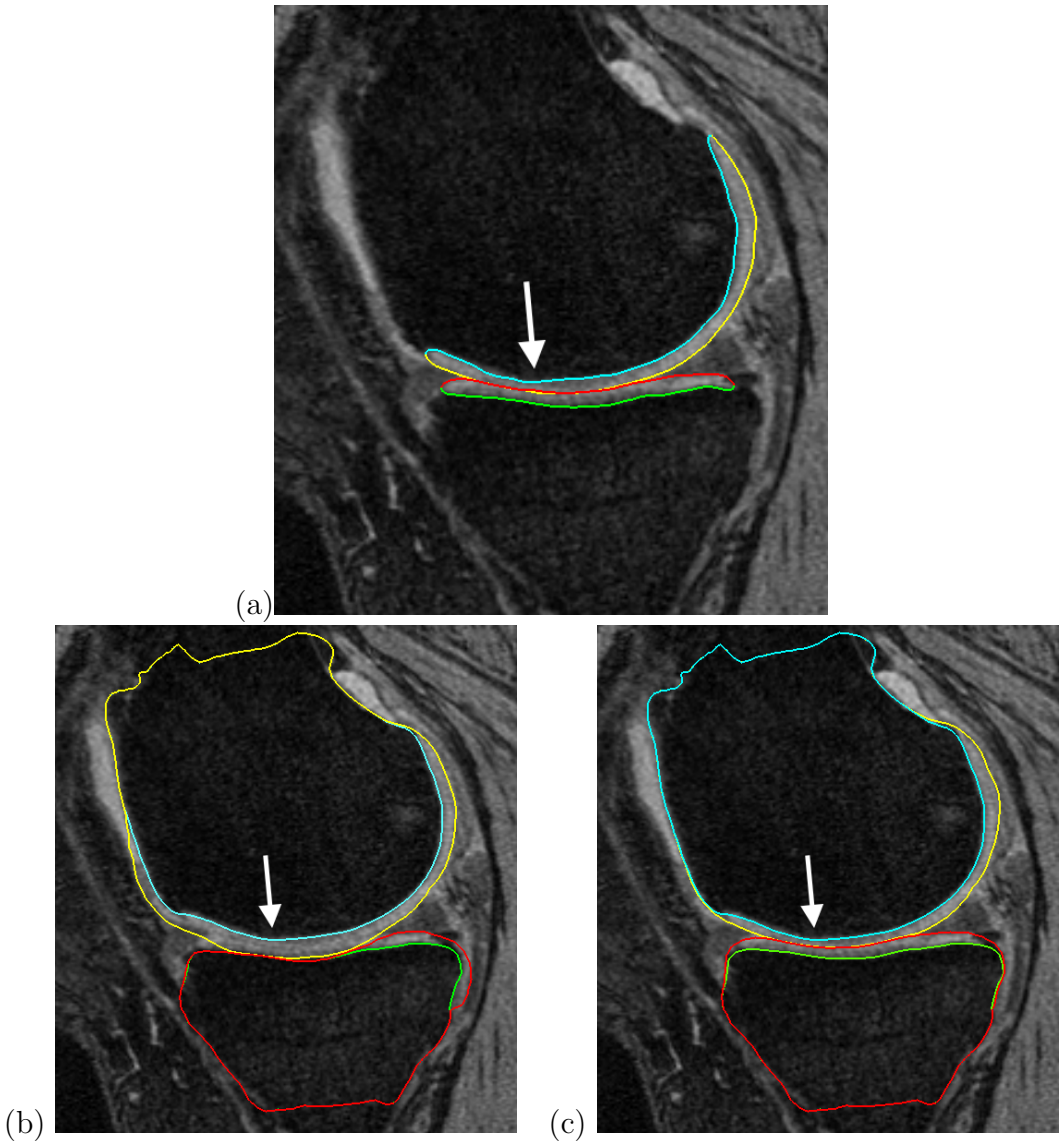


Figure 2.6: Segmentation accuracy in a representative subject. (a) Independent standard. (b) Gradient-costs LOGISMOS segmentation. (c) Learned-costs LOGISMOS segmentation. Region marked by the arrow shows clear improvement in the segmentation quality when using the learned costs.

CHAPTER 3
AUTOMATED ANALYSIS OF SMALLER LOAD BEARING REGIONS TO QUANTIFY THICKNESS LOSSES

Thickness analysis of cartilage is crucial for quantifying osteoarthritis. Analyzing the whole cartilage structure drowns out the thickness losses that occur in the specific regions of the knee joints. These sub-regions have been identified as areas which bear the maximum stresses during motion with a common consensus by the osteoarthritis research community on what regions of the femur and tibia that need to be analyzed in isolation [16]. Several techniques [67, 68] for identifying the sub-plates exist all of which require human interaction to provide an initialization for the sub-region analysis.

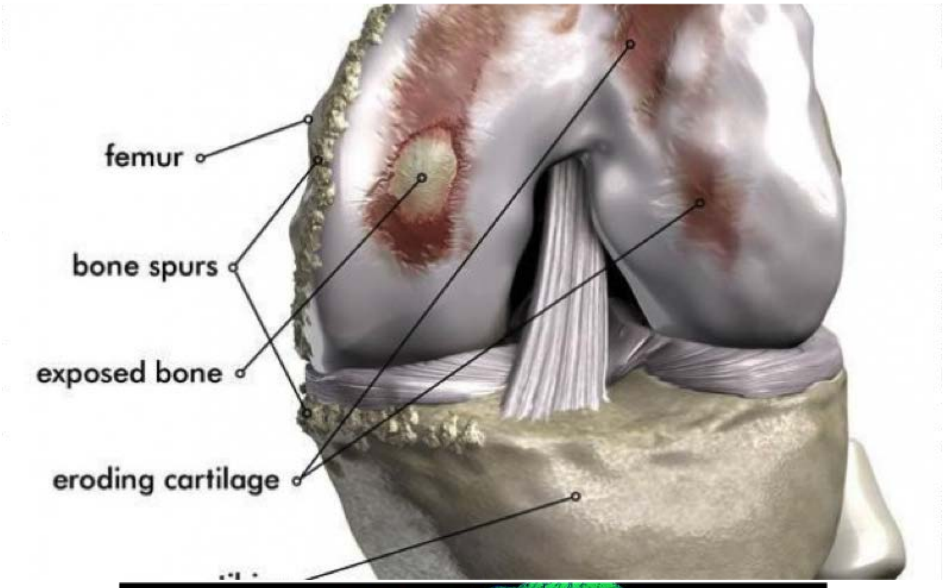
In this chapter a fully automated sub-plates detection algorithm is detailed. The algorithm requires as input the segmented bones and cartilages of the femur and tibia (see Fig. 3.1 for a representative segmented surface). The sub-plate detection algorithm is run on the resulting LOGISMOS segmented surface meshes which automatically extracts and analyzes the nomenclature compliant sub-regions providing a fast reliable technique for studying focal thickness losses in the load bearing regions of the knee.

3.1 Trochlear Notch Identification

The first step in the sub-plate identification is the detection of the trochlear notch. The notch is at the base of a groove along which the patella (knee cap) slides over the femur providing for a smooth contact surface (Fig. 3.2). It is important to identify the correct positioning of the notch because all the cartilage analysis on the femoral condyle is done posterior to this plane since the image is known to suffer from partial volume effects anterior to it.

The main anatomic feature we exploit is anterior to posterior (AP) curvature of the groove of the femoral bone (Fig. 3.2b). The notch is at the base of the curvature before

a.



b.

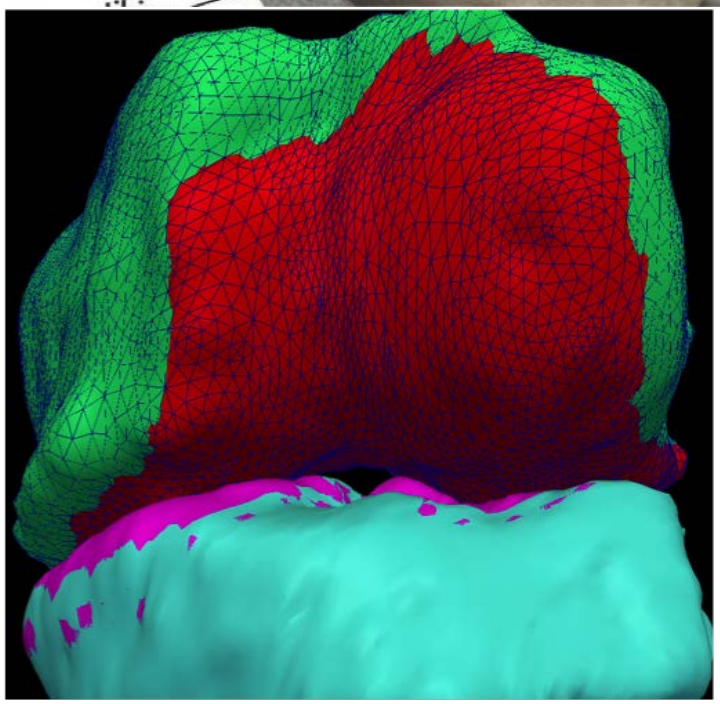


Figure 3.1: a. Shows a real osteoarthritic knee with the eroding cartilage exposing the bone and the growth of osteophtyes on the outer edge of the femur and tibia *Source: ConforMIS via BoneSmart.org.* b. Shows a mesh based representation of the LOGISMOS segmented knee.

the bone ridge structure rises sharply. An implicit cutting plane is used to isolate the analysis region and is represented by $\hat{n}.p$ where $\hat{n}$ is the normal direction from the plane and p is a point on the plane. On isolating the groove regions using simple implicit plane geometry cutting, we draw a family of contours along this surface. The sharp rise of the bone structure along the contour near the base of the ridge gives a large change in contour positioning value traversing in the AP direction. To increase robustness we find the positions of largest change on all of the closely positioned groove contour lines and average them to find the desired trochlear notch (Fig. 3.3).

3.2 Sub-Plate Detection using Implicit Cutting Plane Geometry

Using an implicit plane normal to the AP plane at the trochlear notch separates the posterior regions. From the same position, using another implicit plane with its normal parallel to the AP direction separates the posterior region into the medial and lateral condyle. For each femoral condyle, the load bearing regions is defined as 60 % of the distance in the AP direction from the trochlear notch to the posterior-most point of the respective condyles. Isolating them using cutting planes normal to the AP plane at the 60 % region isolates the load bearing regions of the central medial and lateral femur (Fig. 3.2).

The tibia is divided into the medial and lateral compartments using cutting planes positioned at the trochlear notch with its normal perpendicular to the AP plane Fig. 3.4. After subdividing the plates into the medial and lateral regions, we isolate the central tibia and the peripheral sub-regions. The central 20% elliptical area of the medial and the lateral plate is computed around the center of mass of each respective plate. The radius of the major and the minor axis is computed as a ratio of bounds of the respective medial/lateral compartments. The major axis radius is computed from the compartment bounds along the AP direction and similarly, the minor axis radius is computed from the ratio of the compartment bounds perpendicular to the AP direction. Furthermore, the remaining regions are isolated using a 45° and 135°

a.

b.

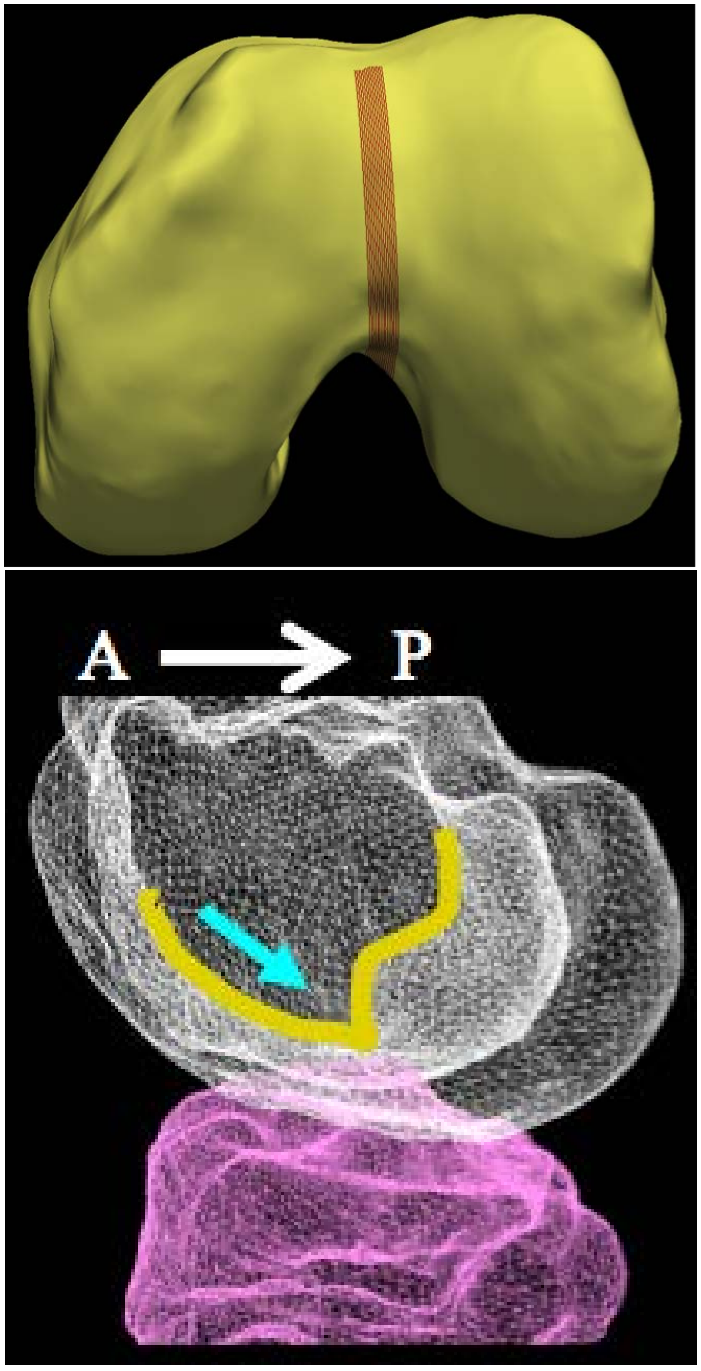


Figure 3.2: a. Example LOGISMOS segmented bone mesh with the trochlear notch region highlighted. b. Bone mesh highlighting the trochlear groove (overlayed in yellow) in the anterior to posterior direction. The blue arrow indicates the trochlear notch on the groove at the base of the groove curvature before the bone ridge structure rises sharply.

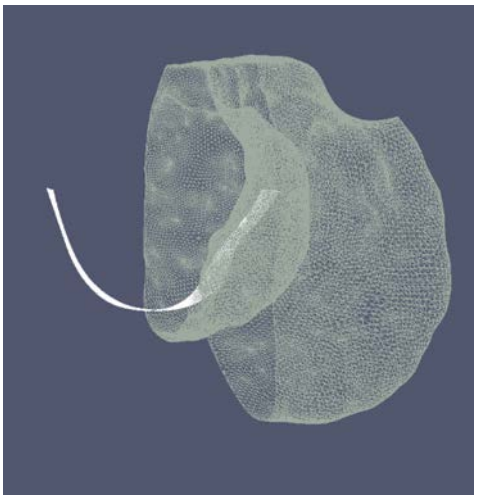
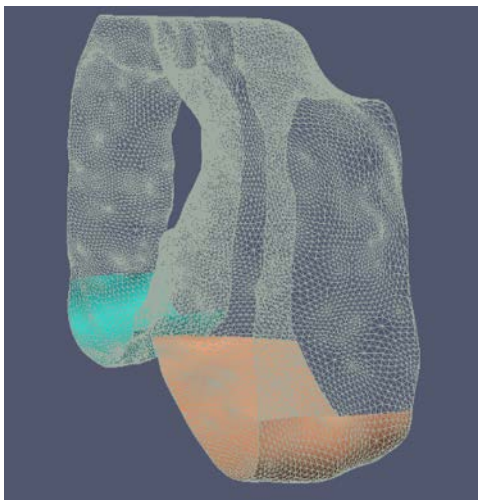
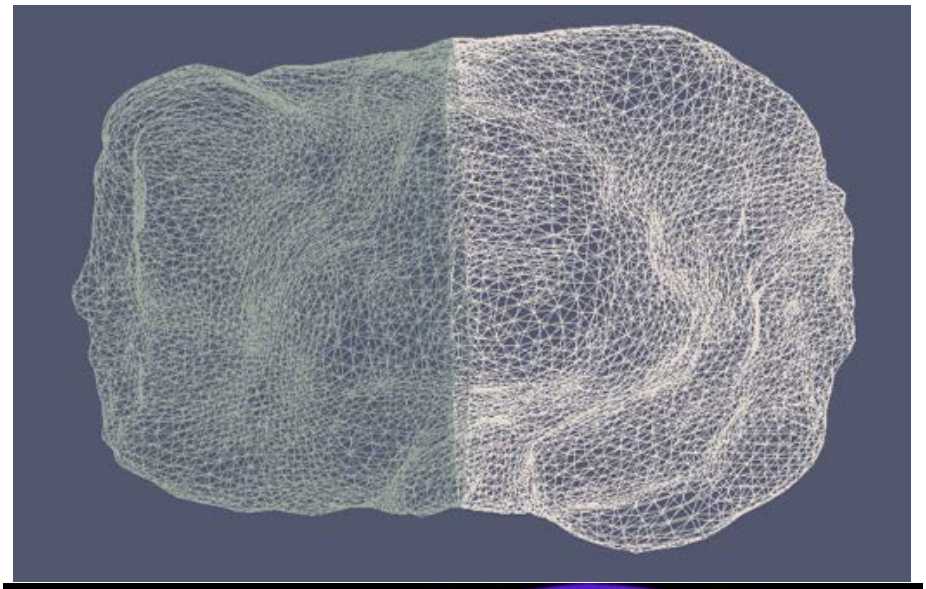
a.  b. 

Figure 3.3: *a.* Trochlear notch Identification exploits the AP curvature of the groove on the femoral bone. After isolating the groove region, the trochlear notch is identified by taking an average gradient on a family of closely positioned contour lines along this surface. *b.* The load bearing regions of the femur identified by isolating 60 % of the distance from the trochlear notch to posterior most in the AP direction on each condyle respectively.

cutting planes around the center of mass of their respective compartments to give all four peripheral sub-regions Fig. 3.4. The final list of all the extracted sub-plates from the automated sub-plate detection algorithm is visually shown in Fig. 3.5.

For most of our analysis throughout the thesis we focus on reporting the thickness losses in the four sub-regions, i.e. the medial/lateral femur and tibia.

a.



b.

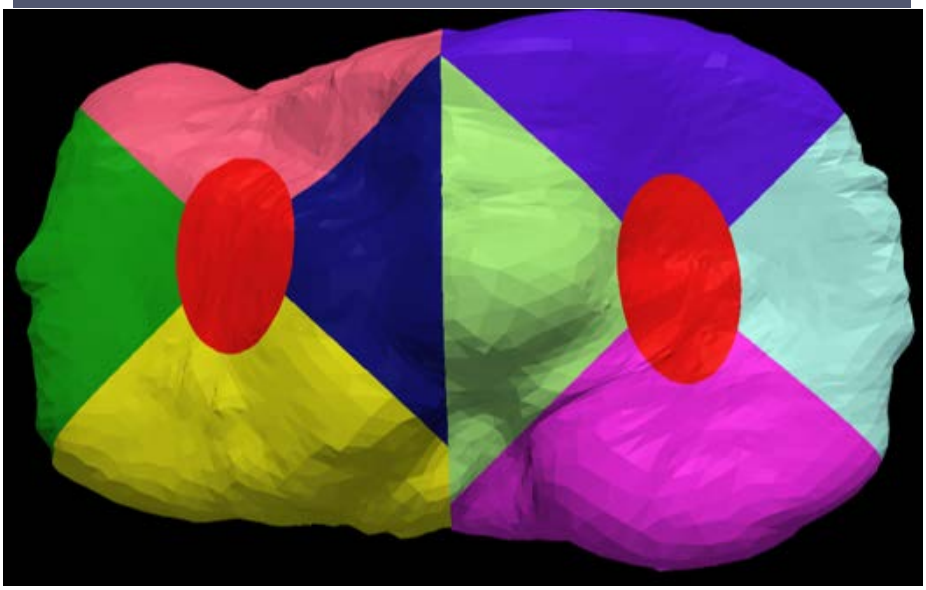


Figure 3.4: *a.* Using the trochlear notch positioning and implicit cutting planes parallel to the AP direction the medial and the lateral tibia can be isolated. *b.* The central 20% elliptical region of the tibial plates with the major and minor axis computed using the bounds of the separated plates regions. Further using 45° and 135° cutting planes around the center of mass isolates the peripheral regions of the medial and lateral plates respectively.

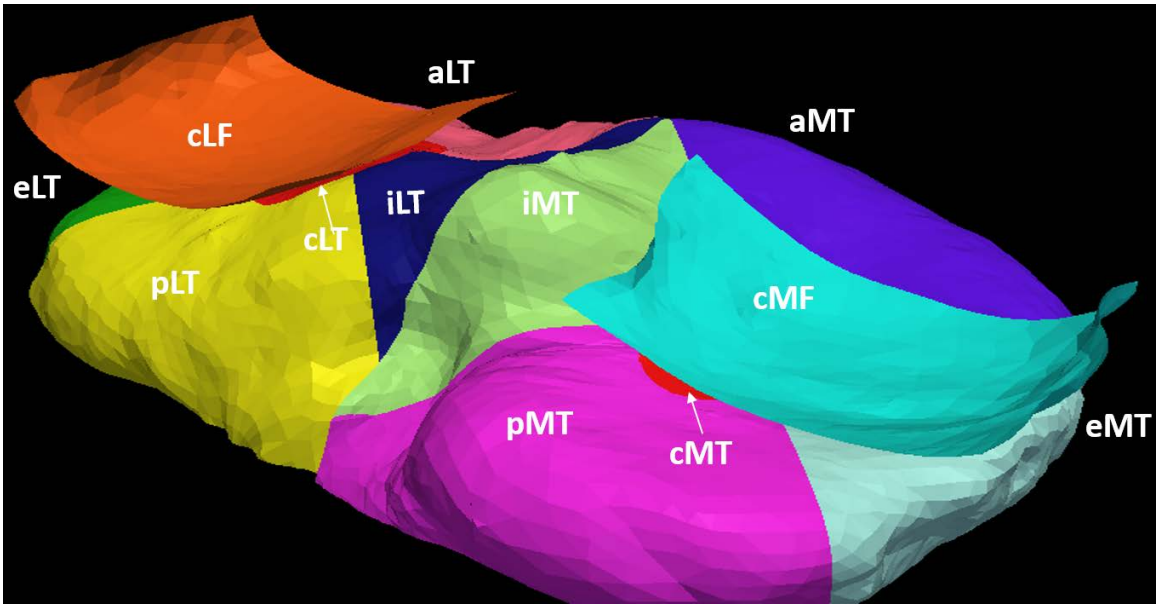


Figure 3.5: Automated sub-plate division of the cartilage. The 60% central lateral (cLF) and central medial femur (cMF) are shown. Each Medial (MT) and lateral tibia (LT) regions are subdivided as: central (cLT/cMT), interior (iLT/iMT), exterior (eLT/eMT), anterior (aLT/aMT) and posterior (pLT/pMT) regions respectively.

CHAPTER 4
4D LONGITUDINAL SEGMENTATION OF KNEE MRI

4.1 Motivation

In Chapter 2, the cost function design and the 3D LOGISMOS segmentation algorithm for knee MRI were explained. All patients in the OAI were enrolled in a longitudinal study with follow-up MRI scans at regular time point intervals. The availability of multi-time point information can be leveraged to provide additional information. This becomes particularly beneficial in patients with progressively worsening osteoarthritis. With no current FDA approved disease modifying OA drug all of these patients eventually undergo total knee replacement. With disease progression, artifacts such as cartilage and bone marrow lesions, synovitis, cartilage thinning appear that have a similar appearance to cartilage making it a challenging task even for expert radiologists. Using information of earlier time points of the same patient when the knee is relatively healthier and with lesser artifacts helps to add extra contextual information thereby reducing the inter-time point variability ensuring that the cartilage losses are within physiologically possible ranges. To the best of our knowledge, this is the first attempt at a 4D knee MRI segmentation algorithm encoding the similarity information in a graph search framework resulting in reduced inter-time point variability and improved segmentation accuracy.

In order to perform 4D segmentation, the crucial first step is to register the pre-segmented mesh surfaces and the corresponding image volumes across time-points to ensure that the 4D constraints establish correspondences temporally between the similar regions of the knee. Further, given that the bone surface is fairly rigid (only slight changes in bone shape were reported for patients with osteoarthritis with time) we use rigid registration to spatially similar region correspondences.

Across all time-points the same configuration is maintained in terms of the geometric graph parameters and topology, enforcing contextual information temporally

is done by drawing inter-time point arcs. Both the registration and the graph arc construction are detailed in the sections below.

4.2 Establishing Point Correspondences

After pre-segmentation of each of the time-points being analyzed, iterative closest points (ICP) algorithm [5] was used to register the pre-segmented mesh surfaces to enforce inter-time point constraints. The algorithm iteratively registers meshes as point clouds until convergence such that the points from one set are closest to the other in the least squares sense. Because the least squares optimization has a tendency to get stuck at local minimum a two-step approach was employed. With a large translational or rotational movement between the two-time points, there is a tendency to mismatch the surfaces, i.e. femur matched to the tibia or vice versa if they were the closest in terms of the movement.

To prevent the mismatch, the first step used the femur mesh for ICP registration. The transform matrix was applied on both the femur and tibia meshes. Given that they were almost close to each other the ICP registration was run again using both the femur and tibia points together as a single point cloud. This ensured that vertex to vertex correspondences and thereby column correspondences were established. Fig. 4.1 illustrates the establishment of correspondence between two-time points. The same two step transformation matrices were applied to the image volume to ensure that both the pre-segmented meshes and the volume were registered to a common space.

4.3 Incorporating Inter-Time Point Context

For the 4D segmentation, the datasets at each time point have their respective 3D multi-surface multi-object constraints explained in the chapter 2. To constrain across time points we use inter-time point arcs between the nodes of the columns connecting the registered bones & cartilage surfaces at different time points with each

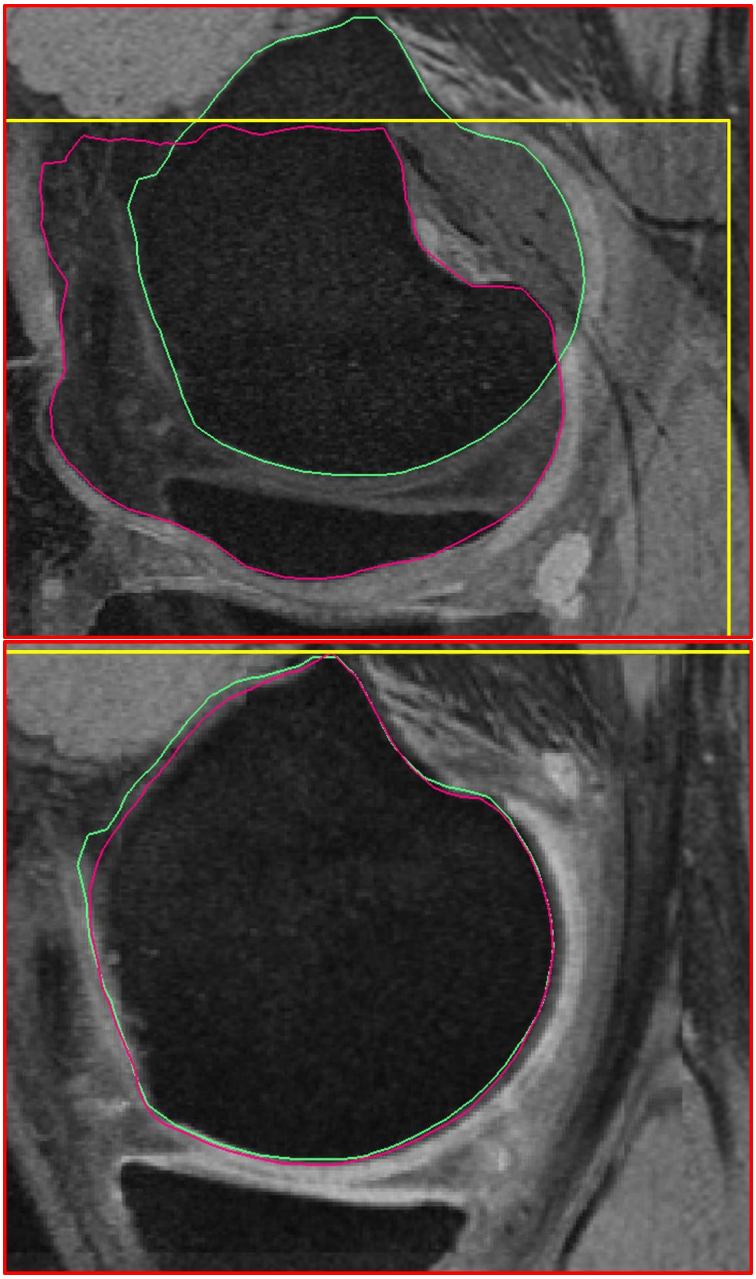


Figure 4.1: Establishing point correspondences between the time-points after pre-segmentation. The meshes are registered using ICP to establish column correspondences. The same transformation is applied on the corresponding volumes. (a) shows the two time-points before registration. The meshes are colored differently and the volume border is highlighted in yellow to indicate its positioning. After ICP, (b) shows the registered meshes and the volume with border positioning highlighted.

other. Using the prior information we can set the different minimum and maximum permissible limits for the bones and cartilages respectively so that the undergoing losses stay within a physiologically feasible range. The arcs introduced define a $(\delta_{timeMin}, \delta_{timeMax})$ where the inter-time-point maximum allowed change $(\delta_{timeMax})$ was set to 0.6mm based on available clinical literature [17].

Extending the terminology used earlier to define the different arc constraints, the inter-time-point arcs between two longitudinally corresponding columns t_1, t_2 are defined as:

$$\begin{aligned} E^{inter\_time\_points} = &\{\langle V_i(t_1, m, n_{cart}, k), V_i(t_2, m, n_{cart}, k - \delta_{timeMin})\rangle\} \\ &\cup\{\langle V_i(t_1, m, n_{cart}, k - \delta_{timeMax}), V_i(t_2, m, n_{cart}, k)\rangle\} \\ &\cup\{\langle V_i(t_1, m, n_{bone}, k), V_i(t_2, m, n_{bone}, k - \delta_{timeMin})\rangle\} \\ &\cup\{\langle V_i(t_1, m, n_{bone}, k - \delta_{timeMax}), V_i(t_2, m, n_{bone}, k)\rangle\} \end{aligned}$$

After the graph edges and nodes have been constructed, the 4D LOGISMOS segmentation problem is solved by computing a minimum $s - t$ cut in a derived arc weighted digraph.

4.4 Experimental Validation

88 datasets with baseline (BL) and 12-months follow-up (12M) scans from the OAI were used (176 3D MRI in total). Independent standards released by the OAI were available for all the datasets. The bone surface segmentation used the initially employed gradient based costs. For the cartilage the hierarchical random forests classifier proposed in chapter 2 was used. The graph parameters used are shown in Table. 2.4 with the additional inter-time-point constraint set at 0.6 mm.

Given that OA is a slowly progressing disease, the expectation was that there would not be a significant appearance of artifacts of thickness losses for the available

set of 108 datasets over a one year study period where the true benefits of using 4D can be appreciated. To validate the benefits of 4D over a longer study period, 399 patients with BL, 12M, 24M, 36M and 48M follow-up scans from the OAI were used resulting in a total of 1995 3D MRI. All patients were 3D LOGISMOS segmented and simultaneous five time-point 4D LOGISMOS segmented followed by sub-plate analysis. Assuming that the loss of thickness due to OA was almost linear over time, we hypothesized that the correlation coefficient R would have a higher number of datasets close to 1.0 for 4D indicating that the additional contextual information helps improve segmentation accuracy over all the time-points.

4.5 Results

Tables 4.5, 4.5 compare signed and unsigned border positioning errors w.r.t to the independent standard for 3D and 4D LOGISMOS segmentation at BL and 12M follow-up time points respectively. As expected the overall segmentation accuracy is almost the same between the two methods except for a few of the plate regions highlighted in bold showing statistically significant differences over 3D (paired t-test comparison with p=0.05). Fig. 4.2 compared the histogram of correlation coefficient (R) values for the different sub plates from 4D and 3D segmented results. We see that for 4D the R value has a larger number of patients towards the right of the charts especially for the medial femur (MF) and medial tibia (MT) indicating more accurate representation of the thickness losses. Further Table 4.3 shows that the paired t-tests between the R values of 3D versus 4D for all the sub-plates showed statistically significant differences.

Fig. 4.3 qualitatively shows the improvement of 4D LOGISMOS over 3D. Note that there is no obvious edge distinguishing the tibia and femur cartilage. Using temporal context from the other time points we see that the proposed method is able to correctly position the cartilage for the tibia and femur in 4D.

Table 4.1: Cartilage surface positioning errors (in mm) of 4D versus 3D LOGISMOS at Baseline. Paired t-test significance value was set at $p = 0.05$.

	4D	3D	p-value
Femur signed	**0.01±0.18**	0.01±0.19	0.027
Femur unsigned	**0.53±0.11**	0.54±0.11	$\ll 0.001$
cMF signed	**-0.01±0.25**	-0.03±0.27	0.015
cMF unsigned	0.51±0.16	0.52±0.17	0.092
cLF signed	-0.26±0.22	-0.25±0.23	0.108
cLF unsigned	0.42±0.10	0.42±0.10	0.790
Tibia signed	0.08±0.17	**0.07±0.16**	0.039
Tibia unsigned	0.59±0.14	0.60±0.14	0.140
cMT signed	-0.13±0.28	-0.14±0.29	0.125
cMT unsigned	**0.50±0.18**	0.51±0.20	0.039
cLT signed	0.00±0.30	0.00±0.31	0.338
cLT unsigned	**0.45±0.17**	0.46±0.18	$\ll 0.001$

Table 4.2: Cartilage surface positioning errors (in mm) of 4D versus 3D LOGISMOS at 12Month. Paired t-test significance value was set at $p = 0.05$.

	4D	3D	p-value
Femur signed	-0.02±0.17	-0.04±0.17	0.062
Femur unsigned	0.55±0.11	0.55±0.11	0.601
cMF signed	**0.02±0.27**	-0.04±0.30	0.016
cMF unsigned	0.52±0.15	0.53±0.16	0.614
cLF signed	-0.30±0.21	-0.30±0.23	0.853
cLF unsigned	0.43±0.12	0.44±0.13	0.318
Tibia signed	0.07±0.19	0.06±0.18	0.464
Tibia unsigned	0.60±0.16	0.60±0.14	0.310
cMT signed	-0.14±0.29	-0.15±0.31	0.619
cMT unsigned	0.52±0.17	0.53±0.17	0.341
cLT signed	-0.06±0.31	-0.06±0.30	0.629
cLT unsigned	0.45±0.16	0.44±0.16	0.256

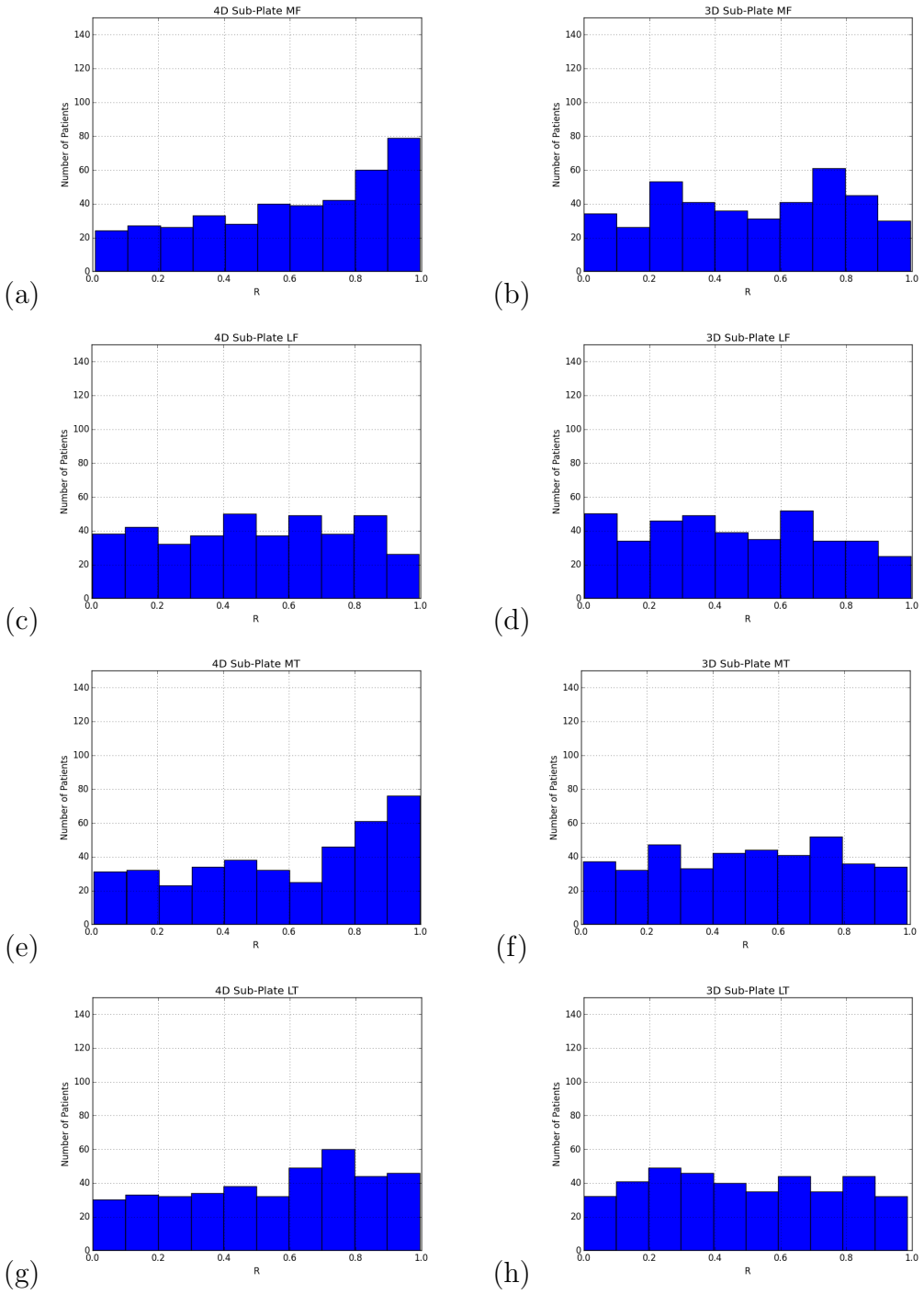


Figure 4.2: R value histogram of the different sub plates in 4D (*left column*) and 3D (*right column*) for medial & lateral femur (MF, LF), and medial & lateral tibia (MT, LT). We see that the 4D based correlation coefficient has higher number of patients with a R value closer to one indicating agreement with the linear thickness loss over time points.

Table 4.3: Paired t-tests on the R values computed from 3D versus 4D LOGISMOS segmentation. Every plate showed statistically significant differences indicating that the 4D segmentation does help improve the accuracy over multiple time-points and are able to capture the thickness losses better.

R-value 4D vs 3D	MF	LF	MT	LT
p-value	<0.001	0.0113	<0.001	<0.001

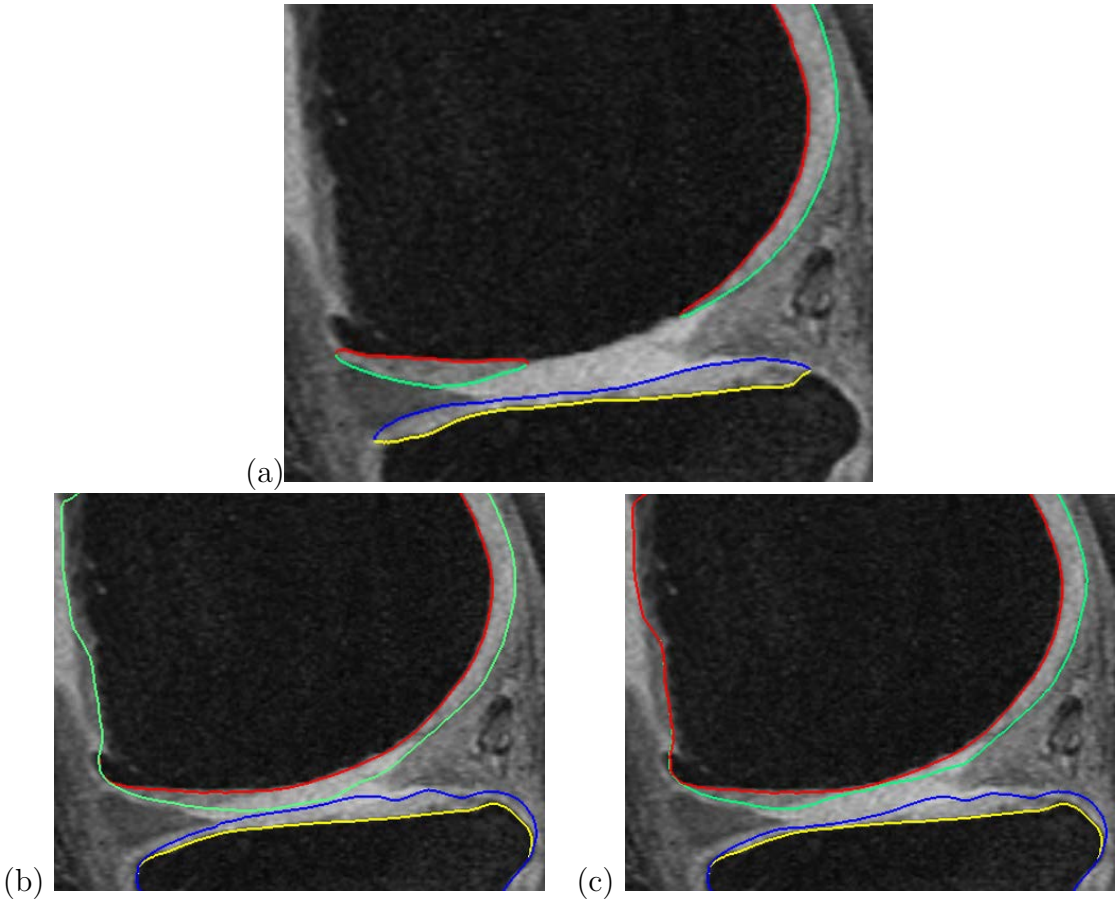


Figure 4.3: Qualitative improvement of 4D segmentation versus 3D. *(a)* Ground Truth Segmentation, *(b)* 3D segmentation, *(c)* Proposed 4D segmentation clearly shows limited ability of the 3D approach while the 4D approach benefited from the temporal context and accurately delineated the cartilages of the femur and tibia.

CHAPTER 5
JUST ENOUGH INTERACTION FOR KNEE MRI

5.1 Motivation

Automated LOGISMOS segmentation is shown to be robust for most of the cases of osteoarthritic knee MRI segmentation. However given today's state of the art no fully automated algorithm is accurate 100 % of the time. Interactive correction methods are designed to help ease the post-processing needed. Several techniques have been proposed in the literature and were reviewed in the introduction. Several of the methods correct for segmentation inaccuracies by directly matching the object boundaries with the interaction which after several iterations result in the surfaces having local topological errors.

The proposed interaction algorithm uses a graph based LOGISMOS framework [40, 70] with the user clicked points hereafter called *nudge points* interacting directly with the underlying graph framework. This method has guarantees of global optimality for every interaction and differs from the traditional voxel-by-voxel editing by requiring just enough (i.e., limited) interaction (JEI) to correct the original automated segmentation if needed. The proposed method may appear similar to the Boykov's graph cut [7], however, their interaction algorithm is not able to guarantee global optimality for multiple surfaces and objects. LOGISMOS-JEI always guarantees global optimality when handling multiple objects and surfaces. The JEI architecture and GUI are designed to be platform and application agnostic and their details are covered in [71].

The JEI method starts with an initial automated LOGISMOS segmentation. Post segmentation, the resulting optimized graph state (called residual graph) is saved for the purposes of JEI. The JEI algorithm was extended from Sun et al. [63] to handle multiple surface multiple object interactions. Further, a new interaction mechanism was developed along with a faster graph optimization library to provide almost imme-

diate feedback on the interaction. The overall JEI workflow is very similar for both the 3D and 4D LOGISMOS-JEI with a few modifications to handle the 4D.

5.2 JEI Work-flow

The electric lines of force (ELF) based geometric graph, image volume and residual graph were loaded into a custom designed GUI (see Fig. 5.1) to inspect segmentation quality and perform JEI. The GUI is designed as a modular system [71] where each module performs a specialized set of atomic tasks independently without having to know the details of the other module. All inter-module communication and data exchange were performed by the Boost.Signals2 library *signal and slot* mechanism. Building on this modular GUI system we have designed knee MRI specific modules for the handling of MRI volume and surfaces, interpreting the nudge points and subsequently changing the cost functions in the appropriate columns to correct the segmentation surfaces. Details of the workflow are as follows:

5.2.1 User provided nudge points

The user identified correction is provided as a set of nudge points which guide the segmentation to the correct position. Fig. 5.3a shows the GUI magnified with the volume and the automated LOGISMOS segmentation results overlayed. The particular slice indicated is a case with severe OA having bright fluid regions improperly segmented as cartilage. The blue line indicates interpolated nudge points indicated by the user approximately identifying the correct cartilage region.

5.2.2 3D local graph cost modification

To identify the underlying graph columns influenced by the nudge points (defined as a contour), a k-dimensional (kD) tree algorithm is used which stores all the geometric graphs positions. In previous JEI applications [63] the graph was constructed on a regular 3D grid where the nearest graph columns could be identified quite easily.

However given the complex shape of the knee objects and the ELF graph constructed based on it, a more sophisticated query of the closest columns is needed which does not compromise on speed. The kD tree allows for a $O(log\ n)$ query on the N nearest graph nodes (empirically determined) for every nudge point. Once identified, the costs (i.e. unlikeliness) at corresponding columns associated with these nodes are modified as

$$c(i,j) = \begin{cases} 0, & \text{if } D((i,j), n(i,j)) < \Delta \\ 1, & \text{otherwise} \end{cases},$$

with $c(i,j)$ defined as the cost of node j on column i, $D((i,j), n(i,j))$ the distance between node closest to the nudge point (i,j) and its nearby intersecting nodes $n(i,j)$ within the Δ tolerance.

5.2.3 Max-flow re-computation

Following the local graph cost modification the max-flow is recomputed in 3D within a few milliseconds since a small number of graph columns have been modified. The updated surfaces are rendered back onto the GUI. As seen in Fig. 5.3b the correction made is reflected in the updated cartilage surfaces.

The above workflow is repeated to correct the tibial cartilage errors as well. In the intermediate steps following the correction of the femur, the tibia bone and cartilage surfaces appear to worsen. This can be attributed to a combination of the existing graph costs and their respective local constraints. Since the tibia cartilage surface has no clear defined edge cost in that region, the surface result moved along with the femur corrected cartilage surface. However, once the nudge points provided the appropriate locations for cost modification the erroneous surfaces were corrected (Figs. 5.3c,d). Note that the corrections made on a single 2D slice resulted in the entire locally affected 3D neighborhood being corrected. This can be appreciated in the corresponding circled regions of the surface model.

5.2.4 Undo-Redo Interaction Capabilities

For a more consistent segmentation interaction, we save the user interaction steps during the JEI. A stack was used to save the inputted nudge points and the surface ID at each editing step. The editing was reverted by popping the stack which restored the previous costs on the local graph columns. Re-optimizing the graph resulted in the previous surfaces. However, the popped stack was not deleted unless a different interaction was continued. If the interaction needed to be redone, the pointer simply moved to the previous position on the stack and repeated the same steps as above to redo the correction.

When a new automated LOGISMOS surface is loaded that was previously edited the user can load the editing stack to bring the interaction to most up-to-date edited state. The edits can be continued from that point. We anticipate that this feature would be very useful in reducing the inter-observer and intra-observer variability.

5.3 Longitudinal JEI

For the 4D LOGISMOS-JEI, the GUI was extended to enable visualization and interaction of all patient time-points simultaneously. The viewer also enabled synchronized scrolling across datasets. Fig. 5.2a shows eight time-points of the same patient (baseline, 12, 18, 24, 36, 48, 72 and 96 month follow-ups) being simultaneously visualized. Each individual thumbnail view can be expanded (see Fig. 5.2b) to a detailed larger GUI (identical to 3D GUI) for interaction.

Once the user decides on which particular time-point and a slice that required editing the thumbnail view is expanded to provide a full editing capable GUI (Fig. 5.2b). The interaction mechanism is similar to the 3D JEI where a set of nudge points on a single 2D slice modifies the graph node costs in the local 3D neighborhood columns of the given time-point. Further, since the longitudinal JEI has a single large underlying residual graph with temporal inter-time-point constraints the corresponding local 3D neighborhood column locations at the other time-points are also corrected. To recall,

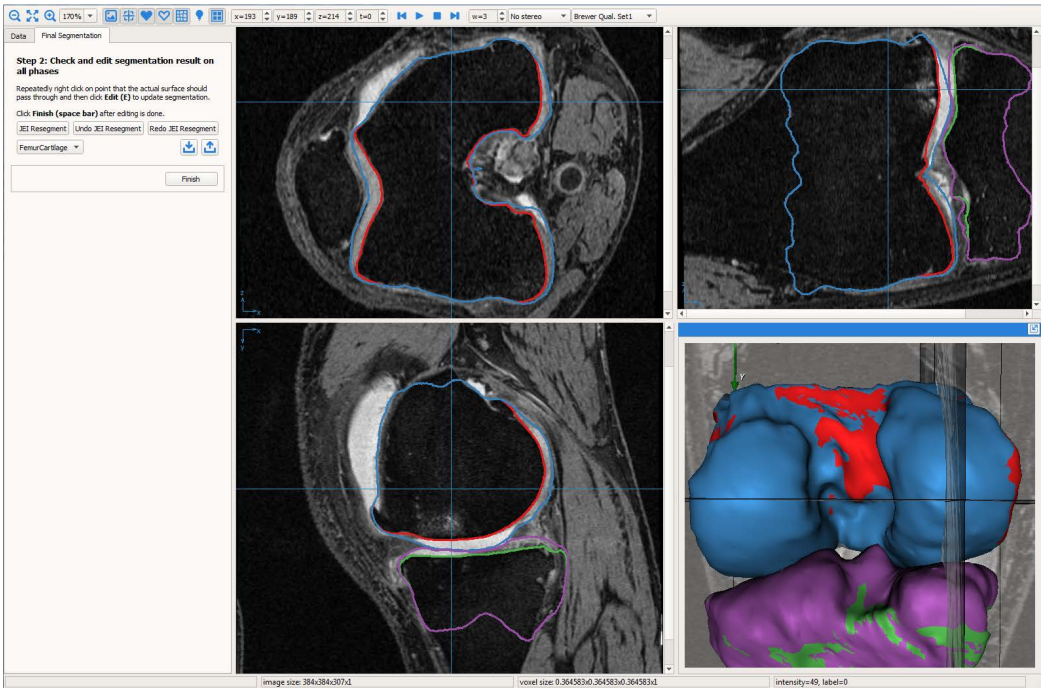


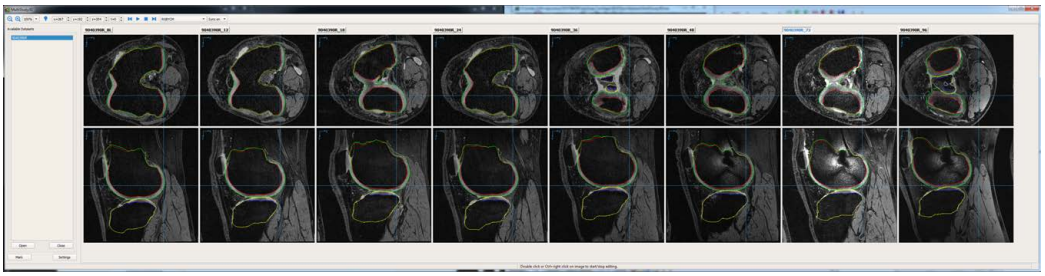
Figure 5.1: The graphical user interface for 3D JEI with the image volumes and the surface meshes overlayed.

the 4D automated LOGISMOS registers all the patients time points to a common space. Although the graph nodes and columns do not hold any spatial information, the ELF based geometric graph required the registration information to parse the clicked nudge points positions to the correct space. This additional meta information of the transformation matrices that were used for the registration of the various time points in 4D was also loaded to the 4D GUI.

5.4 Experimental Methods

14 patient MRI volumes with varying degrees of OA from the OAI with independent standard available were used in the study. All subjects were scanned using the DESS protocol with a voxel resolution of $0.36 \times 0.36 \times 0.7$ mm^3. All datasets were first segmented using the automated 3D LOGISMOS followed by JEI editing. Further, the same set of 14 patients were segmented at BL and 12M by 4D LOGISMOS and then

(a)



(b)

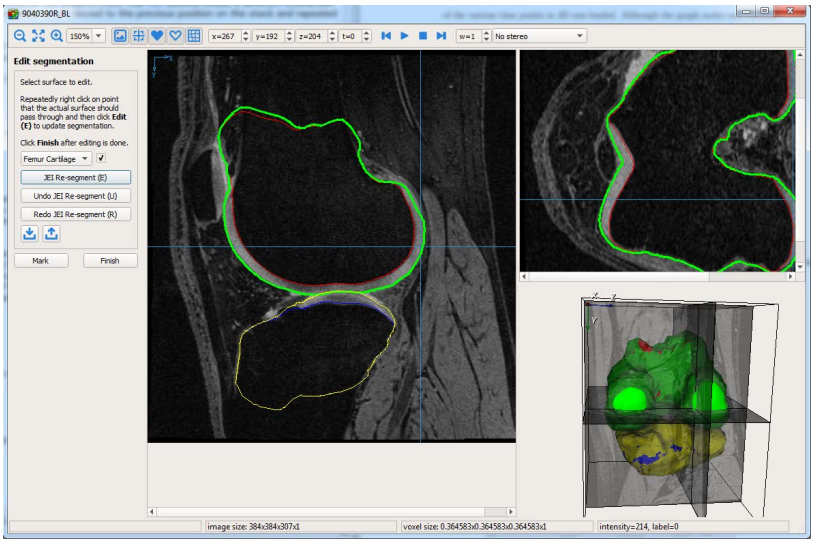


Figure 5.2: 4D LOGISMOS-JEI. (a) Longitudinal JEI viewer screen-shot showing a thumbnail of eight time-points of a single patient simultaneously. (b) Smaller editing window for each 3D time-point.

simultaneously 4D JEI edited. Cartilage border positioning errors and the thickness accuracy of the sub-plates after JEI editing were reported.

5.5 Results

Signed and unsigned cartilage surface positioning errors against the independent standards were reported in Tables 5.1, 5.2. The results were reported for 3D Baseline JEI and 4D JEI of baseline and 12M simultaneously edited on the central medial/lateral femur (cMF/cLT) and the central medial/lateral tibia (cMT/cLT). Paired t-tests (significant set at $p = 0.05$) were computed to test improvements in the accuracy by comparing the cartilage surface positioning errors of gradient costs based automated

Table 5.1: Signed surface positioning errors (in mm) of JEI-corrected versus gradient costs based automated LOGISMOS segmentation. Bold entries mark statistical significantly better performance of the pairwise comparisons.

Cartilage Signed	**cMF**		**cLF**		**cMT**		**cLT**	
Error [mm]	JEI	Automated	JEI	Automated	JEI	Automated	JEI	Automated
Baseline 3D	**-0.05 ± 0.21**	-0.90 ± 0.96	**-0.19 ± 0.12**	-0.61 ± 0.32	0.02 ± 0.17	0.41 ± 0.89	**-0.01 ± 0.18**	0.22 ± 0.46
Baseline 4D	**-0.11 ± 0.13**	-0.59 ± 0.71	**-0.28 ± 0.13**	-0.57 ± 0.37	-0.05 ± 0.18	0.13 ± 0.67	-0.01 ± 0.10	0.14 ± 0.40
12 Months 4D	0.02 ± 0.37	-0.46 ± 0.73	**-0.25 ± 0.17**	-0.56 ± 0.44	-0.12 ± 0.12	0.07 ± 0.61	0.01 ± 0.17	0.20 ± 0.44

Table 5.2: Unsigned surface positioning errors (in mm) of JEI-corrected versus gradient costs based automated LOGISMOS segmentation. Bold entries mark statistical significantly better performance of the pairwise comparisons.

Cartilage Unsigned	**cMF**		**cLF**		**cMT**		**cLT**	
Errors [mm]	JEI	Automated	JEI	Automated	JEI	Automated	JEI	Automated
Baseline 3D	**0.41 ± 0.06**	1.14 ± 0.97	**0.33 ± 0.06**	0.68 ± 0.23	**0.37 ± 0.07**	0.92 ± 0.53	**0.33 ± 0.09**	0.63 ± 0.21
Baseline 4D	**0.46 ± 0.19**	0.91 ± 0.48	**0.38 ± 0.12**	0.69 ± 0.21	**0.40 ± 0.12**	0.76 ± 0.25	**0.31 ± 0.08**	0.59 ± 0.16
12 Months 4D	**0.44 ± 0.23**	0.87 ± 0.46	**0.37 ± 0.12**	0.72 ± 0.22	**0.38 ± 0.10**	0.71 ± 0.25	**0.36 ± 0.10**	0.65 ± 0.16

LOGISMOS for the same 14 patient datasets. Bold entries marked statistically significantly better performance of the pairwise comparisons. All the unsigned surface positioning errors significantly improved. While the signed errors improved for all the sub-plates after JEI correction with only the bold ones showing statistical significance.

The sub-plate analysis algorithm was run on the independent standard surfaces as well. Fig. 5.4 shows a scatter plot of thickness accuracy compared against the independent standard for the cMT and cLT. We see that all the thickness values showed good agreement with the independent standard. Further, the 14 datasets took on an average of 16 min for 3D JEI editing while the combined Baseline and 12M editing of 4D JEI took an average time of 22 min. We expect that with more time points edited simultaneously the time required for interaction per dataset to reduce in comparison with individually 3D JEI editing each of them.

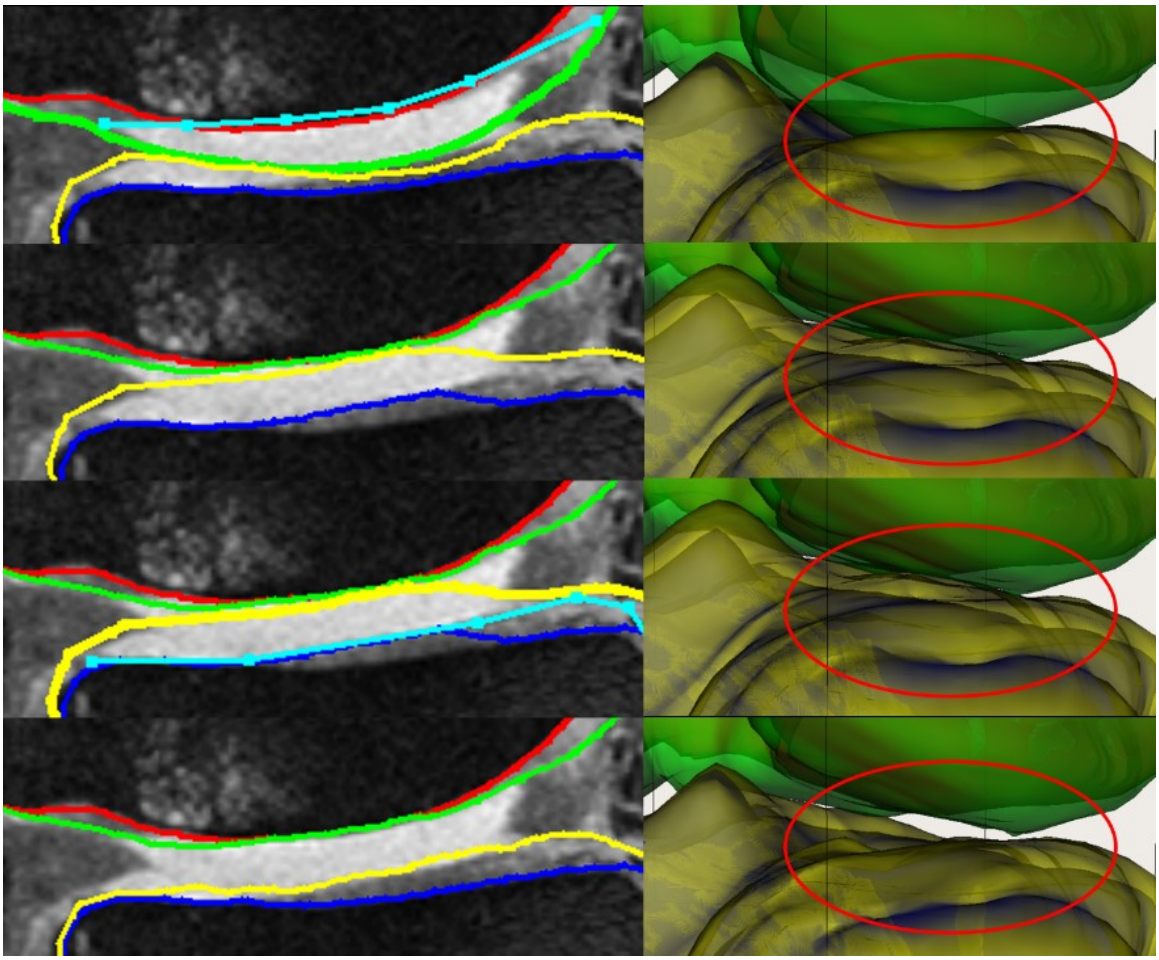


Figure 5.3: JEI work-flow to correct segmentation inaccuracy. The circled region indicates the 3D neighborhood correction based on a single 2D slice editing. The four surfaces shown as contours and 3D objects are the femur bone and cartilage (colored red and green respectively) and the tibia bone and cartilage (colored blue and yellow respectively). The nudge points marked in the intermediate steps are shown in cyan. A detailed video demonstration of the interaction steps is available at *http://bit.ly/2blYXFz*.

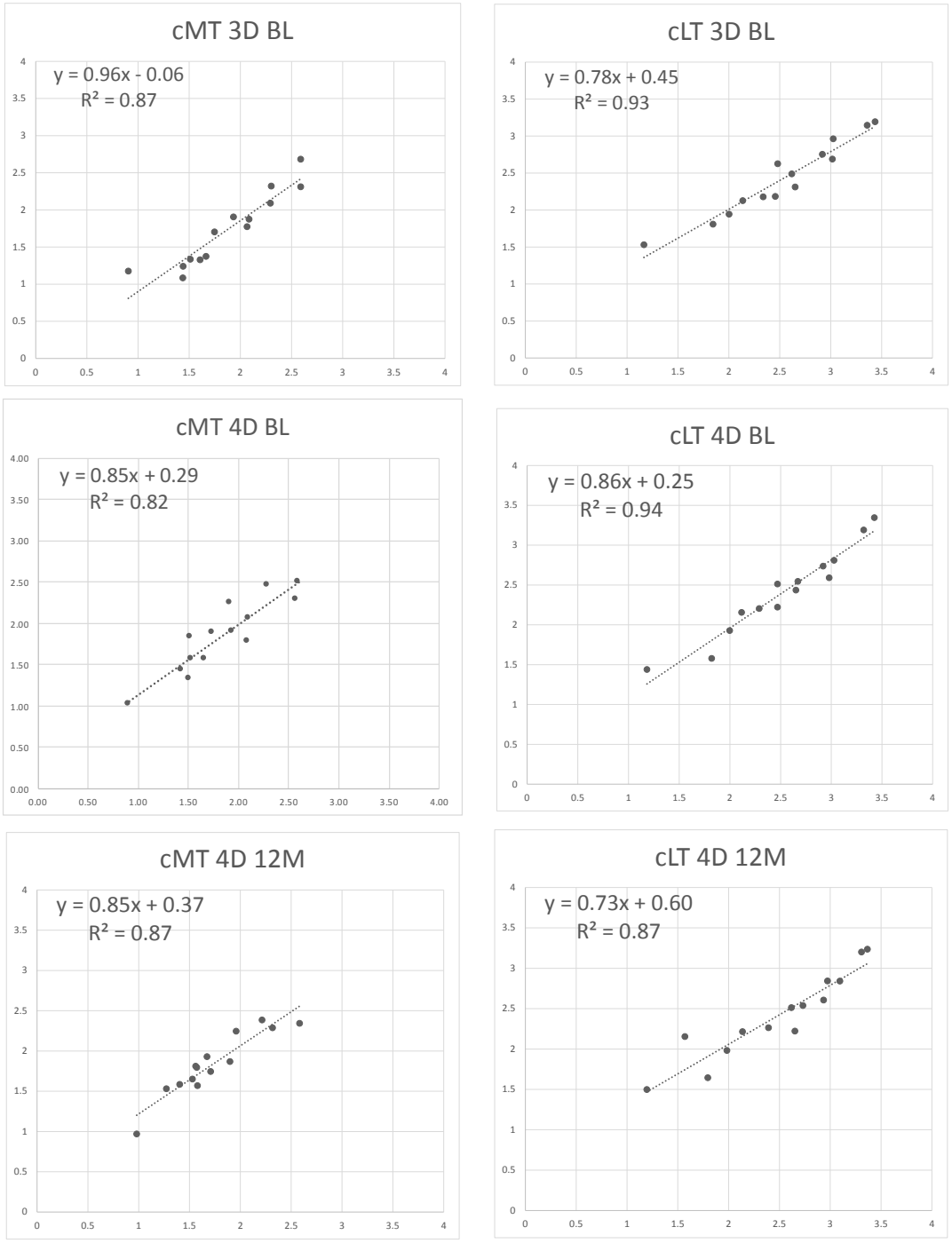


Figure 5.4: Thickness accuracy compared against the independent standard (in mm) for a central medial and lateral tibia sub-plates (cMT, cLT). The x-axis represents the thickness of the independent standards while y-axis represents the JEI-corrected thickness values.

CHAPTER 6
PUTTING IT ALL TOGETHER: QUANTITATIVE ANALYSIS TOOLS TO UNDERSTAND CARTILAGE THICKNESS LOSSES

6.1 Motivation

Given that the state of the art diagnosis currently used in clinics are semi-quantitative grading techniques to diagnose osteoarthritis there is a pressing need to deliver highly sensitive quantitative imaging-based bio-markers to improve diagnosis and aid in the early detection of osteoarthritis. Commonly analyzed are thickness in selected sub-regions of the knee which was covered in Chapter 3. However along with the loss of cartilage being the main indicator of osteoarthritis there is increasing evidence that osteoarthritis is a multi-parameter disease with several other non-cartilage imaging bio-markers showing promise as early of OA in patients (see Sec 1.4 for background details). In this chapter quantitative analysis of bone shape as a potential non-cartilage bio-marker is studied with statistically significant changes in bone shape between populations demonstrated. Further using cartilage thickness from previous time-points we use regression analysis to predict thickness of a later time-point.

6.2 Statistical Shape Analysis of Femur Bones

Statistical shape analysis allows a highly localized measurement of tissue atrophy and may be important in quantifying disease progression. Therefore population studies need reliable methods to quantify shape changes in OA. This section presents a novel, fully automated pipeline to quantify local differences in the femoral bone shape in OA. LOGISMOS segmented femoral bones are used for shape analysis computation using spherical harmonics point distribution model (SPHARM-PDM) [62]. The hypothesis is that significant differences will be observed around the trochlear groove. In addition to confirming this a priori hypothesis, the fully 3D approach enables an exploratory study that may reveal previously unknown structural changes in OA. Many methods exist for building statistical shape models including pairwise [42, 64]

and group-wise approaches such as minimum description length (MDL) [12, 24] and entropy-based correspondence [10, 46] (see [2] for an excellent review). Given the spherical topology and relatively simple shape of the femur bone, SPHARM-PDM is a natural choice for femoral shape analysis. Note that a fundamentally different approach would be volumetric registration of images (see [33] for a review). The current application is focused on the shape rather than image intensity characteristics and thus focuses on surface-to-surface registration instead.

The approach consists of three main steps: a. LOGISMOS segmentation resulting in the bones and cartilage surfaces of the femur and tibia. Although the analysis is only concerned with the analysis of shape changes in the femur bone for the current study, using the contextual information from multiple surfaces helps improve the segmentation accuracy. b. SPHARM-PDM approach [62] to obtain point-wise correspondences between subjects to build a statistical shape model. c. Finally, statistical analysis is performed using a non-parametric permutation test [47] to compare the clinically different populations with each other.

6.2.1 SPHARM-PDM correspondence and statistical analysis

A statistical shape model is built using the SPHARM-PDM framework [62] [1]. Given a 3D surface of spherical topology, SPHARM-PDM generates a spherical parameterization in an area-preserving manner while minimizing distortions. Orientation is determined using first-order ellipsoid alignment. Given a population of surfaces that were parameterized using this approach, surface locations that map to the same (ϕ, θ) coordinates on the sphere correspond to each other.

We begin by minimally smoothing the binary bone segmentation result from LOGISMOS. The interior holes in the segmentation are filled. Any remaining handles are detected and removed by either cutting them open or filling them in, based on whichever approach would modify fewer voxels in the segmentation [28]. This re-

[1] https://www.nitrc.org/projects/spharm-pdm

sults in binary segmentations of spherical topology, which are then converted to surface meshes via marching cubes. The SPHARM-PDM framework then computes the spherical parameterization, using the spherical harmonics of the form

$$\vec{v}(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \vec{c}_l{}^m Y_l^m(\theta,\phi), \tag{6.1}$$

where $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$ are the spherical coordinates, l and m are the order and degree of the spherical harmonic, and $Y_l^m(\theta,\phi)$ is the associated Legendre polynomial. The coefficients $\vec{c}_l{}^m$ are computed using least squares optimization which seeks to minimize metric distortion while preserving area.

The resulting spherical parametrization is then uniformly re-sampled using an icosahedron subdivision to determine sample points on the sphere. The alignment is achieved using the first order ellipsoid from spherical harmonic coefficients by rigid Procrustes method.

Statistical shape analysis was performed by computing the modified Hotelling T^2 two-sample metric at every surface sample [47] to test whether two groups of surfaces differ from each other locally. The metric is tested using permutation tests with the null hypothesis that the two groups of surfaces are statistically identical to each other.

6.2.2 Experimental Methods

576 knee MRI volumes from 192 subjects available at baseline (BL) , 12 months (12M) and 24 months (24M) follow-up clinical visits as part of the OAI Biomarkers consortium project [27] was used for the current study. All subjects were imaged using a double echo steady state (DESS) sequence with a resolution of $0.36 \times 0.36 \times 0.7$ mm. All subjects were classified as progressor (PROG, n=79) or non-progressor (NPROG, n=113) by the OAI experts based on changes in the KL grade and increased knee pain between BL-24M [27]. We limited the current study to patients that had a baseline grade of $KL = 2$.

The goal of the study was to identify statistically significant differences in the femur bone shape that may provide early indications of OA progression. Based on initial reports in the literature, we hypothesized that the trochlear groove region would show statistically significant differences. This hypothesis was tested by analyzing differences between the PROG and NPROG groups at each time-point. We further hypothesized that any differences found in the earlier time-points would become more pronounced as the disease progressed.

Left and right knees were analyzed separately to explore any potential laterality differences. 20 spherical harmonics were used per surface. We used 1000 iterations for generation of the parametric mesh. The icosahedron sampled sphere used a subdivision factor of 20 for a total of 4,002 vertices on the surfaces. For the statistical permutation tests we used 20,000 iterations with a significance value set at $p = 0.05$. The Shape Population Viewer[2] was used for quality control of the SPHARM-PDM parametrization, and any failures were excluded (n=15 at BL, n=18 at 12M, n=15 at 24M).

6.2.3 Results

Each MRI volume was segmented with LOGISMOS and the spherical parameterization was obtained with the SPHARM-PDM framework (approx. 6 hours/surface). Following quality control on all results, statistical analysis was conducted (approx. 20 min).

Fig. 6.1 shows the results of the correspondence quality obtained with SPHARM-PDM. A representative set of eight left knee surfaces is shown with their ϕ coordinate maps. The correspondence quality can be evaluated by the observing the consistent placement of the $\phi = 0$ line (where the red and blue regions come together) in anatomically corresponding regions on each surface, despite the considerable shape variation between the knees.

[2] https://www.nitrc.org/projects/shapepopviewer

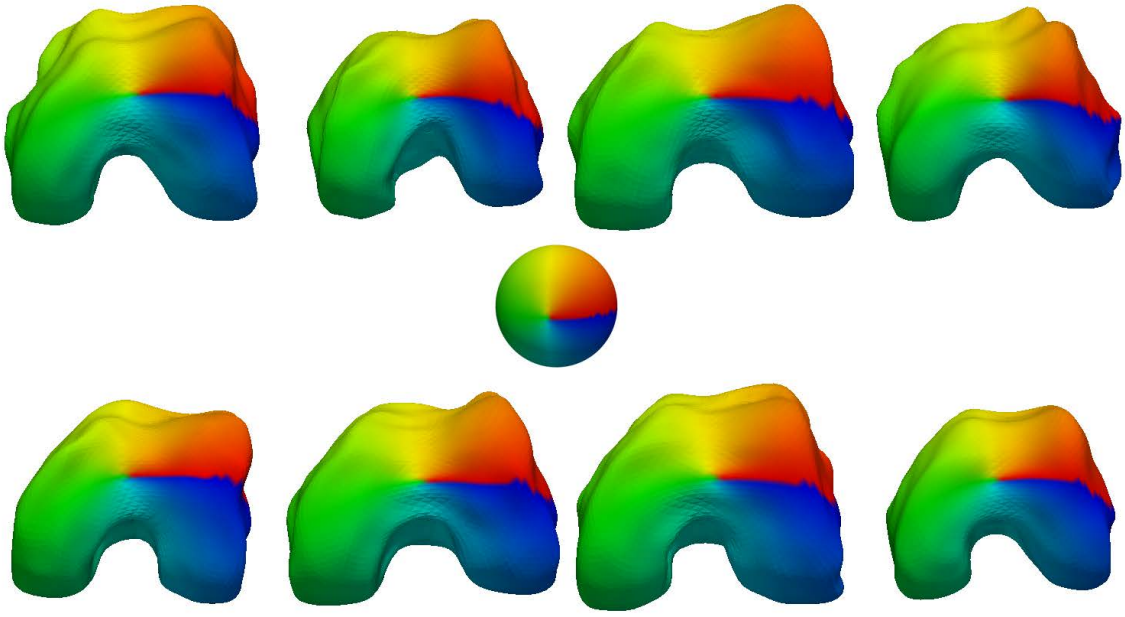


Figure 6.1: Qualitative correspondence evaluation on an illustrative subsample. The color map represents the ϕ coordinates on the parameterized surfaces.

Fig. 6.2 shows the p-values of significant surface differences between the PROG and NPROG groups at each time point. The trochlear groove region shows highly significant differences. The observed differences spread spatially and gain power at later time points. Weaker differences include the condyles on the left knees as well as lateral and medial peripheral regions of the right knees.

6.3 Predicting Thickness Losses

Cartilage thickness is the main imaging bio-marker that is used to understand the progression of OA. Given its importance, it is beneficial to predict the thickness at later time-points to understand the losses or in clinical drug trials to predict the efficacy of the treatment. The goal of this study was to use regression analysis to predict the thickness of the load bearing sub-plates at 48M using sub-plate thickness

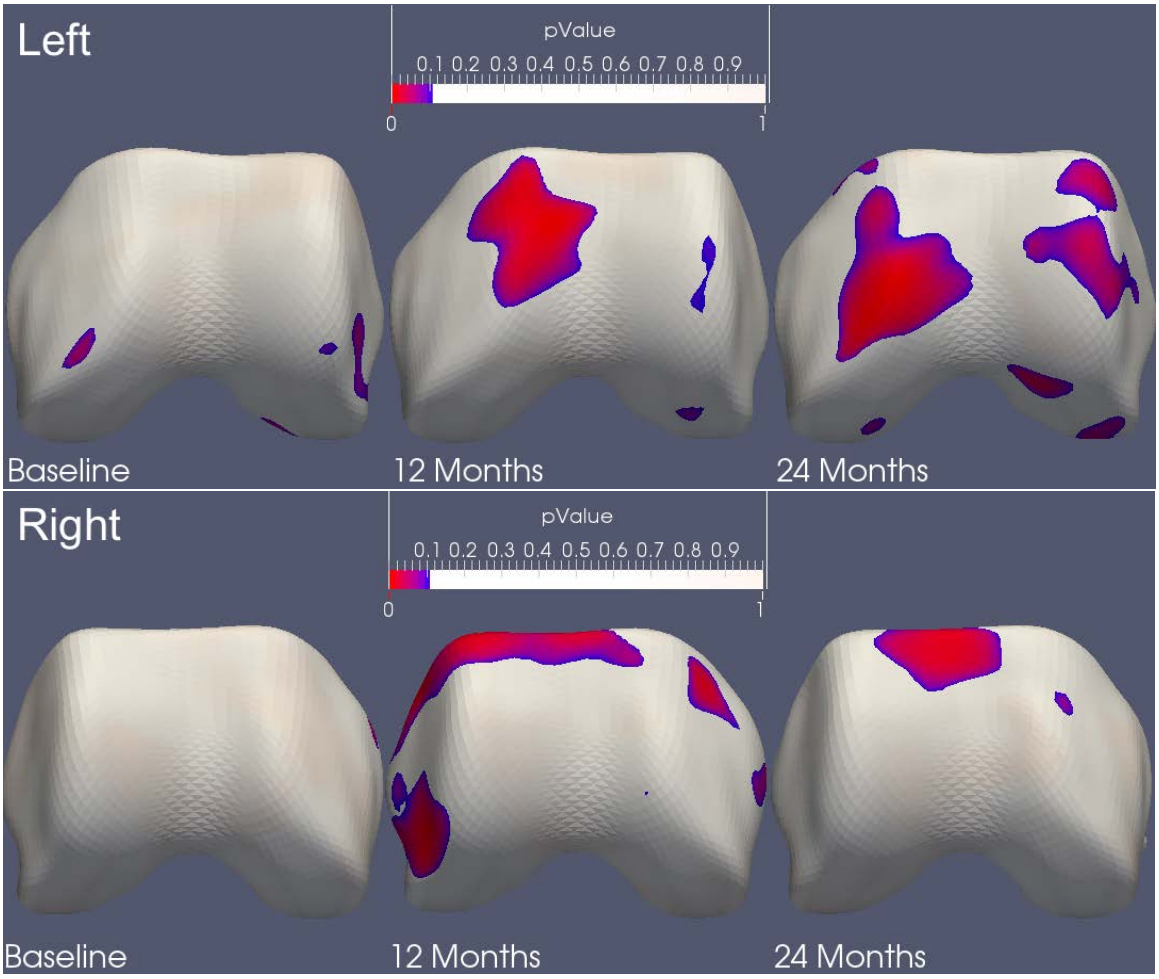


Figure 6.2: p-value map of shape differences between groups (PROG vs NPROG) at BL, 12M and 24M.

value computed for the same patient at BL, 12M, 24M, and 36M. The study explored two main experiments: a. Different types of regression techniques that best capture the relation between the dependent variables (thickness at BL-36M) and the independent variable (thickness at 48M), b. the effect of using 3D segmentation or 4D simultaneous multi-time point segmentation to compute the thickness.

6.3.1 Regression Analysis

Linear Regression Models Linear regression models the relationship between the target output variable ($\hat{y}$) as a linear combination of the inputs (X). The predicted

value can be modeled as:

$$\hat{y}(w, x) = w_0 + w_1 x_1 + \cdots + w_p x_p \tag{6.2}$$

where $w = (w_1, ..., w_p)$ are the coefficients that are used to combine the different sub-plate thickness values to predict the output thickness y, X are the thickness values of the inputs. The coefficient weights are solved by fitting a model which minimizes the residual sum of squares between the predicted response and the observed responses in the least squares sense by solving:

$$min_w ||Xw - y||^2 \tag{6.3}$$

Support Vector Machine Regression Support vector machines are supervised learning methods that can be used for regression. They work by constructing a set of higher dimensional hyperplanes which are optimized by defining the loss function that ignores errors which lie within a certain distance of the true outputs (also called as epsilon intensive loss functions). The SVM regression is optimized by posing the problem as a quadratic programming problem. The problem is reformulated into the lagrangian dual thereby avoiding the inequality constraints and simplifying the optimization process [65]. A detailed explanation of the support vector machine regression is provided in [58].

6.3.2 Experimental Methods

600 patients from the OAI at BL, 12M, 24M, 36M and 48M follow-up clinical visits which are part of the FNIH bio-markers project [27] was used for the current study. Out of the 600 patients, 201 of them had one or more missing follow-up clinical visits and were dropped resulting in a total of 1995 (399 x 5-time points) knee MRIs.

All the time points were segmented using 3D LOGISMOS and simultaneous five

time point 4D LOGISMOS segmentation. All 1995 segmentation results were quantitatively analyzed using the sub-plate detection algorithm for thickness from 3D and 4D segmentations separately. Scikit-learn package was use to run the regression analysis [49]. Support vector regression with radial basis function kernels (SVR-rbf) and Linear regression (LR) were used for the prediction analysis. The first experiment was to compare the SVR-rbf with the linear regression to understand if a linear model was sufficient in capturing the thickness. Secondly the prediction accuracy of the different sub-plates were studied by taking the BL-36M as training and predicting the 48M thickness. This was compared with the computed 48M accuracy. The analysis used a leave-one out strategy (398 training & 1 testing). Regression analysis was done on 3D and 4D sub-plate thickness to study the benefits of inter-time point contextual information enforced during the 4D segmentation. Further to test the benefits of each time-point thickness used during training, a set of experiments were run with reduced training sets to train a regression model that predicts the 48M thickness.

6.3.3 Results

Every experiment was validated on four major sub-plates - medial/lateral femur (MF/LF) and medial/lateral tibia (MT/LT).

Comparing the Regression Performance in Thickness Prediction: Fig. 6.3 shows the comparison of SVR-rbf (*left column*) and the LR (*right column*) for sub-plate thickness values computed on 4D LOGISMOS segmentation results. We see that both the methods predict the thickness of the femur sub-plates (MF & LF) with similar accuracy. However, for the tibial sub-plates, the linear regression model is able to better predict the thickness value at 48M with the best performance amongst all eight experiments achieved for the MT plate prediction using linear regression which had a $R^2 = 0.91$ and slope almost equal to one. The remaining sets of experiments used linear regression for the experiments.

Prediction Accuracy of 4D versus 3D LOGISMOS Thickness using Linear Regression: In this experiment, we studied the benefits of having contextual information by comparing the 4D LOGISMOS segmented sub-plate thickness with the equivalent 3D thickness. The linear regression training used the BL to 36M thickness values as training. Fig. 6.4 shows the regression accuracy of 4D (*left column*) versus 3D (*right column*) thickness. We see that the 4D based thickness prediction was able to perform better for all the sub-plates. Along with the experiments conducted in Chapter 4 where we showed the improvements on using 4D over 3D LOGISMOS, the results shown here on prediction accuracy are convincing improvements that can be attributed to providing more consistently accurate thickness values.

Effect of Reducing the Time-Points Used for Training: For all the previous experiments irrespective of the regression method or the type of thickness used, the training always had as inputs the thickness values from BL, 12M, 24M and 36M. In this experiment, we attempt to deconstruct how much of an added benefit the time points provide. All experiments in this section used 4D LOGISMOS based thickness with linear regression. Fig. 6.5 shows the effect of using only the BL, 12M and 24M thickness values for predicting the 48M on all four sub-plates. The prediction of the medial regions were more adversely affected than the lateral regions. Further reducing the number of time-points to two we study the effects of training using BL and 12M (Fig. 6.6) or using BL and 24M (Fig. 6.7). Although both of these performed far worse than training with more time-points, it was interesting to note that using BL and 24M was markedly better than using BL and 12M for prediction. Overall this experiment produced quite interesting trends in terms of understanding how the prediction of 48M thickness is affected by each of the time-points. The results clearly show that using all four time-points gave the best prediction. When the 36M was dropped, the medial side predictions were more adversely affected while using only

two time-points the information from BL and 24M was more useful in capturing the thickness losses that occurred at 48M.

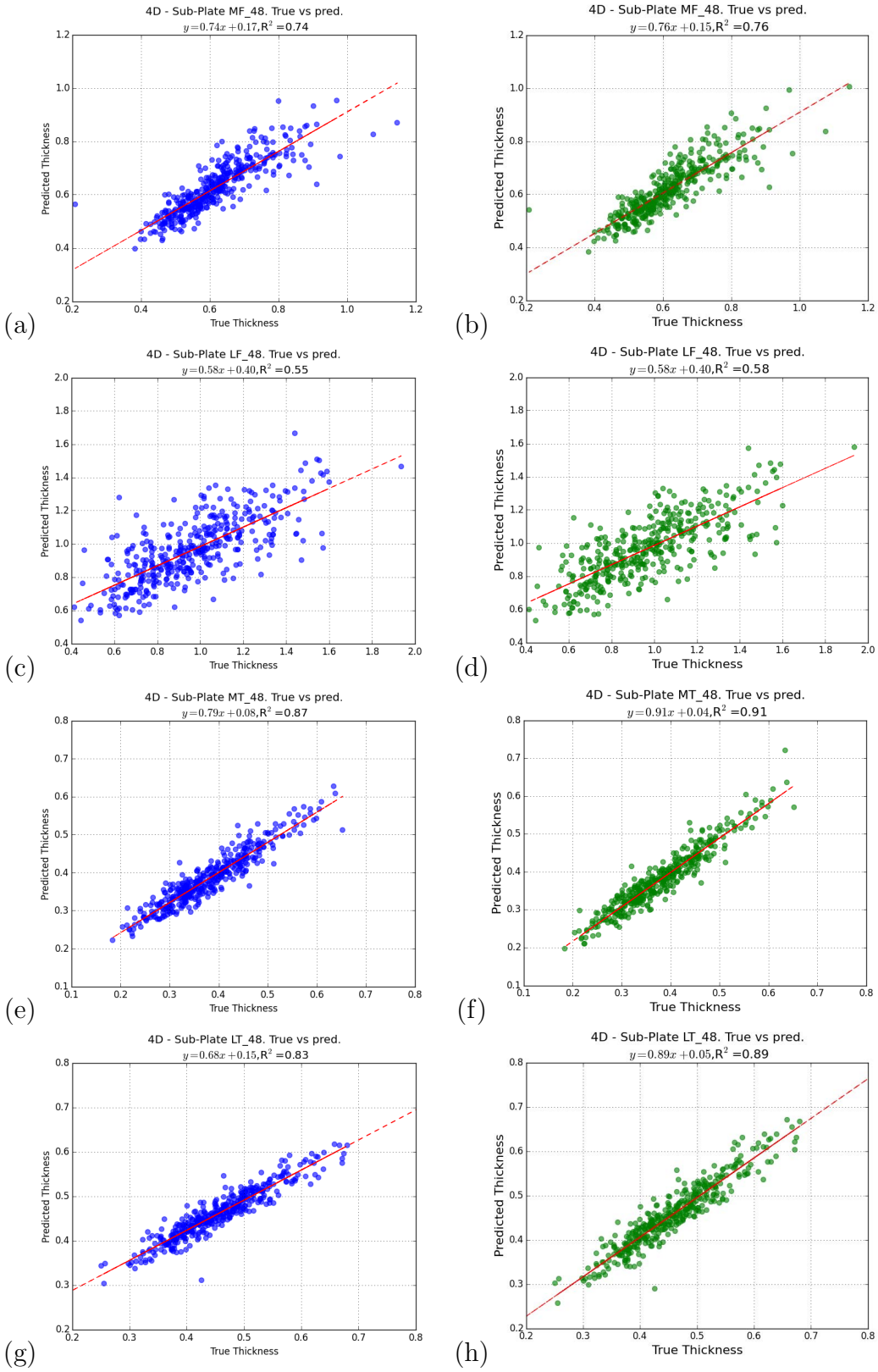


Figure 6.3: Comparing the prediction accuracy of SVR-rbf (*left*) versus linear regression (*right*). The femur sub-plates had more or less the same performance while the linear regression prediction the tibial sub-plate thickness with higher accuracy.

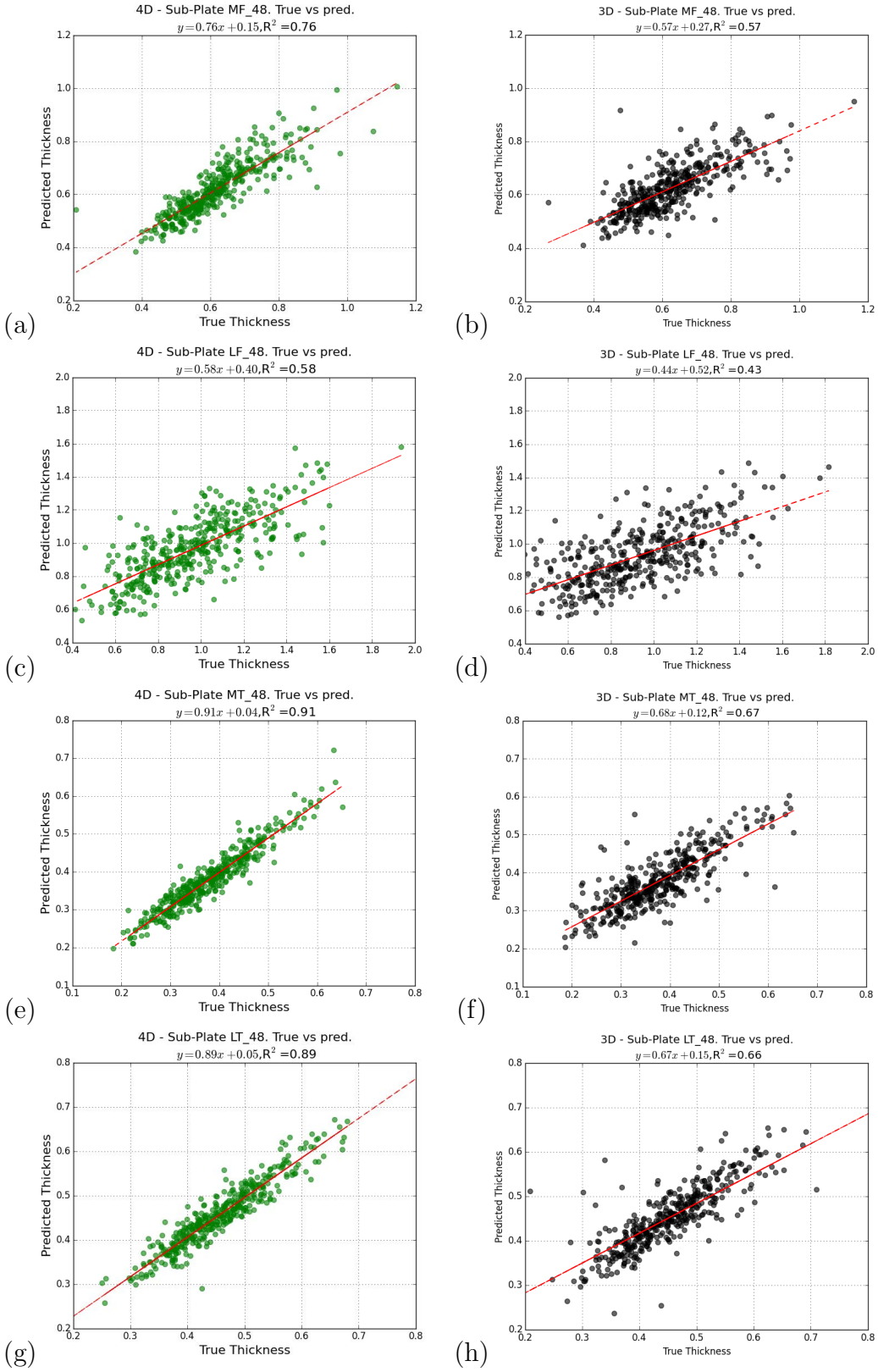


Figure 6.4: Linear regression accuracy of 4D (*left column*) versus 3D (*right column*) thickness. We see that the 4D based thickness prediction was able to perform better for all the sub-plates.

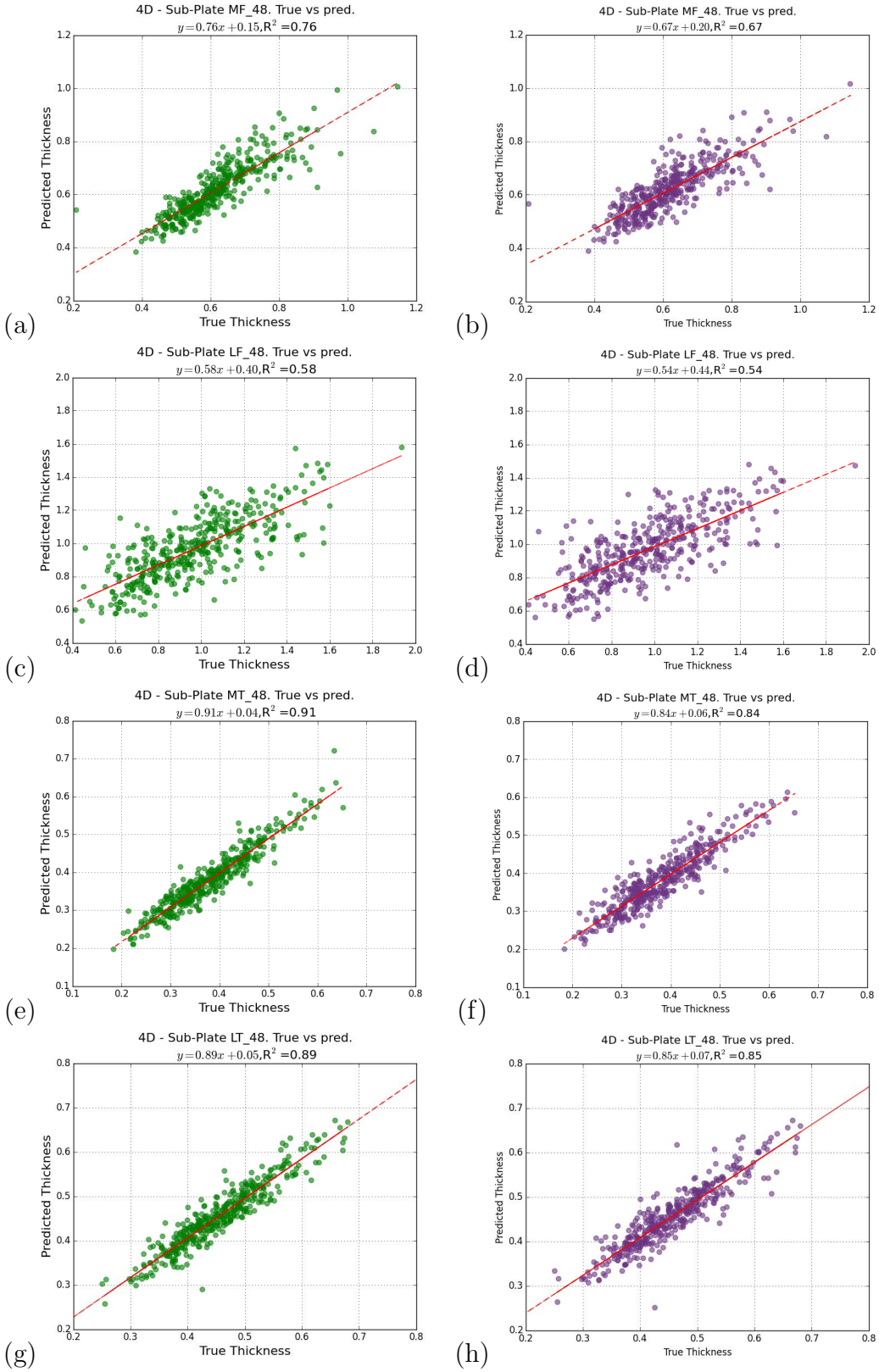


Figure 6.5: Comparing the effect of using only the BL, 12M and 24M (*right*) thickness values for predicting the 48M on all four sub-plates. The prediction of the medial regions were more affected than the lateral regions in comparison with the fully trained 4D linear regression prediction shown on the *left* column.

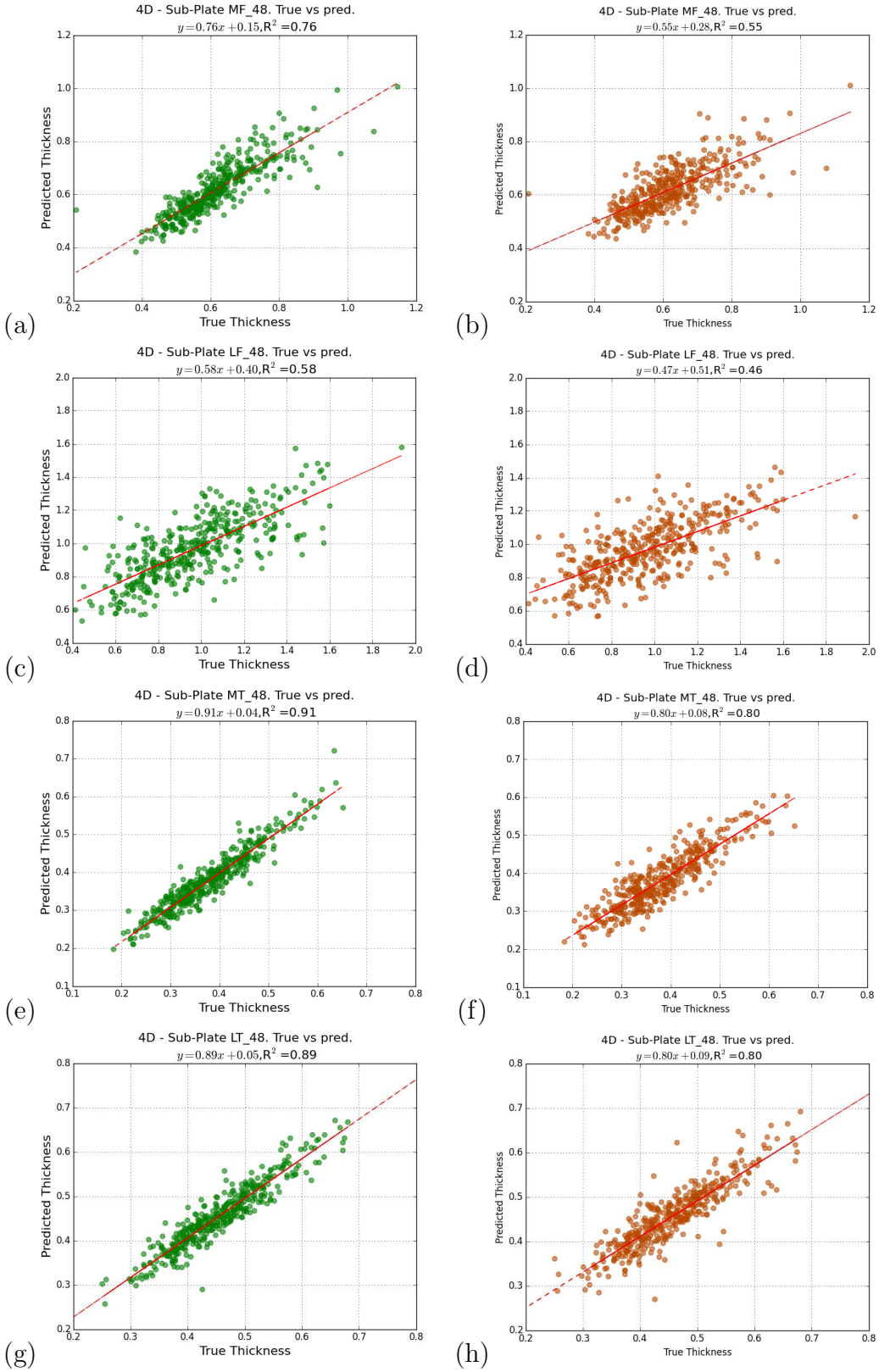


Figure 6.6: Reduced time-points (BL and 12M) to study the effects of training are shown in the *right* column. The fully trained 4D linear regression prediction is shown on the *left* column for comparison.

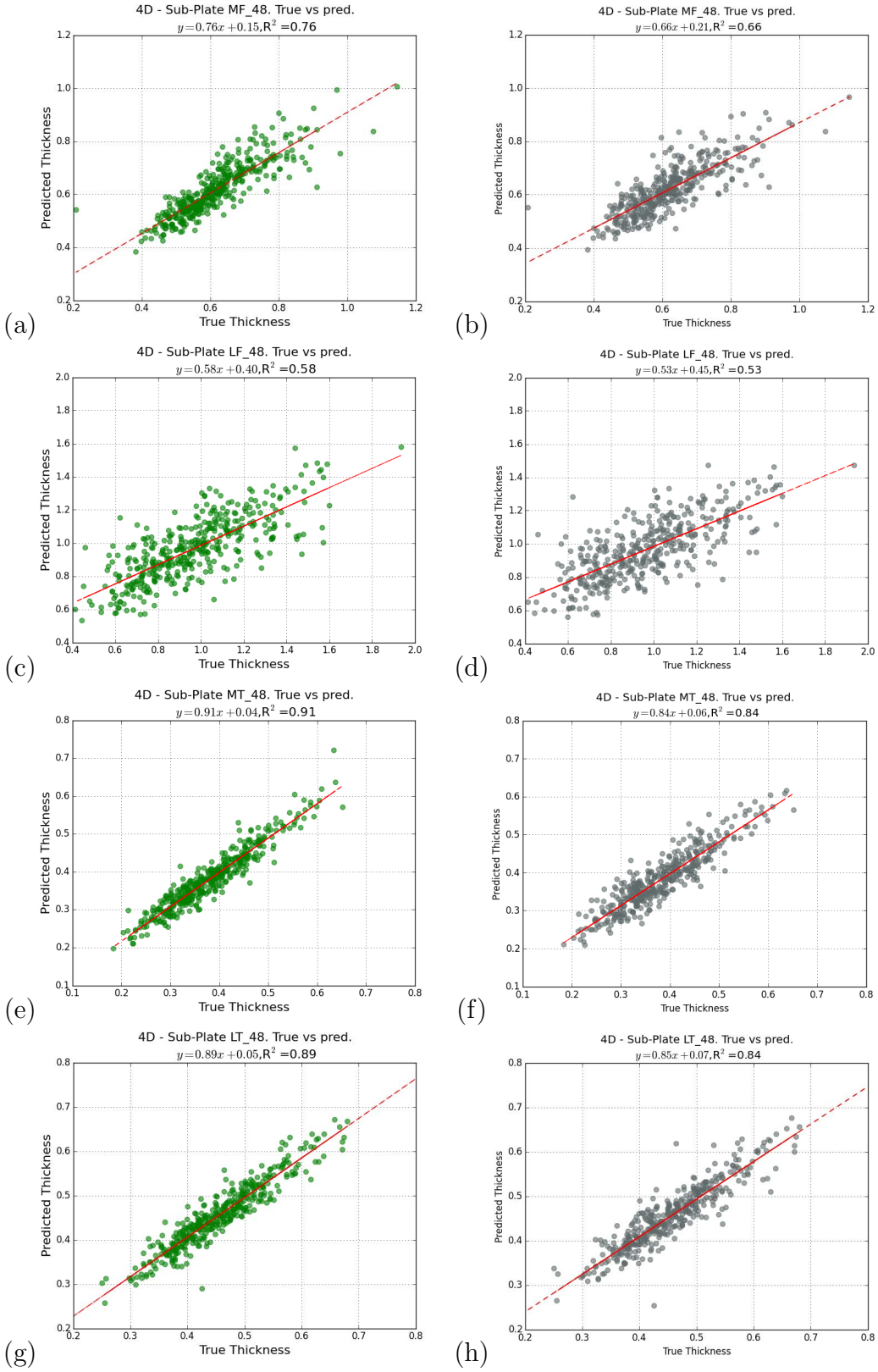


Figure 6.7: Reduced time-points (BL and 24M) to study the effects of training are shown in the *right* column. The fully trained 4D linear regression prediction is shown on the *left* column. Comparing with the Fig. 6.6, the information from BL and 24M is more useful in capturing the 48M thickness.

CHAPTER 7
CONCLUSIONS, FUTURE DIRECTIONS & CONTRIBUTIONS

7.1 Conclusions

In this thesis work we demonstrated the benefits of using automated LOGISMOS segmentation for knee MRI to study patients with osteoarthritis. In Chapter 2, the basic 3D LOGISMOS for knee MRI was explained. Further a hierarchical set of random forest classifiers was designed to provide robust cartilage cost functions for the LOGISMOS segmentation. The first stage of the classifier consisted of a neighborhood approximation forests followed by a clustered random forest collecting features along the geometric graph columns. Our experiments showed that the combination of global and local contextual features from the hierarchical classifiers produced statistically significant improvements in segmentation accuracy in comparison with the simple gradient based costs and the single stage random forest classifiers. Chapter 4 extended the 3D LOGISMOS segmentation to 4D LOGISMOS to simultaneously segment multiple time points of patient follow up visits. With disease progression several artifacts similar in intensity to the cartilage begin to appear in the later time points of MRI volumes. By enforcing the inter-time point constraints as graph edges between the 3D LOGISMOS graphs, the segmentation over the time points become more consistently accurate. Given that the 3D LOGISMOS is already very accurate, we were unable to show drastic differences between 3D and 4D over a limited set of 54 patients at baseline and 12Months. However we did demonstrate that the overall consistency of the 4D segmentation accuracy improves when analyzing 399 patients over 5 time points (baseline, 12 month, 24 month, 36 month and 48 month follow-up scans) although no independent standards were available. With the assumption of a linear deterioration in cartilage thickness the correlation coefficient R was computed for 3D and 4D segmented patient datasets. The 4D segmentation results had a higher number of patients thickness that showed a R value close to one. Although the as-

sumption of linear thickness losses may not be wholly correct, for the experiment it did demonstrate as a proof of concept the benefits of using 4D segmentation. A different analysis approach with independent standards available for all patients at every time-point will validate the benefits more accurately.

Although LOGISMOS 3D and 4D segmentation are very robust a few cases do exist where interactive post segmentation correction was necessary. In Chapter 5, we designed just-enough interaction editing to correct the segmentation inaccuracies. Unlike previous methods, the designed algorithm is able to maintain global optimality over the existing methods for multiple objects and surfaces. The user interaction was used to change the costs functions in the underlying graph thereby ensuring global optimality with each interaction. An additional benefit of this method was avoiding topological errors that occur because of repeated surface corrections. The graphical user interface designed for the JEI was built upon a highly modular multi-platform design. 14 JEI corrected patients in 3D showed significant improvement over the automated LOGISMOS segmentation. While using the 4D JEI for the same 14 patients allowed for simultaneous correction of baseline and 12 month segmentations. The results were similar to the 3D JEI with the additional benefit of completing the 4D JEI on an average of 22 min when the 3D JEI took 16 min. These JEI corrected segmentations were used as training annotations for the hierarchical random forest classifiers described in Chapter 2. The added benefit is that the resulting classifier based segmentation would have higher accuracy and therefore require lesser interaction. Using this feedback of JEI correction and automated segmentation, we anticipate that the number of training examples can be expanded quickly providing a larger annotated training dataset much faster.

Chapters 3, 6 were focused on developing automated quantitative algorithms. Chapter 3 developed a fully automated sub-plate detection algorithm to identify the most common load bearing regions of the cartilage that are likely to be affected by

osteoarthritis. Several semi-automated algorithms for the sub-plate detection exist in the literature. The developed method was fully automated requiring only the LOGISMOS segmented surfaces as inputs. These sub-regions helped identify locally the losses in cartilage over the different time-points of the patient. Chapter 6 developed spherical harmonics based point distribution model (SPHARM-PDM) to analyze the femur bone shape changes between the progressing and non-progression osteoarthritis population. Our proposed method was able to quantitatively analyze local regions of the femur bone in 3D. Permutation tests between the two patient populations showed statistically significant differences in the regions around the trochlear groove of the bone confirming the hypothesis. From our experiments, the shape changes spread spatially over time. For the left knees analyzed, significant changes were also found in the femoral condyle which could be a factor in the cartilage losses with disease progression.

Predicting the cartilage thickness is crucially important especially in clinical trial to study the efficacy of the interventional drugs. We tested liner regression models and support vector machine regression with radial basis function kernels for predicting the thickness of the cartilage sub-plates at 48 month using thickness information from the previous time-points. Our experiments found that using a linear regression model provided better prediction accuracy than using support vector machine regression. Our experiments compared the prediction accuracy when using 4D LOGISMOS versus 3D LOGISMOS segmented sub-plate thickness. The prediction accuracy of the using the 4D LOGISMOS segmented thickness was higher compared to the 3D version on the equivalent sub-plate regions. Further we experimented on the benefits of using different time-point thickness for training the regression model. While training only using the baseline, 12 month, 24 month thickness to predict the 48 month, we found that the medial sub-plate prediction accuracy was more adversely affected. Further when experimenting with using only baseline and 12 month or baseline and

24 month thickness, the latter combination of thickness information predicted the 48 month thickness more accurately. Overall the best prediction in accuracy was obtained using a liner regression model with sub plate thickness computed from 4D LOGISMOS segmentation using all four previous time points (BL-36M) to predict the 48 month thickness.

7.2 Limitations

Given the different algorithms used to perform the segmentation and quantitative analysis throughout the thesis, it is important to highlight the limitations and the stability of sub-steps in the different algorithms and their effect on the final results.

7.2.1 LOGISMOS Segmentation

Recall that the 3D LOGISMOS segmentation consisted of four major steps: Volume of interest detection (VOI), pre-segmentation, image re-sampling and multi-object multi-surface graph segmentation. The VOI detection is followed by affine fitting of the mean shape mesh. As long as the fitted mean shape is able to capture the regions of the objects to be segmented fairly accurately, the subsequent steps and the final segmentation are quite robust with respect to the the variations in VOI bounds. However if the VOI failed, the subsequent steps could not be corrected.

The pre-segmentation captured the dark bone to bright cartilage gradient boundary along the ELF search columns drawn on the fitted mean shape mesh. In order to improve the robustness the inter-column arcs maximum allowed for a capture range of 2.1 mm which accounted for the approximate initially fitted surface. Less that 5% of all cases failed at the pre-segmentation step when the initial VOI bounds captured the objects of interest reasonably accurately.

The gradient based costs functions used were quite robust to variations in the node spacing along the ELF graph columns. However, for the classifier based method, the node spacing was also inherently trained upon due to which the performance was

sensitive to changes in the node spacing between the training and the testing sets.

For the 4D LOGISMOS an extra parameter of inter-time-point arcs needed to be set to constrain the changes across the follow-up MRI's temporally. From our initial experiments, setting a very low value resulted in an over smoothing temporally resulting in increased segmentation inaccuracy. However if it was set too high the inter-time-point arcs had no effect and the results were similar to the individual 3D segmentation at their respective time-points. The final inter-time-point arcs were set at 0.6 mm.

7.2.2 Sub-plate Detection

The sub-plate detection algorithm automatically identified nomenclature compliant regions of the cartilage in which thickness losses due to osteoarthritis were studied. The central medial and lateral femur are computed as 60 % of the region from the trochlear notch to the posterior most point on each of the respective condyles. To test the stability of the sub-plate detection, a reduced 45% of the region was studied. The experiment replicated the study in Table 2.5 where statistically significant improvements in segmentation accuracy were demonstrated using the hierarchical random forest classifiers as costs functions over gradient based costs. Table 7.2.2 reports the cartilage surface position errors (w.r.t. independent standard, 108 patients) on 3D LOGISMOS segmented surfaces using hierarchical random forest classifiers and gradient costs. The reported sub-plate errors were on 45% of the region. In these 45% sub-regions as well using learned classifier as cost function showed statistically significant (paired t-test, p=0.05) improvements in segmentation accuracy.

7.2.3 Femoral Bone Shape Analysis

The bone shape analysis used the LOGISMOS segmented mesh surface which was analyzed using SPHARM-PDM. The SPHARM-PDM algorithm is a parametric mapping of the mesh objects onto a sphere using harmonics. The uniform re-sampling

Table 7.1: Cartilage border positioning errors (in mm) for femoral sub-plates achieved by hierarchical classifier, gradient cost and single stage RF classifier. This experiment used the 45% sub-plate regions on the femoral condyles to test the stability of the algorithm.

	NAF+RF	Gradient	p-value
Femur signed	**-0.01±0.18**	-0.31±0.24	$\ll 0.001$
Femur unsigned	**0.55±0.11**	0.69±0.13	$\ll 0.001$
cMF signed	**0.04±0.70**	-0.37±1.06	$\ll 0.001$
cMF unsigned	0.62±1.17	0.87±1.13	0.099
cLF signed	**-0.28±0.25**	-0.56±0.39	$\ll 0.001$
cLF unsigned	**0.43±0.14**	0.70±0.22	$\ll 0.001$

of the sphere results in point-to-point correspondences between the two study populations. Statistically significant local differences were obtained using permutation tests. In Chapter 6, statistically significant local changes between the progressing and non-progressing OA populations were demonstrated. The main regions of change occurred around the trochlear notch region with the changes between the populations gaining power and spreading spatially over time. The experiment used 79 progressors (PROG) and 113 non-progressors (NPROG) to study the shape changes.

To test the stability of the algorithm, the earlier experimental data was divided into two approximately equal sets. Set 1 consisted of 40/57 patients in the PROG/NPROG group respectively and the set 2 consisted of 39/56 patients. The resulting groups were examined in the shape population viewer with patients (n=19 at BL, n=21 at 12M, n=17 at 24M) removed due to sub-step failure. Figs. 7.1, 7.2 shows the statistically significant differences between the two groups for left and right knees of both sets. We see that for the first set the differences in population is similar to the earlier experiment shown in Fig. 6.2. Although differences were found in set 2, the changes were not as pronounced. This indicates the limitations of the method in terms of the number of patients or – at the same time – the impact of random selection of patients in the relatively small test groups. Although all patients were of

KL2, the heterogeneity of KL grading could also be a factor in the differences between the two sets.

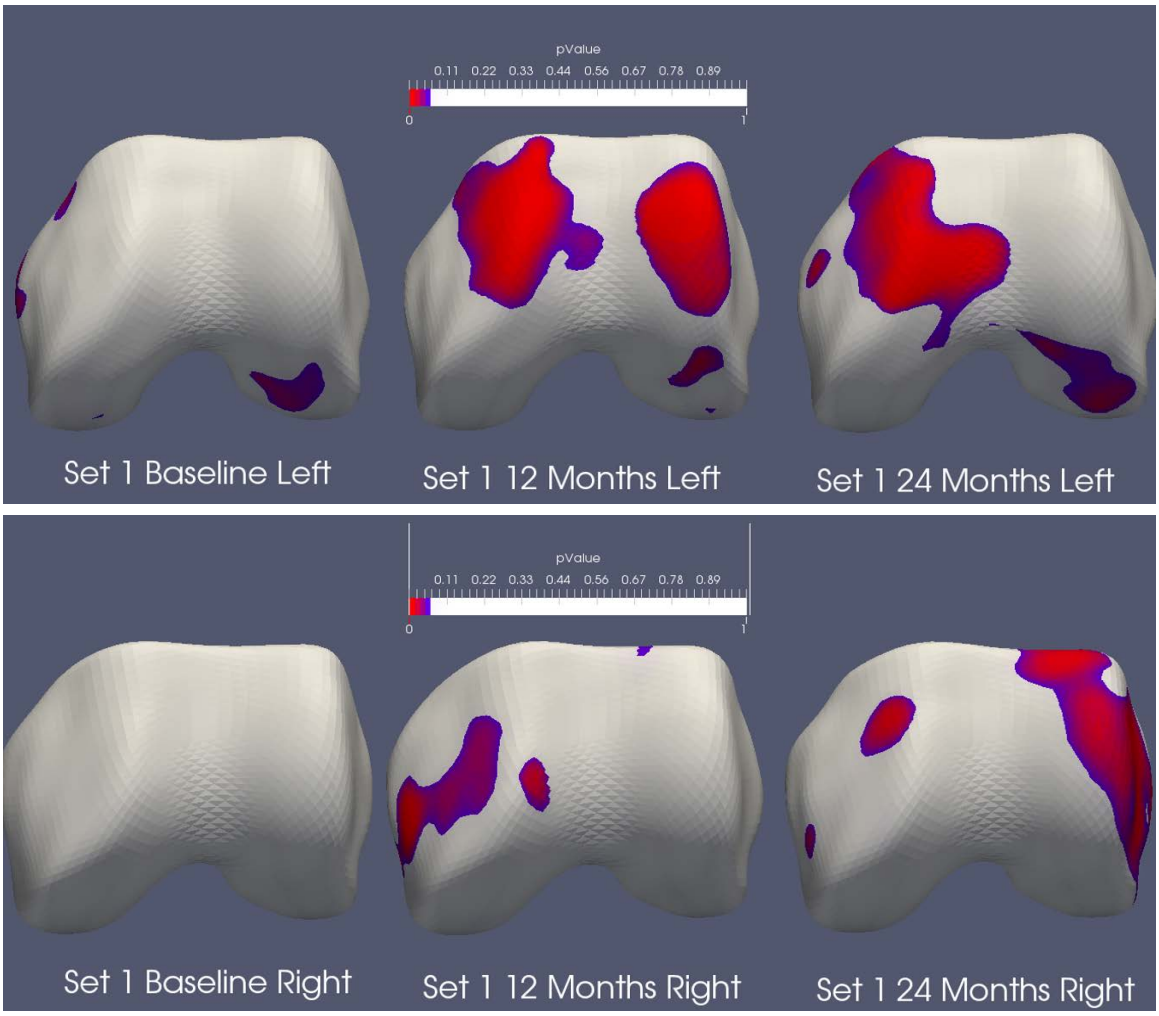


Figure 7.1: Comparison of the PROG vs NPROG datasts in set 1. The localized statistically significant regions shown are similar to the results obtained when using the entire experimental datasets.

7.2.4 Regression Analysis

The major limitation in the regression analysis were the lack of independent standards. Although the combination of experiments in Chapter 4 shows that using 4D

LOGISMOS based sub-plates are better at predicting 48M, the accuracy of predicting the true change in thickness needs to be understood and that is only possible with an available gold standard. Further it may be that using only the thickness for predicting may not be sufficient and other textual and intensity features may help improve the accuracy once we have an accepted gold standard for comparison.

7.3 Future Directions

There are several broad themes on which this work can be extended upon.

7.3.1 Machine Learning

Random forests have played a very important role in significantly improving the cartilage cost functions in the LOGISMOS framework. However, there are several other aspects of the study where machine learning can help. With advances in artificial intelligence community, a reinvented neural network called deep learning is proving to be extremely adept. The random forests based costs functions could be replaced by convolutional neural networks [37] and validated against the independent standard.

7.3.2 LOGISMOS Graph Extension

The patients in the OAI cohort also have been scanned on several other pulse sequences such as T_2 mapping, and $T_{1\rho}$. These sequences are good at identifying changes in the collagen matrix but have poor spatial resolution. The 4D LOGISMOS can be extended to 5D where every time point will have two sequences. The complementary set of information from them can be exploited along with the corresponding images at several follow-up time-points to improve the segmentation accuracy in both types of MRI sequences for all the time-points of the patient.

Most of this thesis focused on identifying surfaces of the bones and cartilages of the femur and tibia to study thickness losses. However there are several other

artifacts such as bone marrow lesions, synovial fluid leakage, meniscus extrusion, and osteophytes which are all indicators of the disease. These artifacts do not have a well-defined structure. The earlier works of Qi et al. [59] combining graph cuts with graph search can be used to identify many of these structures if it was known before hand that they appear in the given MR volume. However most of these artifacts may not even exist in early stage osteoarthritis. The challenge will be to extend the current graph techniques to catch these structures if and when the exist while still working successfully as graph search to identify the surfaces when the artifacts do not exist.

7.3.3 Just Enough Interaction

JEI is proven to be very useful for providing fast interactive corrections. However one of the most time consuming steps of the work-flow is to scroll through the several slices identifying the regions that need correction. JEI could have a learning based step where the algorithm is trained to identify regions that may required editing. It could be argued that if the learning based algorithm could identify the region then that information might as well be incorporated into the cost function design of the automated LOGISMOS. However this specialized learning might differ vastly from the cost functions that are currently providing robust segmentation and may actually worsen the accuracy of the segmentation given the highly specific type of artifact the algorithm would be designed to detect.

Real-time reporting and analysis can be incorporated into the JEI viewer providing quantitative metrics. With every JEI correction these changes can be reflected in real time on the thickness analysis. This feature will prove to be especially useful for 4D JEI where the reporting can graphically show the progression of thickness losses quantitatively. Further smart 3D interactions can be designed for large 3D capable visualization wall where the main challenge would be to move beyond clicking nudge points on a few slices in 2D to providing a truly 3D information that would help

identify and correct the problematic regions quickly.

7.3.4 Analysis

There are several avenues of exploration for analysis. Adding more textural and intensity-based features to the thickness prediction can be explored to understand if it would improve prediction accuracy. Further, a subset of OAI datasets also have bio-chemical markers available at multiple time-points. Combining this information in a meaningful manner may provide more information on the prediction as well as the progression of the disease. Although linear regression provided superior accuracy to support vector machine based regression, other more powerful regression models could be employed to improve the accuracy.

In order to understand the different parameters that trigger the onset of osteoarthritis unsupervised deep learning could be used. After training the network to predict KL grades, the intermediate layer representations can be studied to understand what structures the network utilized to identify differences between the different osteoarthritis stages.

We already covered femoral bone shape change in this work. One immediate extension of the shape analysis work will be to study the tibial bone shape as well and gain insights on how it changes with disease progression. Only a small subset of the entire cohort was used in this thesis work. Running the segmentation and analysis on the entire cohort would be the next step. More meaningful clinical prospective studies can be undertaken applying our quantitative analysis tools to confirm the diagnosis and further our understanding of the efficacy of the developed tools. JEI editing can be done on more patient datasets providing a larger set of highly accurate independent standard which is very useful for validating the existing techniques especially to better understand the true benefits of 4D.

7.4 Contributions

The following highlights the contributions of this thesis:

1. Designed a new hierarchical random forest based classifier combining the neighborhood approximation forests with the clustered random forest classifier on graph nodes whose probability output was used as cartilage cost functions in a LOGISMOS framework – as a result, significantly improving the segmentation accuracy over existing methods.

2. Extended the 3D LOGISMOS to multiple time-point simultaneous segmentation (4D) resulting in improved segmentation accuracy.

3. Designed 3D just-enough interaction for knee MRI using a modular graphical user interface framework for 3D visualization and editing.

4. Extended the 3D just-enough interaction capabilities to 4D allowing simultaneous visualization and interactive editing of multiple time point images and results.

5. Developed a fully automated pipeline combining LOGISMOS with SPHARM-PDM to analyze femoral bone shape to quantify statistically significant local shape differences between two study populations.

6. Designed an automated nomenclature compliant sub-plate detection algorithm to analyze thickness losses in the load bearing regions of the knee.

7. Designed a regression model to accurately predict the cartilage thickness of various sub-plates at a particular time-point using thickness information of previous time-points as training.

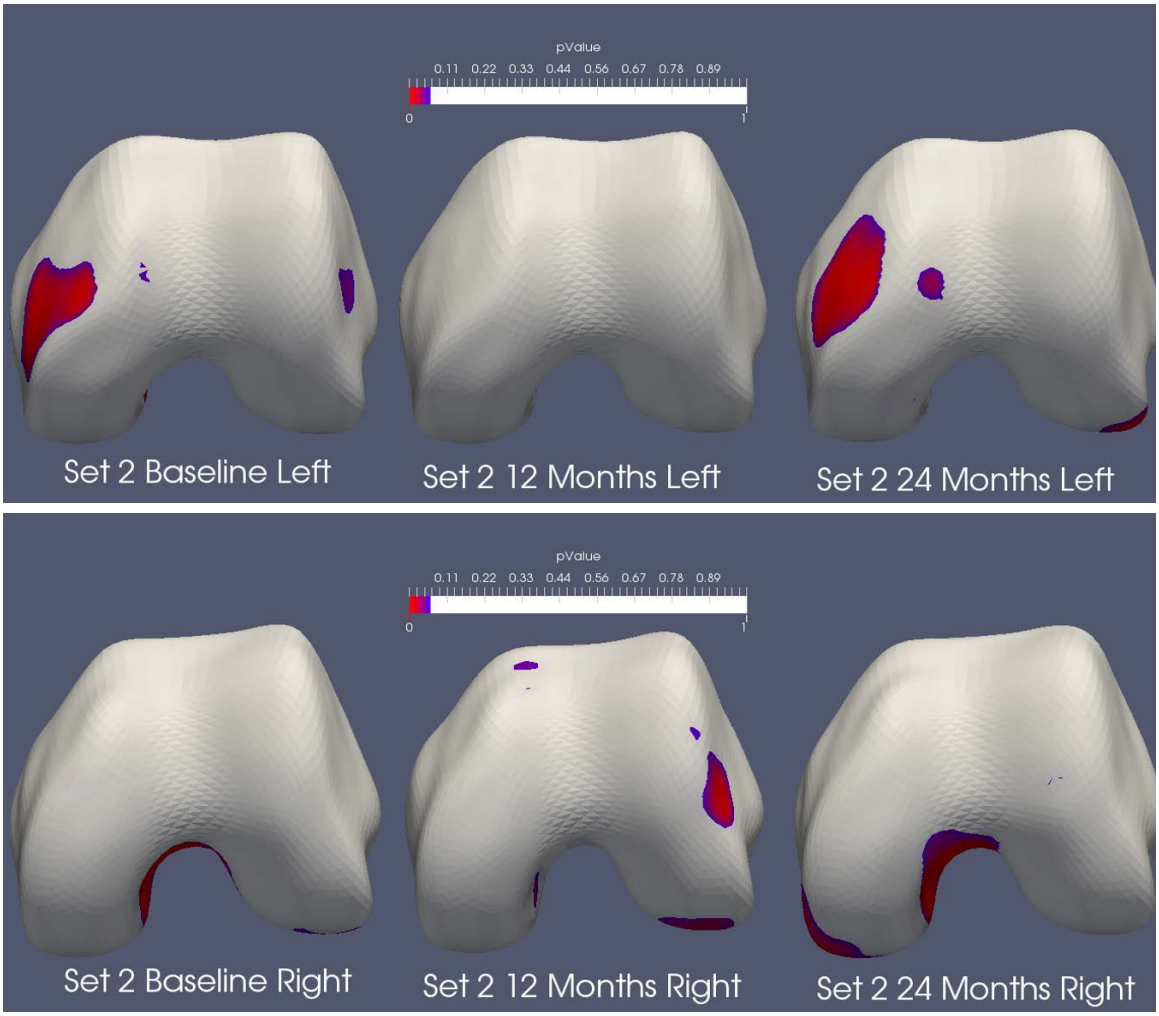


Figure 7.2: Comparison of PROG vs NPROG datasets in set 2. This dataset shows that the results of the shape analysis are dependent on the number of patients. It also highlights the heterogeneity of the KL grading system with different results between the patients of set 1 and set 2 although both start at KL 2.

REFERENCES

[1] "Arthritis Prevalence and Activity Limitations United States, 1990," *JAMA: The Journal of the American Medical Association*, vol. 272, no. 5, p. 346, aug 1994.

[2] "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, aug 2009.

[3] R. D. Altman and G. E. Gold, "Atlas of individual radiographic features in osteoarthritis, revised," *Osteoarthritis and Cartilage*, vol. 15, Supple, no. 0, pp. A1–A56, 2007.

[4] W. A. Barrett and E. N. Mortensen, "Interactive live-wire boundary extraction," *Medical Image Analysis*, vol. 1, no. 4, pp. 331–341, 1997.

[5] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, feb 1992.

[6] M. A. Bowes, G. R. Vincent, C. B. Wolstenholme, and P. G. Conaghan, "A novel method for bone area measurement provides new insights into osteoarthritis and its progression." *Annals of the rheumatic diseases*, vol. 74, no. 3, pp. 519–525, mar 2015.

[7] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary amp; region segmentation of objects in N-D images," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, 2001, pp. 105–112 vol.1.

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[9] R. J. Buck, B. T. Wyman, M.-P. Hellio Le Graverand, M. Hudelmaier, W. Wirth, and F. Eckstein, "Osteoarthritis may not be a one-way-road of cartilage loss comparison of spatial patterns of cartilage change between osteoarthritic and healthy knees," *Osteoarthritis and Cartilage*, vol. 18, no. 3, pp. 329–335, mar 2010.

[10] J. Cates, P. T. Fletcher, M. Styner, M. Shenton, and R. Whitaker, "Shape modeling and analysis with entropy-based particle systems." in *Information processing in medical imaging*, vol. 20. Springer, 2007, pp. 333–45.

[11] F. M. Cicuttini, G. Jones, A. Forbes, and A. E. Wluka, "Rate of cartilage loss at two years predicts subsequent total knee arthroplasty: a prospective study." *Annals of the rheumatic diseases*, vol. 63, no. 9, pp. 1124–1127, sep 2004.

[12] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 525–537, 2002.

[13] L. C. Dijkgraaf, L. G. M. de Bont, G. Boering, and R. S. B. Liem, "The structure, biochemistry, and metabolism of osteoarthritic cartilage: A review of the literature," *Journal of Oral and Maxillofacial Surgery*, vol. 53, no. 10, pp. 1182–1192, oct 1995.

[14] P. Dodin, J. Martel-Pelletier, J.-P. Pelletier, and F. Abram, "A fully automated human knee 3D MRI bone segmentation using the ray casting technique." *Medical & biological engineering & computing*, vol. 49, no. 12, pp. 1413–24, dec 2011.

[15] T. C. Dunn, Y. Lu, H. Jin, M. D. Ries, and S. Majumdar, "T2 Relaxation Time of Cartilage at MR Imaging: Comparison with Severity of Knee Osteoarthritis," *Radiology*, vol. 232, no. 2, pp. 592–598, aug 2004.

[16] F. Eckstein, G. Ateshian, R. Burgkart, D. Burstein, F. Cicuttini, B. Dardzinski, M. Gray, T. M. Link, S. Majumdar, T. Mosher, C. Peterfy, S. Totterman, J. Waterton, C. S. Winalski, and D. Felson, "Proposal for a nomenclature for magnetic resonance imaging based measures of articular cartilage in osteoarthritis." *Osteoarthritis and cartilage*, vol. 14, no. 10, pp. 974–83, oct 2006.

[17] F. Eckstein, R. M. Boudreau, Z. Wang, M. J. Hannon, W. Wirth, S. Cotofana, A. Guermazi, F. Roemer, M. Nevitt, M. R. John, C. Ladel, L. Sharma, D. J. Hunter, and C. K. Kwoh, "Trajectory of cartilage loss within 4 years of knee replacement a nested casecontrol study from the Osteoarthritis Initiative," *Osteoarthritis and Cartilage*, vol. 22, no. 10, pp. 1542–1549, oct 2014.

[18] F. Eckstein, A. Guermazi, G. Gold, J. Duryea, M.-P. H. L. Graverand, W. Wirth, and C. G. Miller, "Imaging of cartilage and bone: promises and pitfalls in clinical trials of osteoarthritis," *Osteoarthritis and Cartilage*, vol. 22, no. 10, pp. 1516–1532, 2014.

[19] F. Eckstein, A. Guermazi, and F. W. Roemer, "Quantitative MR Imaging of Cartilage and Trabecular Bone in Osteoarthritis," *Radiologic Clinics of North America*, vol. 47, no. 4, pp. 655–673, jul 2009.

[20] J. Folkesson, E. Dam B., O. Olsen Fogh, P. Pettersen C., and C. Christiansen, "Segmenting Articular Cartilage Automatically Using a Voxel Classification Approach," *IEEE Transactions on Medical Imaging*, vol. 26, no. 1, pp. 106–115, 2007.

[21] Y. Freund and R. E. Schapire, "A decision theoretic generalization of on-line learning and an application to boosting." *Computer Systems Science*, vol. 57, no. 1, pp. 119–139, aug 1997.

[22] J. Fripp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee." *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 55–64, jan 2010.

[23] A. A. Guccione, D. T. Felson, J. J. Anderson, J. M. Anthony, Y. Zhang, P. W. Wilson, M. Kelly-Hayes, P. a. Wolf, B. E. Kreger, and W. B. Kannel, "The effects of specific medical conditions on the functional limitations of elders in the Framingham Study." *American journal of public health*, vol. 84, no. 3, pp. 351–8, mar 1994.

[24] T. Heimann, I. Oguz, and I. Wolf, "Implementing the automatic generation of 3D statistical shape models with ITK," in *Open Science Workshop*. Citeseer, 2006, pp. 1–8.

[25] D. S. Hochbaum, "A New-Old Algorithm for Minimum-Cut and Maximum-Flow in Closure Graphs," *Networks*, vol. 37, no. 4, pp. 171–193, 2001.

[26] D. J. Hunter, Y. Q. Zhang, X. Tu, M. LaValley, J. B. Niu, S. Amin, A. Guermazi, H. Genant, D. Gale, and D. T. Felson, "Change in joint space width: Hyaline articular cartilage loss or alteration in meniscus?" *Arthritis & Rheumatism*, vol. 54, no. 8, pp. 2488–2495, 2006.

[27] Investigators for the OAI, "Osteoarthritis Biomarkers Consortium FNIH Project : Study Design," Tech. Rep., 2015.

[28] S. Jaume and P. Rondao, "Open topology: A toolkit for brain isosurface correction," in *International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI)*, 2005.

[29] P. M. Jungmann, S. C. Tham, H. Liebl, M. C. Nevitt, C. E. McCulloch, J. Lynch, and T. M. Link, "Association of trochlear dysplasia with degenerative abnormalities in the knee: data from the Osteoarthritis Initiative." *Skeletal radiology*, vol. 42, no. 10, pp. 1383–92, oct 2013.

[30] Kang Li, Xiaodong Wu, D. Chen, and M. Sonka, "Optimal Surface Segmentation in Volumetric Images-A Graph-Theoretic Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119–134, jan 2006.

[31] S. Kashyap, I. Oguz, H. Zhang, and M. Sonka, "Automated Segmentation of Knee MRI Using Hierarchical Classifiers and Just Enough Interaction Based Learning: Data from Osteoarthritis Initiative," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II*, 2016, pp. 344–351.

[32] J. S. Kellgren, J. H.; Lawrence, "Radiological assessment of osteo-arthrosis," *Annals of the rheumatic diseases*, vol. 16, no. 4, pp. 494–502, 1957.

[33] A. Klein, J. Andersson, B. a. Ardekani, J. Ashburner, B. Avants, M. C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, vol. 46, no. 3, pp. 786–802, jul 2009.

[34] E. Konukoglu, B. Glocker, D. Zikic, and A. Criminisi, "Neighbourhood approximation using randomized forests," *Medical image analysis*, vol. 17, no. 7, pp. 790–804, 2013.

[35] M. Kothari, A. Guermazi, G. von Ingersleben, Y. Miaux, M. Sieffert, J. E. Block, R. Stevens, and C. G. Peterfy, "Fixed-flexion radiography of the knee provides reproducible joint space width measurements in osteoarthritis," *European Radiology*, vol. 14, no. 9, pp. 1568–1573, 2004.

[36] R. E. Leach, T. Gregg, and F. J. Siber, "Weight-Bearing Radiography in Osteoarthritis of the Knee," *Radiology*, vol. 97, no. 2, pp. 265–268, nov 1970.

[37] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[38] J. G. Lee, S. Gumus, C. H. Moon, C. K. Kwoh, and K. T. Bae, "Fully automated segmentation of cartilage from the MR images of knee using a multi-atlas and local structural analysis method," *Medical Physics*, vol. 41, no. 9, pp. –, 2014.

[39] S. Lee, S. H. Park, H. Shim, I. D. Yun, and S. U. Lee, "Optimization of local shape and appearance probabilities for segmentation of knee cartilage in 3-D MR images," *Computer Vision and Image Understanding*, vol. 115, no. 12, pp. 1710–1720, dec 2011.

[40] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images – A graph-theoretic approach," {*IEEE*} *Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 119–134, 2006.

[41] C. A. McKenzie, A. Williams, P. V. Prasad, and D. Burstein, "Three-dimensional delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) at 1.5T and 3.0T," *Journal of Magnetic Resonance Imaging*, vol. 24, no. 4, pp. 928–933, 2006.

[42] D. Meier and E. Fisher, "Parameter space warping: Shape-based correspondence between morphologically different objects," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 31–47, 2002.

[43] K. Moisio, F. Eckstein, J. S. Chmiel, A. Guermazi, P. Prasad, O. Almagor, J. Song, D. Dunlop, M. Hudelmaier, A. Kothari, and L. Sharma, "Denuded subchondral bone and knee pain in persons with knee osteoarthritis." *Arthritis and rheumatism*, vol. 60, no. 12, pp. 3703–3710, dec 2009.

[44] J. B. Moseley, K. O'Malley, N. J. Petersen, T. J. Menke, B. A. Brody, D. H. Kuykendall, J. C. Hollingsworth, C. M. Ashton, and N. P. Wray, "A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee," *New England Journal of Medicine*, vol. 347, no. 2, pp. 81–88, jul 2002.

[45] T. Neogi, M. A. Bowes, J. Niu, K. M. De Souza, G. R. Vincent, J. Goggins, Y. Zhang, and D. T. Felson, "Magnetic Resonance Imaging-Based Three-Dimensional Bone Shape of the Knee Predicts Onset of Knee Osteoarthritis: Data From the Osteoarthritis Initiative," *Arthritis & Rheumatism*, vol. 65, no. 8, pp. 2048–2058, aug 2013.

[46] I. Oguz, J. Cates, T. Fletcher, R. Whitaker, D. Cool, S. Aylward, and M. Styner, "Cortical correspondence using entropy-based particle systems and local features," in *2008 5th IEEE International Symposium on Biomedical Imaging From Nano to Macro*. IEEE, 2008, pp. 1637–1640.

[47] D. Pantazis, R. Leahy, T. Nichols, and M. Styner, "Statistical surface-based morphometry using a nonparametric approach," in *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*. IEEE, 2004, pp. 1283–1286.

[48] V. Pedoia, D. Lansdown, M. Zaid, C. McCulloch, R. Souza, C. Ma, and X. Li, "Three-dimensional MRI-based statistical shape model and application to a cohort of knees with acute ACL injury," *Osteoarthritis and Cartilage*, vol. 23, no. 10, pp. 1695–1703, oct 2015.

[49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and Others, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[50] C. W. A. Pfirrmann, M. Zanetti, J. Romero, and J. Hodler, "Femoral Trochlear Dysplasia: MR Findings1," *Radiology*, vol. 216, no. 3, pp. 858–864, 2000.

[51] J.-P. Raynauld, J. Martel-Pelletier, F. Abram, M. Dorais, B. Haraoui, D. Choquette, P. Bias, K. H. Emmert, S. Laufer, and J.-P. Pelletier, "Analysis of the precision and sensitivity to change of different approaches to assess cartilage loss by quantitative MRI in a longitudinal multicentre clinical trial in patients with knee osteoarthritis." *Arthritis research & therapy*, vol. 10, no. 6, p. R129, 2008.

[52] R. R. Regatte, S. V. S. Akella, A. J. Wheaton, G. Lech, A. Borthakur, J. Kneeland, and R. Reddy, "3D-T1ρ-relaxation mapping of articular cartilage: In vivo assessment of early degenerative changes in symptomatic osteoarthritic subjects1," *Academic Radiology*, vol. 11, no. 7, pp. 741–749, jul 2004.

[53] F. W. Roemer, C. K. Kwoh, M. J. Hannon, D. J. Hunter, F. Eckstein, T. Fujii, R. M. Boudreau, and A. Guermazi, "What comes first? Multitissue involvement leading to radiographic osteoarthritis: magnetic resonance imaging-based trajectory analysis over four years in the osteoarthritis initiative." *Arthritis & rheumatology (Hoboken, N.J.)*, vol. 67, no. 8, pp. 2085–2096, may 2015.

[54] J. C. Ross, R. S. J. Estépar, A. D\'\iaz, C.-F. Westin, R. Kikinis, E. K. Silverman, and G. R. Washko, "Lung extraction, lobe segmentation and hierarchical region assessment for quantitative analysis on high res.˜{CT} images," *MICCAI*, pp. 690–698, 2009.

[55] A. Schenk, G. Prause, and H.-O. Peitgen, "Efficient semiautomatic segmentation of {3D} objects in medical images," in *MICCAI*, 2000, pp. 186–195.

[56] T. Schwarz, T. Heimann, R. Tetzlaff, A.-M. Rau, I. Wolf, and H.-P. Meinzer, "Interactive surface correction for {3D} shape based segmentation," in *SPIE*, 2008.

[57] L. Sharma, F. Eckstein, J. Song, A. Guermazi, P. Prasad, D. Kapoor, S. Cahue, M. Marshall, M. Hudelmaier, and D. Dunlop, "Relationship of meniscal damage, meniscal extrusion, malalignment, and joint laxity to subsequent cartilage loss in osteoarthritic knees," *Arthritis and Rheumatism*, vol. 58, no. 6, pp. 1716–26, jun 2008.

[58] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[59] Q. Song, J. Bai, M. K. Garvin, M. Sonka, J. M. Buatti, and X. Wu, "Optimal multiple surface segmentation with shape and context priors," *IEEE transactions on medical imaging*, vol. 32, no. 2, pp. 376–386, 2013.

[60] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, Analysis, and Machine vision*. Cengage Learning, 2014.

[61] C. Stehling, "Scoring Systems to Semiquantitatively Grade Cartilage Pathology with MRI," *Cartilage Imaging*, 2011.

[62] M. Styner, I. Oguz, and S. Xu, "Framework for the statistical shape analysis of brain structures using SPHARM-PDM," *Insight Journal*, vol. m, no. 1071, pp. 242–250, 2006.

[63] S. Sun, M. Sonka, and R. R. Beichel, "Graph-Based IVUS Segmentation with Efficient Computer-Aided Refinement," *IEEE transactions on medical imaging*, vol. 32, no. 8, p. 10.1109/TMI.2013.2260763, aug 2013.

[64] P. M. Thompson, C. Schwartz, and a. W. Toga, "High-resolution random mesh algorithms for creating a probabilistic 3D surface atlas of the human brain." *NeuroImage*, vol. 3, no. 1, pp. 19–34, 1996.

[65] V. N. Vapnik, *Statistical Learning Theory*. New York: J Wiley and Sons, 1998.

[66] Q. Wang, D. Wu, L. Lu, M. Liu, K. L. Boyer, and S. K. Zhou, *Semantic Context Forests for Learning-Based Knee Cartilage Segmentation in 3D MR Images*. Cham: Springer International Publishing, 2014, pp. 105–115.

[67] W. Wirth, O. Benichou, C. K. Kwoh, A. Guermazi, D. Hunter, R. Putz, and F. Eckstein, "Spatial patterns of cartilage loss in the medial femoral condyle in osteoarthritic knees: data from the Osteoarthritis Initiative." *Magnetic resonance in medicine*, vol. 63, no. 3, pp. 574–81, mar 2010.

[68] W. Wirth and F. Eckstein, "A technique for regional analysis of femorotibial cartilage thickness based on quantitative magnetic resonance imaging." *IEEE transactions on medical imaging*, vol. 27, no. 6, pp. 737–44, jun 2008.

[69] X. Wu, D. Chen, and K. Li, "The layered net surface problems in discrete geometry and medical image segmentation," *journal of computational geometry*, 2007.

[70] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka, "LOGISMOS —Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces: Cartilage Segmentation in the Knee Joint," {*IEEE*} *Trans. Med. Imag.*, vol. 29, no. 12, pp. 2023–2037, 2010.

[71] H. Zhang, S. Kashyap, A. Wahle, and M. Sonka, "Highly Modular Multi-Platform Development Environment for Automated Segmentation and Just Enough Interaction," in *Interactive Medical Image Computing: 3th International Workshop, IMIC 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 2016*, 2016.