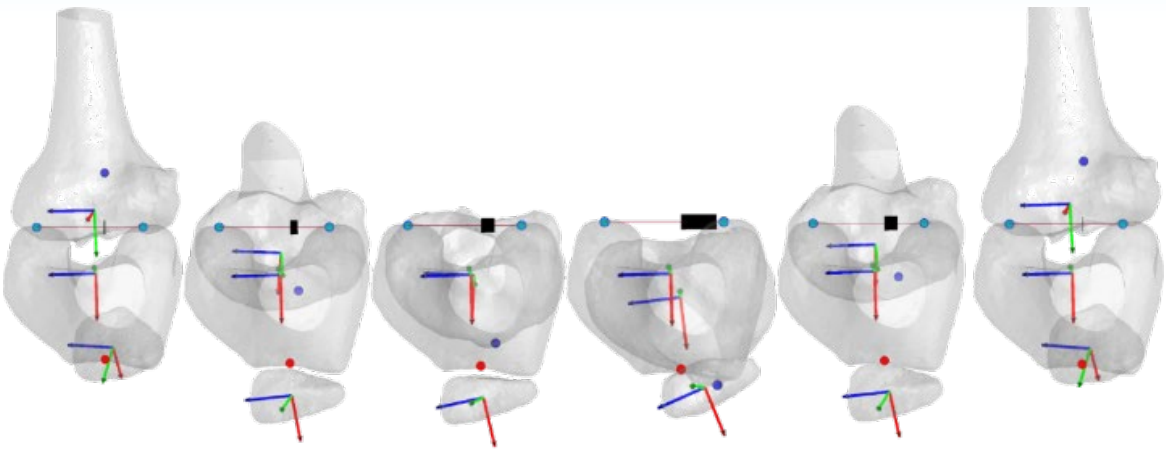
Automatic Segmentation and Motion Analysis of the Knee Joint based on MRI and 4DCT Images

Hao Chen



AUTOMATIC SEGMENTATION AND MOTION ANALYSIS OF THE KNEE JOINT BASED ON MRI AND 4DCT IMAGES

Hao Chen

AUTOMATIC SEGMENTATION AND MOTION ANALYSIS OF THE KNEE JOINT BASED ON MRI AND 4DCT IMAGES

DISSERTATION

to obtain

the degree of doctor at the University of Twente,

on the authority of the rector magnificus,

prof. dr. T.T.M. Palstra,

on account of the decision of the graduation committee,

to be publicly defended

on Thursday 6th February 2020 at 14:45

by

Hao Chen

born on 26th October 1988

in Zhaoqing, China

This dissertation has been approved by:

Supervisors:
prof. dr. ir. N. J. J. Verdonschot
prof. dr. Y. Kang

Co-supervisor:
dr. A. M. J. Sprengers

The work presented in this dissertation was conducted at the Department of Biomechanical Engineering of the University of Twente, and carried out within the BioMechTools project. The project was funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n.323091 and China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW).

Financial support for the publication of this dissertation by the Department of Biomechanical Engineering of the University of Twente is gratefully acknowledged.

Cover: designed by Hao Chen and Jia Li
Printed by: Ipskamp
ISBN: 978-90-365-4947-9
DOI: 10.3990/1.9789036549479
URL: https://doi.org/10.3990/1.9789036549479

Graduation Committee:

Chairman / secretary

prof. dr. G.P.M.R. Dewulf	University of Twente

Supervisors:

prof.dr.ir. N.J.J. Verdonschot	University of Twente
prof. dr. Y. Kang	Northeastern University

Co-supervisor:

dr. A.M.J. Sprengers	Amsterdam University Medical Center

Committee Members:

prof. dr.ir. C. H. Slump	University of Twente
prof. dr.ir. G. J. Verkerke	University of Twente
prof. dr. ir. B. M. ter Haar Romeny	University of Technology Eindhoven
prof. dr. ir. A. J. Nederveen	Amsterdam University Medical Center
dr. T. van Tienen	Radboud University Medical Center

Table of Contents

Summary

Lower limbs disorders, such as osteoarthritis (OA), knee bursitis, meniscal lesions, are musculoskeletal diseases that affect the hips, knees and legs, and usually lead to chronic pain conditions and limited mobility. Knee joint pains can interfere with daily activities, from sports to climbing stairs to walking. In this sense, a relatively minor knee defect can have a great impact on everyday life and create negative conditions for further health deterioration. In the past decades, rapid advances in medical imaging and increased understanding of musculoskeletal mechanics, and as a result diagnosis and effective treatment of knee joint disorders have improved substantially. Nevertheless, the available tools for researchers, radiologists and orthopaedic surgeons to assess the biomechanical condition of patients still lack in delivering a quantitative analysis of muscle and ligament ruptures, cartilage deterioration and patella misalignment, forcing the clinicians to rely mostly on their own subjective judgement which can lead sub-optimal care. One of the main bottlenecks in delivering quantitative digital models of the knee joint is the fast and accurate segmentation of the various knee structures from medical images. Manual segmentation is intrinsically subjective and time-consuming despite the developments in software, and accurate segmentation requires both clinical and user interface experience. As a part of the ERC BioMechTools project, the work described in this thesis is dedicated to advancing existing methods and developing new methods for robust, accurate and maximally automated segmentation and clinical measure determination relevant to pathology assessment in the knee joint.

One of the strengths of MRI is its capability to render different contrast of human tissue. To bridge the gap between clinical MRI protocols developed for fast, reliable acquisition and qualitative assessment and requirements of finite element research, segmentation frameworks for two clinical MRI protocols were developed: Proton Density Weighted and fat suppressed T1/T2 Weighted MRI. Chapter 2 describes a non-supervised 3D local intensity clustering based level set method to determine the bone structure of the knee from PDW weighted MRI. The algorithm first determines the trabecular bone areas using 3D local intensity clustering based level set method, together with correcting any slow varying inhomogeneity in MRI image and then estimates the thin cortical bone boundary with sub-pixel accuracy from intensity profiles along normal vectors originating from the trabecular bone shape. An average dice score of 0.96 was achieved in this study, together with segmentation time of 250 seconds.

In fat suppressed T1/T2 weighted MRI images of the knee, the trabecular bone is translucent, requiring a different approach. In Chapter 3 a deep neural network was employed

to extract the bone structures using the public dataset, made available with the grand challenge SKI10 (http://www.ski10.org/). The extensive comparisons of different networks showed that a simple U-Net architecture with fined tuned hyper-parameters can achieve state-of-the-art results, but still required improvement on specific pathological data sets deviation from the training data. A resampling strategy and the introduction of adversarial losses further improved the bone segmentation by enlarging the contextual information of the network and ensuring the shape consistency. An average dice score of 0.98 and 0.88 for bone and cartilage segmentation respectively was achieved by using the proposed deep learning method, in an inference time of around 60 s, using a Nvidia 1080 Ti graphical card.

Despite the considerable improvement in magnetic resonance imaging techniques of the lower extremity over the last decades, computed tomography (CT) remains an often-preferred image modality for knee joint related disease. Although the bone geometry can be clearly identified by the CT image, cartilage is hardly visible, which prevents providing the diagnostic potential of cartilage status and construction of 3D knee joint models from CT images. In Chapter 4, an alternative method to predict the status of cartilage degeneration based on the bone geometry only using a posterior shape model was proposed. The obtained accuracy of 0.64 mm for femur and 0.58 mm for tibia outperformed the average cartilage thickness distribution method in the tibiofemoral contact area. This provides both clinicians and modelling researchers with more information on the influence of cartilage status on knee joint kinematics, giving the clinician more basis to assess the need for further testing (MRI) and modelling researchers a spatially more resolved distribution of cartilage thickness to optimize their models.

Another focus of this thesis was to automate the knee joint motion analysis, generating clinically related parameters from CT scans. From the various biomechanical parameters described in literature, a fully automated workflow for extraction of the commonly used tibial tubercle – trochlear groove (TT-TG) distance was developed. First, a simulation sensitivity analysis was done on the influence of the alignment between the leg axis and the craniocaudal axis of the CT scanner on TT-TG determination. Chapter 5 showed how minor disagreement between the leg and craniocaudal axis of the CT scanner can have a significant impact on the TT-TG distance calculation (1 mm per degree) and suggested a necessity of pre-correction of the axial cross section before manual measurements.

Subsequently, since the extra works to compensate the deviation of the CT axial plane from the anatomical axial plane is not trivial, an automatic determination of the anatomical coordinate frame (Chapter 6) was designed to automate the correction of disagreement of the planes. Furthermore, according to the experiments in Chapter 5 and Chapter 6, the intra- and inter- observations of the key landmarks of TT-TG distance were significant to clinical

measure outcomes, and the algorithm for the determination of these landmarks was designed. Overall, the proposed semi-automatic workflow provides more robust results and was able to successfully discriminate patients from healthy subjects in terms of previously described normal and pathological TT-TG distance. However, the workflow described in Chapter 6 still requires considerable manual input, hampering the application to larger patient groups and robustness of the method.

To reduce the required manual input, a deep learning based knee joint segmentation was developed for 3D and 4D CT of the knee joint (Chapter 7). A mean DSC for femur, tibia, and patella of 0.98 ± 0.01, 0.97 ± 0.02, and 0.97 ± 0.01, respectively was achieved. Meanwhile, the difference of TT-TG distance between the manual segmentation and automatic segmentation was relatively small (0.3 ± 0.3 mm), compared with the manual intra-observation errors of 1.2 ± 1.0 mm and the manual inter-observation of 2.4 ± 1.8 mm. Overall, a fully automatic workflow is proposed in this thesis capable of accurate and robust assessment of knee joint functioning based on static and a novel dynamic TT-TG distance, validated on an in-house acquired data sets of 4 subjects. The total processing time for extraction of these measures from static and dynamic CT is around 30 minutes.

In conclusion, as a part of the BioMechTools project, several novel image analysis techniques are presented in this thesis for quantitative analysis of the knee joint from MRI and CT images. The automatic segmentation of bone and cartilage from PDW and fat-suppressed T1/T2 weighted MRI advances their diagnostic merit and provides a starting point for subject-specific finite element models. Automatic captures of bone shapes and landmarks from static and dynamic CT provides the possibility of standardised assessment of kinematic knee joint functioning, increasing the diagnostic merit of the TT-TG distance. However, for successful implementation into the clinic, further efforts will have to be made to confirm the clinical findings in a larger patient group and a user-friendly GUI will have to be developed to make these methods available to clinical personnel in an easy to use and straightforward fashion.

Samenvatting

Aandoeningen van de knie, zoals artrose, slijmbeursontsteking of bursitis, meniscus rupturen en kraakbeenletsels zijn musculoskeletale aandoeningen die de heupen, knieen en benen aantasten en dikwijls leiden tot chronische pijn en beperkte mobiliteit. Pijn in de gewrichten kan de dagelijkse activiteiten zoals wandelen, traplopen en sporten verstoren. Daarom kan een relatief kleine pathologie van de knie een grote impact hebben op het dagelijks leven en de kwaliteit van leven negatief beïnvloeden. In de afgelopen decennia heeft de snelle vooruitgang in medische beeldvorming geleid tot een beter begrip van de musculoskeletale biomechanica en als gevolg daarvan een aanzienlijke verbetering in de diagnose en effectieve behandeling van kniegewrichtsaandoeningen teweeggebracht. In de huidige beschikbare verzameling van instrumenten voor onderzoekers, radiologen en orthopedische chirurgen om de biomechanische toestand van de patiënten te evalueren ontbreekt het nog steeds aan een kwantitatieve analyse van spier en –ligament schade, pathologische status van het kraakbeen en patella afwijkingen, waardoor de artsen op hun eigen ervaring blijven aangewezen wat tot suboptimale zorg kan leiden.

Een van de belangrijkste knelpunten bij het leveren van kwantitatieve digitale modellen van het kniegewricht is de snelle en nauwkeurige segmentatie van de verschillende kniestructuren op basis van medische beeldvorming. Handmatige segmentatie is intrinsiek subjectief en tijdrovend, ondanks voortdurende ontwikkelingen in software, en nauwkeurige segmentatie vereist zowel grote klinische ervaring als ervaring met de user interfaces van verscheidene softwarepakketten. Als onderdeel van het ERC BioMechTools project is het werk dat in dit proefschrift wordt beschreven gericht op het verbeteren van bestaande en het ontwikkelen van nieuwe methoden voor robuuste, nauwkeurige en maximaal geautomatiseerde segmentatie alsmede het vertalen van complexe 3D beelden in objectieve kwantitatieve klinische maten die standaardisatie van pathologie in het kniegewricht verder mogelijk maken.

MRI onderscheid zich ten opzichte van andere beeldvormingstechnieken in de veelzijdigheid aan contrasten tussen de verschillende weefsels. Acquisitiesnelheid is de grootste beperkende factor van MRI, en de keuze in het ondervangen hiervan (multi 2D vs 3D, niet isotrope resolutie, half Fourier, etc) vormt een van de oorzaken van de verscheidenheid in MRI protocollen. Eindige elementen modellen ('Finite Element Models') zijn met de in recente jaren toegenomen CPU/GPU rekenkracht er toe in staat om een resolutie van twee a drie ordegroottes meer dan MRI nu (binnen de grove kaders van praktische haalbaarheid) kan leveren, te benutten. Om de kloof te overbruggen tussen klinische MRI-protocollen die zijn ontwikkeld met de nadruk op snelle betrouwbare

acquisitie, en de vereisten voor eindige elementen modellen, zijn in dit proefschrift segmentatie algoritmes ontwikkeld voor een tweetal veel voorkomende klinische MRI-protocollen met een afwijkend contrast: Protondichtheid gewogen contrast ("Proton Density Weighted MRI) en vet onderdrukt T1/T2 gewogen contrast (Fat Supressed T1/T2 weighted MRI).

Hoofdstuk 2 beschrijft het algoritme ontwikkeld voor PDW gewogen contrast. Dit algoritme identificeert eerst het (in dit contrast) hyperintense trabeculaire bot door gebieden van gelijke intensiteit te clusteren en te minimaliseren volgens een level set functie, deze kostfunctie bevat een noise term die gebruikt wordt om eventuele langzaam variererende inhomogeniteit te corrigeren. Vervolgens wordt de dikte van de corticale bot laag geschat met sub-pixel nauwkeurigheid uit intensiteitsprofielen loodrecht op het trabeculaire botoppervlak. Een gemiddelde Dice score van 0,96 werd behaald met dit algoritme, in een gemiddelde segmentatietijd van 250 seconden.

T1 en T2-gewogen MRI-beelden van de knie met vet onderdrukking geven een compleet andere weergave van de knie, en vergt een andere benadering voor segmentatie. Hoofdstuk 3 beschrijft hoe een neuraal netwerk kan worden aangewend om de botstructuren te segmenteren gebruik makend van de SK10 openbare data als trainingsdata (http://www.ski10.org/). De gedetailleerde onderlinge vergelijking tussen verschillende netwerken toont aan dat een eenvoudige U-Net-architectuur met precies afgestemde hyperparameters state-of-the-art resultaten kan bereiken, die echter nog wel onderhevig zijn aan verbetering door uitbreiding van de trainingsdata in termen van volume en diversiteit in pathologie. Data interpolatie en toevoeging van adversariele kosten leidde tot verbetering door de contextuele informatie van het netwerk te vergroten en de vormconsistentie te waarborgen. Een gemiddelde Dice score van respectievelijk 0,98 en 0,88 voor bot- en kraakbeensegmentatie werd bereikt met behulp van de voorgestelde methode, in een inferentie tijd van ongeveer 60 seconden, met behulp van een grafische Nvidia 1080 Ti- kaart.

Ondanks de aanzienlijke verbetering van de MRI beeldvorming van de knie in de laatste decennia, blijft computertomografie (CT) een vaak geprefereerde beeldmodaliteit voor kniegewrichtgerelateerde klachten. Bot geeft een zeer helder CT signaal en de geometrie van het bot kan in hoge resolutie worden gesegmenteerd. Kraakbeen geeft een zeer zwak signaal, en diagnose van de status van kraakbeen of constructie van 3D-kniegewrichtmodellen met kraakbeen uit CT beelden is hierdoor niet mogelijk. In hoofdstuk 4 werd een alternatieve methode voorgesteld om de status van kraakbeendegeneratie te voorspellen op basis van de botgeometrie verkregen uit CT beelden met alleen een posterieur vormmodel. De verkregen nauwkeurigheid van 0,64 mm voor dijbeen en 0,58 mm voor tibia overtrof de gemiddelde kraakbeendikteverdelingsmethode in het tibiofemorale contactgebied. Dit biedt zowel clinici

als onderzoekers in het veld van modellering meer informatie over de invloed van de kraakbeenstatus op de kinematica van het kniegewricht: voor de clinicus aanvullende informatie voor het stellen van een diagnose dan wel inschatten of er extra (MRI) beelden nodig zijn, voor het modellerend onderzoek een verhoogde resolutie in de variatie in kraakbeendikte.

Een andere focus van dit proefschrift was het automatiseren van de bewegingsanalyse van het kniegewricht, met het doel om klinisch gerelateerde parameters te genereren uit de CT-scans. Uit de verschillende biomechanische parameters die in de literatuur zijn beschreven, is een volledig geautomatiseerde workflow ontwikkeld voor de extractie van de veelgebruikte tibiale tuberkel - trochleaire groef (TT-TG) afstand. Hiervoor werd eerst met behulp van een simulatie een sensitiviteits analyse uitgevoerd op de invloed van de uitlijning tussen de beenas en de craniocaudale as van de CT-scanner op de TT-TG-bepaling. Hoofdstuk 5 laat zien dat een klein hoekverschil tussen het been en de craniocaudale as van de CT-scanner een diagnose beinvloedend effect kan hebben op de TT-TG afstandsberekening (1 mm per graad) en hoe een handmatige pre-correctie.

Om een hoge nauwkeurigheid en mate van reproduceerbaarheid te kunnen waarborgen in de correctie in afwijking tussen het axiale CT vlak en het anatomische axiale vlak, is een automatische bepaling van het anatomische coördinatenframe (hoofdstuk 6) ontworpen waarmee afwijking tussen de assen kan worden gecorrigeerd. Experimenten in hoofdstuk 5 en 6 laten verder zien dat de spreiding in handmatige selectie van de oriëntatie punten (anchor points) leidt tot variatie in de bepaalde TT-TG afstand; deze variatie is te groot in klinisch perspectief. Om dit proces efficiënter te maken en de reproduceerbaarheid te verbeteren is een algoritme ontwikkeld om de oriëntatie punten met zo min mogelijk manuele inbreng te bepalen. Over het algemeen biedt de voorgestelde semi-automatische workflow robuustere resultaten en was het mogelijk om patiënten te onderscheiden van gezonde personen op basis van de eerder beschreven normale en pathologische TT-TG-afstanden. Deze workflow zoals beschreven in hoofdstuk 6 vereist echter nog steeds aanzienlijke handmatige invoer, wat vooralsnog een belemmering vormt voor de klinische toepasbaarheid met grotere patiëntengroepen.

Om deze afhankelijkheid van handmatige invoer bij het bepalen van TT-TG afstanden te verminderen, werd een neuraal network voor kniegewricht segmentatie ontwikkeld voor 3D en 4D CT (hoofdstuk 7). Een gemiddelde Dice score voor femur, tibia en patella van respectievelijk 0.98 ± 0.01, 0.97 ± 0.02, en 0.97 ± 0.01 werd respectievelijk bereikt. Het verschil in TT-TG-afstand tussen de handmatige segmentatie en de automatische segmentatie was relatief klein (0.3 ± 0.3 mm) in vergelijking met de handmatige intra-observatiefouten

van 1.2 ± 1.0 mm en de handmatige inter-observatie van 2.4 ± 1.8 mm. De algoritmes beschreven in hoofdstuk 4/5 t/m 7 vormen tezamen een workflow om op nauwkeurige, robuuste en volledig automatische wijze de TT-TG afstand als functie van de knie flexie te verkrijgen uit een 4DCT dataset. Deze workflow is gevalideerd op een in - huis verkregen data set van 4 personen. De totale verwerkingstijd voor het extraheren van deze metingen uit statische en dynamische CT was ongeveer 30 minuten.

Concluderend, als onderdeel van het BioMechTools- project, worden in dit proefschrift verschillende nieuwe beeldverwerkings technieken gepresenteerd ten behoeve van een meer kwantitatieve analyse van MRI en CT beelden van het kniegewricht. De automatische segmentatie van trabeculair en corticaal bot uit MRI met PDW gewogen contrast en kraakbeen uit T1/T2 gewogen contrast met vet onderdrukking bevordert hun diagnostische waarde en verschaft een uitgangspunt voor patiënt specifieke biomechanische modellen van de knie. De workflow for automatische segmentatie van botstructuren en selectie van oriëntatie punten in statische en dynamische CT biedt de mogelijkheid om het kinematisch functioneren van het kniegewricht verder te standaardizeren, en daarmee de diagnostische waarde van de (dynamische) TT-TG afstand substantieel te verhogen. Voor een succesvolle implementatie in de kliniek zullen echter verdere inspanningen moeten worden geleverd om de klinische bevindingen in een grotere patiëntengroep te bevestigen. Bovendien moet een gebruikersvriendelijke GUI worden ontwikkeld om deze methoden beschikbaar te maken voor klinisch personeel in een gemakkelijk te gebruiken en ongecompliceerde wijze.

Chapter 1 Introduction

1

1.1 Background

Musculoskeletal (MS) diseases affect hundreds of millions of people around the world, impacting people's quality of life and independence [1]. According to the report from World Health Organization, musculoskeletal disorders are the second leading cause of disability around the world, and the diseases affect all ages [2]. The resulting limited mobility leads to multimorbidity health states [3], and associates with a high risk of obesity cardiovascular conditions, diabetes and depression [4]. Lower limbs disorders (LLDs), such as osteoarthritis (OA), knee bursitis, meniscal lesions, are MS diseases that affect the hips, knees and legs, and usually lead to chronic pain conditions and limited mobility.

The aetiology of many LLDs is complex and still under research, but the most common risk factors relate to trauma or overuse of the lower limbs, such as repetitive kneeling or squatting, long periods of standing or frequent jumping from a height. Although the exact causes of LLDs are complex, the biomechanical consequences are observable in patients with LLDs, which can be used to discriminate the pathological subjects from healthy ones. Therefore, to develop a tool that can restore the biomechanical parameters of the lower limbs is not only beneficial to the study of the aetiology, but also to the treatment of the LLDs.

Although interest is growing to study of LLDs, including the analysis of aetiology, prevention, and the methodologies to improve treatments and rehabilitation [1], the tools to assess the biomechanical condition of patients used by clinicians and researchers are often crude and subjective leading to non-optimal patient analyses and care. This is because the information from the images is assessed by the human eye, which is not only time consuming but also subject to (human) error and depends heavily on clinical experience.

The work in this thesis aims to develop computer-aided tools to quantify the knee joint anatomy and movement of the knee in an automatic and accurate manner. In the following section, the research framework of this thesis will be briefly discussed (Section 1.2). Two fundamental aspects are considered in this thesis: automatic quantitative description of knee joint movement (Section 1.3) and automatic segmentation of knee joint structure based on medical image data (Section 1.4). Finally, the structure of this thesis is given in Section 1.5.

1.2 Research framework – BioMechTools

The research work presented in this thesis is part of the BioMechTools project (Biomechanical diagnostic, pre-planning and outcome tools to improve musculoskeletal surgery) funded by the European Research Council under the European Union's Seventh Framework Programme (Grant Agreement no. 323091). As Figure 1.1 illustrates, the goal of

the BioMechTools project is to produce advanced diagnostic, pre-planning and outcome tools which allow clinicians and researchers for detailed biomechanical analysis about the abnormal tissue deformations [5], pathological loading of the joints [6], abnormal stresses in the hard and soft tissues [7, 8], and aberrant joint kinematics [9]. The focus in this thesis is to develop a workflow that could automatically quantify the knee joint movement, which involves two fundamental issues in the BioMechTools project: 1) automatic quantification of knee kinematics from medical image data, and 2) automatic segmentation of the knee joint structure from medical image data.

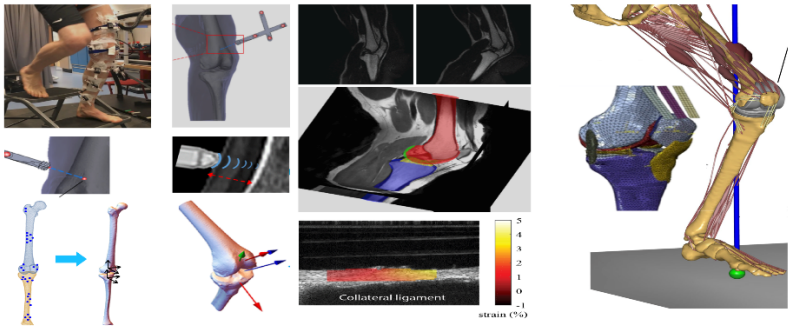


Figure 1.1: Subject-specific modelling approaches in the BioMechTools project.

1.3 Quantification of the knee joint movement

Whereas tissue anatomy or lesions can be identified by the current 3D medical imaging techniques (i.e. computed tomography (CT), Magnetic Resonance Imaging (MRI) and ultrasound), motion-related pathologies are less easily captured by 3D medical imaging modalities, and the biomechanical model of knee joint cannot be created only based on the static data. To acquire the dynamic information of the lower limbs, several techniques are introduced.

1.3.1 Current techniques of capturing the knee joint movement information

(a) Skin markers

Skin markers are widely used in capturing human motion. The principle of skin marker is to track the attached active or reflective markers on the skin by stereo photographic techniques for the trajectory of the movement of the subject. Although the skin markers related system provide a high capturing frequency of the movement, the average rotational

errors were up to 4.4° and 13.1° and the average translational errors were up to 13.0 mm and 16.1 mm because of the soft tissue artefacts [10], which reduces the practical applicability.

(b) Fluoroscopic system

Fluoroscopic systems utilize radiographic images and model based approaches to achieve an accuracy in the order of 1 mm for translations and 1 degree for rotations with the bi-planar fluoroscopic imaging system [11, 12]. However, the radiation exposure of the patients, high-cost cumbersome setup and limited field of view prevent it from a laboratory setting to a clinical setting. Furthermore, the usage of the system relies on the generalization ability of the model based method to reconstruct the bone shape model.

(d) Dynamic MRI

Recently, dynamic MRI is proposed as an alternative technique to image the tibiofemoral motion [9, 13, 14]. During scanning, the subject is asked to flex and extend their knee for about 5 minutes at a frequency of 0.67Hz in order to allow the scanner to sample enough information to perform reconstruction [9]. The obvious advantage of dynamic MRI imaging is that the actual 3D kinematics of the joint can be assessed and the capability of visualization of the soft tissues such as cartilage and muscles during the kinematics. In contrast, the limitations are the limited range of motion and loading, the reduced resolution of the 3D image and the requirement of the repeated flexion and extension tasks during the scan.

(e) Ultrasound based skeletal motion capture

Ultrasound based skeletal motion capture system combines multiple A-mode ultrasound transducers with a conventional motion capture system to capture skeletal motions [15, 16]. With the optical tracking system, the motion information of the leg is captured, and the actual position of the knee bone is calibrated based on A-mode ultrasound information, which can be combined to generate the tibiofemoral kinematics [15, 16]. The proposed method is non-invasive and can be used without the limitation of the FOV, which is apparent with fluoroscopic assessments. Nevertheless, the tracking accuracy depends on the registration method, and so far, the setup is only used in laboratory experiments.

(f) Four dimensional computed tomography

Four dimensional computed tomography (4DCT) is introduced as a technological innovation in the field of CT. It involves a three dimensional computed tomography volume imaged over a period of time [17], which provides the option of capturing continuous knee joint motion. The apparent advantage of 4DCT is that the actual 3D kinematics of the knee joint can directly be extracted from the image. The disadvantage is the limited range of

motion, non-physiological loading conditions, the radiation dose issue and limited field of view (FOV).

1.3.2 Key issues on quantitation of the knee joint movement

The quantitation of the knee joint movement is a quantitative description of the 3D moving relationship among knee joint structures. Several concepts have been proposed to describe the knee joint movement as a function of flexion angle between femur and tibia: distance between tibial tubercle (landmark on tibia) and trochlear groove (landmark on femur) [18]; bisect offset (patellar lateralization relative to the trochlear groove) [18]; patella tilt (angle formed by the lines joining the maximum width of the patella and the posterior condylar line) [18]. Nevertheless, the above concepts are assessed by the human eye, subject to subjective interpretation [18-20], which is time-consuming and depends heavily on clinical experience [21-24]. Therefore, an automatic, objective and robust way to measure knee joint movement is an unmet requirement, and the challenge remains to how robustly and automatically determine:

- The anatomical coordinates of the knee bones;
- The bone landmarks for quantitative description of the knee bone relationship;
- The transformation relationship of the knee joint along with time frames.

1.4 Segmentation framework of the knee joint

To build a specific musculoskeletal model of the knee joint, the first step is to perform the segmentation of the knee joint structure. One of the aims in this study is to perform fully automatic segmentation of the knee joint structure based on MRI and CT. The proposed methods for fully automatic segmentation of the knee joint can be roughly categorised into active contour based, prior knowledge (statistical shape model and atlas) based, and deep learning based approaches.

1.4.1 Active contour based methods

Lorigo et al. were the first to apply active contour based method to segment MRI based knee joint images without the utilisation of training data [25]. The texture information based on vector-valued geodesic snakes with local variance was incorporated into the active contour framework to detect the trabecular bone from other structures. This kind of method to detect regions of interest through evolving contours or surfaces under constraints from a given image has been widely accepted in the segmentation field [26-29]. To include cortical bone, Pang et al. added two forces driven active contour model to segment knee structures with fat-

suppressed MR sequences, which included the directional vector field convolution (DVFC) force and coupled prior shape model [30].

Although the active contour based methods show potential to provide the accurate boundary of knee bone structure, they usually require a good initial contour or are selected as an automatic step in a semi-automatic approach, which requires a lot of user interaction and hence, can be time-consuming and effort-demanding [31]. One of the reasons for the requirement of the initial guess for knee bone segmentation on MRI is the image inhomogeneous issue on MRI image. Furthermore, the complex structure of the knee joint and similar intensity between tissues (muscle and trabecular bone or cortical bone and ligament) also contribute to the difficulty on applying the active contour on the knee joint segmentation on MRI image.

1.4.2 Prior knowledge based methods

(a) Statistical shape model

Model based methods attempt to deform the prior statistical shape model (SSM) to the target image object. Fripp et al. utilised the active shape model to segment the bone first and then detected the cartilage edge based on intensity profiles from vertices of the bone mesh [32]. Seim et al. also attempted to deform the statistical shape model to the target bone, and then reconstruct the cartilage based on thickness variation model [33]. Vincent et al. constructed an active appearance model for cartilage and optimised the model parameters to match the image object [34].

(b) Atlas based methods

The atlas based methods attempt to deform the selected atlas objects in order to better register them with the corresponding objects in the target image to be segmented. Lee et al. proposed multiple atlas building and select the best match to the target image after non-rigid registration, after which a local adjustment is applied to optimize the cartilage segmentation [35]. To guarantee the separation of femoral and tibial cartilage, Shan et al. used a three-label segmentation after the atlas based registration [36].

Although encouraging results have been achieved by the model based and atlas based approaches, they might perform poorly when there are high subject variability and significant differences of local features [31, 37]. Furthermore, the statistical shape model based approaches rely on a prior-knowledge of knee bone shape model and require a lot of computation to construct the model, such as rigid, non-rigid registration to find the correspondence [38]. All in all, the generalization ability of the prior knowledge based

method is still an interesting topic in the knee joint area, especially for the case with the limited training sample.

1.4.3 Deep learning based methods

Recently, convolutional neural networks (CNNs) based method had achieved enormous success in biomedical imaging problems, such as classification [39], segmentation [40-43], etc. Regarding knee joint structure segmentation, Prasoon et al. firstly applied 2D CNNs to classify each voxel as the background or tibial cartilage [44]. They applied the 2D CNN along the axial, coronal and sagittal plane independently, and then concatenated to a joint output map feeding into the soft-max classifier to determine the class of voxel. Nevertheless, Ronneberger et al. pointed out the traditional CNN method for segmentation has two drawbacks, large redundancy and a trade-off between localization accuracy and the use of context, and proposed a dense prediction network without any fully connected layers, U-Net [43]. In 2017, Liu et al. attempted to use 2D encoder and decoder network, SegNet [45] and U-Net [43] to perform multi-class tissue classification and then used a deformable model to obtain a smoothed mask [31]. In 2018, Ambellan et al. further optimized the 2D CNN to 3D CNN based method and refined the bone segmentation result with statistical shape model (SSM) [36].

Although very good results were achieved in the past few years based on the deep learning based method, the previous approaches require an optimisation step to refine the bone mask, such as active contour [31] or the statistical shape model [37]. This is because the proposed neural network does not consider the shape consistency. Hence, to build up an end to end network is still an unmet challenge.

1.4.4 Other methods

In addition to the above methods, supervised based method are also proposed in the knee joint segmentation task. Folkesson et al. [46] proposed a k nearest neighbour (kNN) to perform tissues classification based on multiple features, such as position in the image, the eigenvalue of Hessian, and so on. To enhance the robustness of the method, the cartilage of the target subject needs to be shifted to match the cartilage centre of the training dataset. Similarly, Zhang et al. proposed a support vector machine (SVM) with discriminative random field (DRF) approach to segment femoral, tibial and patellar cartilage with similar features setting as [46], but based on multi-contrast MR images [47]. Although above supervised based methods provide alternatives to perform fully automatic segmentation of the knee joint structure, the position matching step may cause large errors, especially with the subject-specific variance.

1.4.5 Key issues on segmentation of knee joint

In the past decades, numerical approaches were proposed to automate the segmentation tasks of the knee joint and have achieved much success on the application to MRI image. The main challenges in terms of segmentation of the knee joint in MRI image remain:

- Inhomogeneous issue in the MRI image;
- High subject variability;
- Unexpected image artefacts;
- Generalization ability of the segmentation framework (different scanner or different sequence)

1.5 Aim and outline of the thesis

The main goal of this thesis is to quantify the knee joint movement in an automatic and robust manner. Two major issues are included: automatic segmentation of the knee joint structure and automatic workflow to quantify the knee joint kinematics. The following research results are shown in the various chapters:

Chapter 2 presents a level set based segmentation method on a clinical used MRI sequence (proton density weighted). In this study, we discussed the robustness of this method on MRI knee bone segmentation in regardless of the inhomogeneity problem.

Chapter 3 presents a deep neural network to perform knee bone and cartilage segmentation on a public dataset (Grand Challenge 2010 [48]). In this study, we introduced the adversarial loss into the network to improve the bone segmentation and proposed to segment the bone and cartilage with different neural networks and different loss functions.

Chapter 4 presents a posterior shape model based method to estimate the cartilage thickness. In this work, we discussed the feasibility of the generation of cartilage only based on knee bone geometry, which might be used to predict cartilage thickness solely from CT images.

Chapter 5 presents a clinically relevant study on the effect of small axial orientation mismatch between the subject and CT scanner on the tibial tubercle -trochlear groove (TT-TG) distance. Furthermore, we discussed the possible solutions concerning this issue.

Chapter 6 presents a semi-automatic workflow to quantitatively assess the patellofemoral instability on 4DCT. In this work, we extended the spatial related clinical related parameters (TT-TG distance and patellar centre – trochlear groove (PC-TG) distance) to spatial and time-

related parameters and demonstrated the feasibility of the automatic calculation of these clinical parameters by utilising a knee bone shape model.

Chapter 7 presents a fully automatic segmentation method to perform knee bone segmentation on both three-dimensional computed tomography (3DCT) and four-dimensional computed tomography (4DCT). Furthermore, we demonstrated the technical feasibility of fully automating the workflow of Chapter 6 by replacing the manual segmentation of the automatic segmentation results. The work in this study also showed an example of how the computer-aided diagnosis toolbox can aid the clinicians and researchers in their assessment of knee pathologies.

Lastly, **Chapter 8** presents a general discussion, where the key findings of the proposed segmentation methods are summarized, and the impact of the automatic workflow concerning motion description of the knee joint on the clinical application are also discussed. In addition, future research is suggested for further improvements of the developed toolbox towards the application to clinical practice.

References

[1] S. J. Dixon and P. Chidiac, "Editorial overview: Musculoskeletal: New therapeutic targets and delivery systems," *Current Opinion in Pharmacology,* vol. 40, pp. ix-xi, 2018, doi: 10.1016/j.coph.2018.06.002.

[2] W. H. Organization., "Musculoskeletal conditions – Fact sheet," 2018.

[3] S. J. Duffield *et al.*, "The contribution of musculoskeletal disorders in multimorbidity: Implications for practice and policy," *Best Practice and Research: Clinical Rheumatology,* vol. 31, pp. 129-144, 2017, doi: 10.1016/j.berh.2017.09.004.

[4] A. M. Briggs *et al.*, "Musculoskeletal Health Conditions Represent a Global Threat to Healthy Aging: A Report for the 2015 World Health Organization World Report on Ageing and Health," *Gerontologist,* vol. 56, pp. S243-S255, 2016, doi: 10.1093/geront/gnw002.

[5] H. Naghibi Beidokhti, D. Janssen, S. van de Groes, J. Hazrati, T. Van den Boogaard, and N. Verdonschot, "The influence of ligament modelling strategies on the predictive capability of finite element models of the human knee joint," *Journal of Biomechanics,* vol. 65, pp. 1-11, 2017, doi: 10.1016/j.jbiomech.2017.08.030.

[6] M. A. Marra *et al.*, "Flexing and downsizing the femoral component is not detrimental to patellofemoral biomechanics in posterior-referencing cruciate-retaining total knee arthroplasty," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 26, pp. 3377-3385, 2018, doi: 10.1007/s00167-018-4900-z.

[7] H. Naghibi *et al.*, "A noninvasive MRI based approach to estimate the mechanical properties of human knee ligaments," *Journal of the Mechanical Behavior of Biomedical Materials,* vol. 93, pp. 43-51, 2019, doi: 10.1016/j.jmbbm.2019.01.022.

[8] K. Gijsbertse *et al.*, "Computer-aided detection of fasciculations and other movements in muscle with ultrasound: Development and clinical application,"

Clinical Neurophysiology, vol. 129, pp. 2567-2576, 2018, doi: 10.1016/j.clinph.2018.09.022.

[9] V. Mazzoli *et al.*, "Accelerated 4D self-gated MRI of tibiofemoral kinematics," *NMR in Biomedicine,* vol. 30, pp. 1-11, 2017, doi: 10.1002/nbm.3791.

[10] D. L. Benoit, D. K. Ramsey, M. Lamontagne, L. Xu, P. Wretenberg, and P. Renström, "Effect of skin movement artifact on knee kinematics during gait and cutting motions measured in vivo," *Gait and Posture,* vol. 24, pp. 152-164, 2006, doi: 10.1016/j.gaitpost.2005.04.012.

[11] S. Guan, H. Gray, F. Keynejad, and M. Pandy, "Mobile Biplane X-ray Imaging System for Measuring 3D Dynamic Joint Motion during Overground Gait," *IEEE Transactions on Medical Imaging,* vol. 0062, pp. 1-1, 2015, doi: 10.1109/TMI.2015.2473168.

[12] G. Li, S. K. Van de Velde, and J. T. Bingham, "Validation of a non-invasive fluoroscopic imaging technique for the measurement of dynamic knee joint motion," *Journal of Biomechanics,* vol. 41, pp. 1616-1622, 2008, doi: 10.1016/j.jbiomech.2008.01.034.

[13] J. Kaiser, R. Bradford, K. Johnson, O. Wieben, and D. G. Thelen, "Measurement of 3D Tibiofemoral Kinematics using Volumetric SPGR-VIPR Imaging," *Magn Reson Med,* vol. 69, pp. 1310-1316, 2013, doi: 10.1002/mrm.24362.Measurement.

[14] E. C. Clarke, J. H. Martin, a. G. d'Entremont, M. G. Pandy, D. R. Wilson, and R. D. Herbert, "A non-invasive, 3D, dynamic MRI method for measuring muscle moment arms in vivo: Demonstration in the human ankle joint and Achilles tendon," *Medical Engineering & Physics,* vol. 37, pp. 93-99, 2015, doi: 10.1016/j.medengphy.2014.11.003.

[15] K. Niu *et al.*, "In situ comparison of A-mode ultrasound tracking system and skin-mounted markers for measuring kinematics of the lower extremity," *Journal of Biomechanics,* vol. 72, pp. 134-143, 2018, doi: 10.1016/j.jbiomech.2018.03.007.

[16] K. Niu, J. Homminga, V. Sluiter, A. Sprengers, and N. Verdonschot, "Measuring relative positions and orientations of the tibia with respect to the femur using one-channel 3D-tracked A-mode ultrasound tracking system: A cadaveric study," *Medical Engineering and Physics,* vol. 57, pp. 61-68, 2018, doi: 10.1016/j.medengphy.2018.04.015.

[17] Y. Kwong, A. O. Mel, G. Wheeler, and J. M. Troupis, "Four-dimensional computed tomography (4DCT): A review of the current status and applications," *Journal of Medical Imaging and Radiation Oncology,* vol. 59, pp. 545-554, 2015, doi: 10.1111/1754-9485.12326.

[18] M. J. Tanaka, J. J. Elias, A. A. Williams, J. A. Carrino, and A. J. Cosgarea, "Correlation Between Changes in Tibial Tuberosity-Trochlear Groove Distance and Patellar Position During Active Knee Extension on Dynamic Kinematic Computed Tomographic Imaging," *Arthroscopy : the journal of arthroscopic & related surgery,* vol. 31, pp. 1748-1755, 2015, doi: 10.1016/j.arthro.2015.03.015.

[19] M. J. Tanaka, J. J. Elias, A. A. Williams, S. Demehri, and A. J. Cosgarea, "Characterization of patellar maltracking using dynamic kinematic CT imaging in patients with patellar instability," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 24, pp. 3634-3641, 2016, doi: 10.1007/s00167-016-4216-9.

[20] A. Myronenko and X. Song, "Point Set Registration: Coherent Point Drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 2262-2275, 2010, doi: 10.1109/TPAMI.2010.46.

[21] L. Yao, N. Gai, and R. D. Boutin, "Axial scan orientation and the tibial tubercle-trochlear groove distance: Error analysis and correction," *American Journal of Roentgenology,* vol. 202, pp. 1291-1296, 2014, doi: 10.2214/AJR.13.11488.

[22] K. Tensho *et al.*, "What Components Comprise the Measurement of the Tibial Tuberosity-Trochlear Groove Distance," *J Bone Joint Surg Am,* vol. 97, pp. 1441-1448, 2015, doi: http://dx.doi.org/10.2106/JBJS.N.01313.

[23] V. R. Carlson, F. T. Sheehan, A. Shen, L. Yao, J. N. Jackson, and B. P. Boden, "The Relationship of Static Tibial Tubercle–Trochlear Groove Measurement and Dynamic Patellar Tracking," *The American Journal of Sports Medicine,* vol. 45, p. 036354651770011, 2017, doi: 10.1177/0363546517700119.

[24] C. M. Anley, G. V. Morris, A. Saithna, S. L. James, and M. Snow, "Defining the Role of the Tibial Tubercle-Trochlear Groove and Tibial Tubercle-Posterior Cruciate Ligament Distances in the Work-up of Patients With Patellofemoral Disorders," *American Journal of Sports Medicine,* vol. 43, pp. 1348-1353, 2015, doi: 10.1177/0363546515576128.

[25] L. M. Lorigo, O. Faugeras, W. E. L. Grimson, R. Keriven, and R. Kikinis, Berlin, Heidelberg. Medical Image Computing and Computer-Assisted Intervention --- MICCAI'98: First International Conference Cambridge, MA, USA, October 11--13, 1998 Proceedings.

[26] H. Yousefi, M. Fatehi, M. Amian, and R. a. Zoroofi, "A fully automated segmentation of radius bone based on active contour in wrist MRI data set," *2013 20th Iranian Conference on Biomedical Engineering (ICBME),* pp. 42-47, 2013, doi: 10.1109/ICBME.2013.6782190.

[27] T. Liu, H. Xu, W. Jin, Z. Liu, Y. Zhao, and W. Tian, "Medical Image Segmentation Based on a Hybrid Region-Based Active Contour Model," *Computational and Mathematical Methods in Medicine,* vol. 2014, pp. 1-10, 2014, doi: 10.1155/2014/890725.

[28] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition,* vol. 43, pp. 1199-1206, 2010, doi: 10.1016/j.patcog.2009.10.010.

[29] A. Aprovitola and L. Gallo, "Knee bone segmentation from MRI: A classification and literature review," in *Biocybernetics and Biomedical Engineering* vol. 36, ed, 2016, pp. 437-449.

[30] J. Pang, J. B. Driban, T. E. McAlindon, J. G. Tamez-Peña, J. Fripp, and E. L. Miller, "On the Use of Coupled Shape Priors for Segmentation of Magnetic Resonance Images of the Knee," *IEEE Journal of Biomedical and Health Informatics,* vol. 19, pp. 1153-1167, 2015, doi: 10.1109/JBHI.2014.2329493.

[31] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, and R. Kijowski, "Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging," *Magnetic Resonance in Medicine,* vol. 79, pp. 2379-2391, 2018, doi: 10.1002/mrm.26841.

[32] 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, pp. 55-64, 2010, doi: 10.1109/TMI.2009.2024743.

[33] H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, "Model-based auto-segmentation of knee bones and cartilage in MRI data," *MICCAI Workshop Medical Image Analysis for the Cinic,* pp. 215-223, 2010.

[34] G. Vincent, C. Wolstenholme, I. Scott, and M. Bowes. Fully Automatic Segmentation of the Knee Joint using Active Appearance Models.

[35] 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, p. 092303, 2014, doi: 10.1118/1.4893533.

[36] L. Shan, C. Zach, C. Charles, and M. Niethammer, "Automatic atlas-based three-label cartilage segmentation from MR knee images," *Medical Image Analysis,* vol. 18, pp. 1233-1246, 2014, doi: 10.1016/j.media.2014.05.008.

[37] F. Ambellan, A. Tack, M. Ehlke, and S. Zachow, "Automated Segmentation of Knee Bone and Cartilage combining Statistical Shape Knowledge and Convolutional Neural Networks: Data from the Osteoarthritis Initiative," *Medical Image Analysis,* vol. 52, pp. 109-118, 2018, doi: 10.1016/j.media.2018.11.009.

[38] T. Albrecht, M. Lüthi, T. Gerig, and T. Vetter, "Posterior shape models," *Medical Image Analysis,* vol. 17, pp. 959-973, 2013, doi: 10.1016/j.media.2013.05.010.

[39] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Transactions on Medical Imaging,* vol. 35, pp. 1299-1312, 2016, doi: 10.1109/TMI.2016.2535302.

[40] K. Kamnitsas, C. Ledig, and V. F. J. Newcombe, "Efficient Multi-Scale 3D CNN with fully connected CRF for Accurate Brain Lesion Segmentation," vol. 36, pp. 61-78, 2017.

[41] D. M. Vigneault, W. Xie, C. Y. Ho, D. A. Bluemke, and J. A. Noble, "Ω-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks," *Medical Image Analysis,* vol. 48, pp. 95-106, 2018, doi: 10.1016/j.media.2018.05.008.

[42] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation.

[43] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation.

[44] A. Prasoon and E. Al., "Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolution Neural Network," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* vol. 8150, pp. 599-606, 2013, doi: 10.1007/978-3-642-40763-5.

[45] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, pp. 2481-2495, 2017, doi: 10.1109/TPAMI.2016.2644615.

[46] J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen, "Segmenting Articular Cartilage Automatically Using a Voxel Classification Approach," *IEEE Transactions on Medical Imaging,* vol. 26, pp. 106-115, 2007, doi: 10.1109/TMI.2006.886808.

[47] K. Zhang, W. Lu, and P. Marziliano, "Automatic knee cartilage segmentation from multi-contrast MR images using support vector machine classification with spatial dependencies," *Magnetic Resonance Imaging,* vol. 31, pp. 1731-1743, 2013, doi: 10.1016/j.mri.2013.06.005.

[48] T. Heimann and B. Morrison, "Segmentation of knee images: A grand challenge," 2010.

Chapter 2 Automated segmentation of trabecular and cortical bone from proton density weighted MRI of the knee

Abstract:

Introduction: Patient-specific implant design and pre- and intra-operative planning is becoming increasingly important in the orthopaedic field. For clinical feasibility of these techniques, fast and accurate segmentation of bone structures from MRI is essential. However, manual segmentation is time-intensive and subject to inter- and intra-observer variation. The challenge in developing automatic segmentation algorithms for MRI data mainly exists in the inhomogeneity problem and the low contrast among cortical bone and adjacent tissues.

Methods: In this study, we proposed a method for automatic segmentation of knee bone structures for MRI data with a 3D local intensity clustering-based level set and a novel approach to determine the cortical boundary utilising the normal vector of the trabecular surface.

Results: Application to clinical imaging data shows that our method is robust to MRI inhomogeneity. In comparing our method to manual segmentation in 18 femurs and tibiae, we found a dice similarity coefficient (DSC) of 0.96 ± 0.01 for the femurs and 0.96 ± 0.02 for tibiae. The average surface distance error was 0.46 ± 0.14 mm for the femurs and 0.47 ± 0.21 mm for the tibiae.

Conclusions: The results of the automatic technique strongly corresponded to the manual segmentation using less than 3% of the time and with virtually no workload.

2.1 Introduction

Fast and accurate segmentation of knee bone structures from MRI data is a topic of increasing interest as its applications continue to broaden from direct diagnostic purposes [1] to the creation of 3D finite element models [2], optimising implant design [3] and pre- and intra-operative planning [4]. However, accurate automated segmentation is hampered by two problems:

1. Intensity inhomogeneity due to MRI inherent problems (coil sensitivity and B1 inhomogeneity) can cause a slow varying intensity gradient as can be noticed from the difference in brightness of trabecular fat in the two white boxes in Figure 2.1)

2. Low contrast between the structures of interest (trabecular bone and infrapatellar fat (green box in Figure 2.1), cortical bone and ligament (blue box in Figure 2.1)).

Multiple approaches have been used for knee joint segmentation such as thresholding, region growing, deformable models, clustering methods and atlas-guided approaches [5]. In 2010, several automated segmentation methods were assessed in the grand challenge competition for segmenting cartilage and bone in knee MRI data [6]. Prior knowledge-based models, such as statistical shape models and atlas-based methods, seemed to outperform pixel-based methods [7]. These methods, however, require data set training and may be less suitable for pathologies that are not incorporated in the training data. Hence, an alternative method to segment the image without training may be challenging but is desirable from both clinical and research perspective.

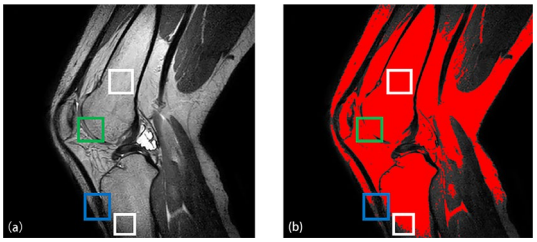


Figure 2.1: (a) Knee MRI image in proton density weighted (PDW) sequence. (b) Segmentation with Ostu's method (red mask is the segmentation mask). (White box represents the inhomogeneity problem; green box represents the issue of low contrast between trabecular bone and infrapatellar fat and blue box represents the low contrast problem between cortical bone and ligament)

To the best of our knowledge, Lorigo et al. were the first to apply active contours to segment MRI-based knee joint images [8] without utilisation of training data. The texture information based on vector-valued geodesic snakes with local variance was incorporated into the active contour framework to detect the trabecular bone from other structures. This

kind of method to detect regions of interest through evolving contours or surfaces under constraints from a given image has been largely accepted in the segmentation field [5, 9-11]. To include cortical bone, Pang et al. added two forces driven active contour model to segment knee structures with fat-suppressed MR sequences, which included the directional vector field convolution (DVFC) force and coupled prior shape model [12]. Furthermore, Dodin et al. proposed a ray casting technique to detect the femur and tibia boundary slice by slice in a sagittal direction [13]. Shan et al. proposed a multi-atlas- based method to extract the femur and tibia mask [14].

Although many attempts were put into automating these segmentations, problems in MRI inhomogeneity and weak edges remain challenging, especially for an effective way to estimate the boundary between the cortical bone and adjacent tissue with similarly low intensity, i.e. ligament.

In this paper, we proposed an automatic segmentation for the trabecular and cortical bone of the femur and tibia in a clinically relevant MR sequence, proton density weighted contrast [15]. We used a local energy-based level set method to obtain a 3D rough segmentation of trabecular bone and correct the image data from the inhomogeneity problem. Subsequently, we generated intensity lines slice by slice based on the rough trabecular masks. Then, we optimized the trabecular boundary based on the intensity lines and propose an iterative process to detect the cortical boundary.

The remainder of this paper is organized as follows: In Section 2, we introduce the MRI data, the segmentation pipeline and related methods. In Section 3, we compare the results of manual segmentation and segmentation using our method on proton density weighted MRI data. Finally, Sections 4 and 5 summarise this research, review the results in comparison to other studies and discuss future plans.

2.2 Materials and Methods

2.2.1 MRI image acquisition

A proton density weighted (PDW) contrast MRI sequence was chosen in this study as it is frequently used to assess pathologies of the knee in a clinical setting [15, 16]. The PDW contrast provides data in which the ligaments, menisci and cartilage can be simultaneously assessed for diagnosis with a reasonable scanning time. In the PDW contrast, all relevant structures are displayed in different intensities ranging from high to low: fatty tissue (i.e. both infrapatellar fat and trabecular bone), cartilage, muscle, ligament and virtually no signal in the cortical bone. Scans were acquired with an eight-channel rigid coil in a 3.0-T Philips scanner. Further sequence details are as follows: $\mathrm{FOV} = 200 \times 200 \times 200\,\mathrm{mm}$, voxel

size = 0.35 × 0.35 × 0.52 mm for six of the data sets and 0.60 × 0.60 × 0.90 mm for the other 12, flip angle = 90, TR/TE = 1000/32.18 ms and scan duration = 6 min. All data were interpolated to 0.90 × 0.90 × 0.90 mm. To test the robustness of the proposed segmentation pipeline, a total of 18 data sets were used in this study. This study was approved by the local IRB and written informed consent was provided by all subjects prior to the study.

2.2.2 Segmentation pipeline

Figure 2.2 shows a schematic representation of the pipeline for the proposed automated method, which includes 3D local intensity clustering-based level set (3DLICLS), inhomogeneity correction, generation of 2D intensity line image along the normal vectors of the rough surface, trabecular mask optimization and cortical mask detection. Also, the required input data, intermediate data and corresponding output data are described in Figure 2.2.

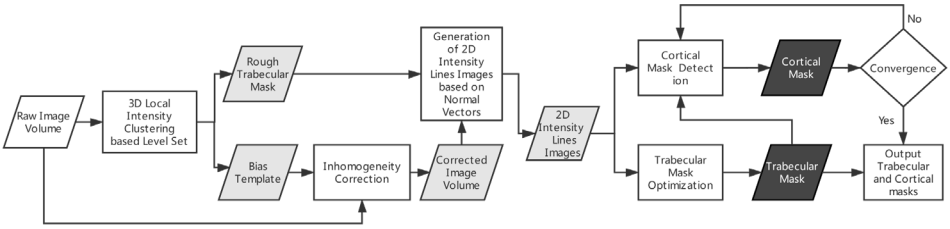


Figure 2.2: Segmentation pipeline for the proposed method. Rectangular boxes represent applied methods. White, grey and black parallelograms represent input data, intermediate data and output results, respectively

2.2.3 Rough segmentation of the trabecular surface and inhomogeneity correction

In this section, we introduce a local energy-based level set method to both obtain a rough segmentation of the trabecular boundary and a bias field to remove any inhomogeneity from the data.

2.2.3.1 3D local intensity clustering-based level set

In 2011, Li et al. proposed a local intensity clustering framework to segment the region of interest simultaneously with solving the inhomogeneity problem [17]. We extended this method to 3D in this study. Suppose the observed volume is V:

$$V = bJ + I_{noise} \tag{2.1}$$

where J represents the actual 3D volume components; b is the 3D bias field, which accounts for the intensity inhomogeneity of the image volume and is slowly varying; and

I_{noise} is the Gaussian noise with zero mean. We proposed to use the mean shift filter [18] to reduce the noise influence in this study, which leads to the model becoming $V = bJ$.

Based on the model, the essential ideas of 3DLICLS to segment interested object in the image with intensity inhomogeneity are introducing a kernel function to define local energy function and introducing a bias variable to define the inhomogeneity template as follows:

$$E(C) = \int_{\Omega}\int_{inside(C)} K_{\sigma}(x-y)|V(x) - b(y)c_{inside}|^2 dydx + \int_{\Omega}\int_{outside(C)} K_{\sigma}(x-y)|V(x) - b(y)c_{outside}|^2 dydx \tag{2.2}$$

where $V: \Omega \in R$ is an input volume, $x, y \in \Omega$, K_{σ} is a Gaussian kernel with standard deviation σ, c_{inside} and $c_{outside}$ represent the constant intensity inside and outside the contour C (such as dark green and light green in Figure 2.3a1), respectively, and b is the inhomogeneity template. The reason for introducing the kernel function is to calculate the energy based on local information, while the reason for introducing the bias variable is to detect the target in the situation with inhomogeneity.

According to level set theory, contour, $C \in \Omega$, can be represented by the zero level set of a Lipschitz function $\phi : \Omega \in R$ [19]. To minimise the cost function E with respect to ϕ, the gradient descent method is applied, $\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi}$, and thus, we can obtain the curve evolution equation as:

$$\frac{\partial \phi}{\partial t} = -\delta_{\varepsilon}(\phi)(e_1 - e_2) \tag{2.3}$$

In order to stabilise the evolution of the level set function, a distance regularised term [20] is incorporated into (3). Furthermore, the Euclidean length term is included to regularise the zero contour of ϕ. Finally, the final evolution equation is as follows:

$$\frac{\partial \phi}{\partial t} = -\delta_{\varepsilon}(\phi)(e_1 - e_2) + v\delta_{\varepsilon}(\phi) div\left(\frac{\nabla \phi}{|\nabla \phi|}\right)^{+} \mu\left(\nabla^2 \phi - div\left(\frac{\nabla \phi}{|\nabla \phi|}\right)\right) \tag{2.4}$$

In (2.4),

$$\begin{cases} e_1(x) = \int_{\Omega} K_{\sigma}(x-y)|V(x) - b(y)c_{inside}|^2 dy \\ e_2(x) = \int_{\Omega} K_{\sigma}(x-y)|V(x) - b(y)c_{outside}|^2 dy \end{cases} \tag{2.5}$$

During the evolution, the representatives of constant and the bias template must be updated. Based on the above assumption model, V could approximately be expressed as the multiplication of b and constant c, and thus, the updated form of c_{inside} and $c_{outside}$ are as follows:

$$\begin{cases} c_{inside} = \dfrac{\int (b * K_\sigma) V H_\varepsilon(\phi) dy}{\int (b^2 * K_\sigma) H_\varepsilon(\phi) dy} \\ c_{outside} = \dfrac{\int (b * K_\sigma) V\big(1 - H_\varepsilon(\phi)\big) dy}{\int (b^2 * K_\sigma)(1 - H_\varepsilon(\phi)) dy} \end{cases} \tag{2.6}$$

and regarding b, the optimal bias filed, $\hat{b}$, that minimised the energy E can be updated as follows:

$$\hat{b} = \frac{\Big(V\Big(c_{inside} H_\varepsilon(\phi) + c_{outside}\big(1 - H_\varepsilon(\phi)\big)\Big)\Big) * K_\sigma}{\Big({c_{inside}}^2 H_\varepsilon(\phi) + {c_{outside}}^2\big(1 - H_\varepsilon(\phi)\big)\Big) * K_\sigma} \tag{2.7}$$

Similar to the previous level set-based method, such as [11], Heaviside function H and Dirac function δ used in above equation are as follows:

$$\begin{cases} H_\varepsilon(x) = \dfrac{1}{2}\left[1 + \dfrac{2}{\pi} arctan\left(\dfrac{x}{\varepsilon}\right)\right] \\ \delta_\varepsilon(x) = \dfrac{1}{\pi} \cdot \dfrac{\varepsilon}{\varepsilon^2 + x^2} \end{cases}, x \in R \tag{2.8}$$

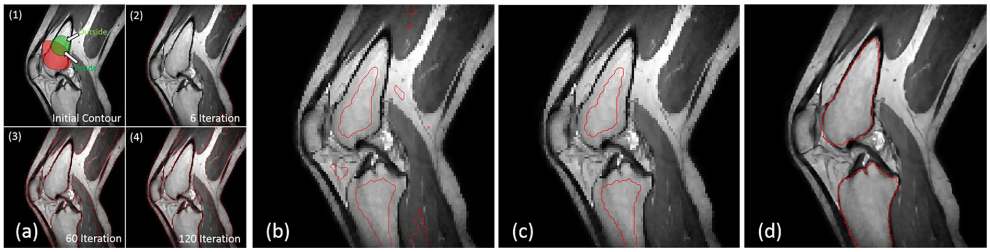


Figure 2.3: (a) Evolution of 3DLICLS: (1): red object is the initial contour example, green circle is used to calculate the inside energy (e_1) and outside energy (e_2); (2), (3), and (4): red contours are the boundary results in iterations 5, 14 and 60. (b) Red contour represents the boundary after erosion. (c) Red contour represents a selected area for femur and tibia. (d) Red contour represents the dilated rough trabecular mask for femur and tibia

The selected iterations from the evolution are shown in Figure 2.3a. Nevertheless, the output result of 3DLICLS includes not only the trabecular bone but also may include the infrapatellar fat, as Figure 2.3a4 shows. In order to obtain rough segmentation of the femur and tibia, a 3D spherical-shaped erosion kernel with a radius of 5 mm was applied (Figure

2.3b. After erosion, the femur and tibia bone area are separated from infrapatellar fat using a connectivity search (Figure 2.3c). Then, an image dilation operation with the same kernel size of erosion finalises the result, a rough segmentation of the trabecular bone of femur and tibia (see Figure 2.3d). The basic theory of 3DLICLS also supports multi-phase detection[17]. In this study, we aimed to use it to position the trabecular bones roughly, and thus, the two-phase model was selected.

2.2.3.2 Inhomogeneity correction

The 3DLICLS process results in a rough segmentation of the trabecular bone of femur and tibia and a bias template of the complete FOV, as shown in Figure 2.4b. This bias template is used to remove inhomogeneity, and the bias-corrected volume is computed as:

$$V_{corrected} = V \,/\, b \tag{2.9}$$

where $V_{corrected}$ is the corrected volume, V is the original volume and b is the bias template from 3DLICLS. The corrected image slice is shown in Figure 2.4c, and the comparison between before and after correction is indicated by the red box in Figure 2.4a and Figure 2.4c.

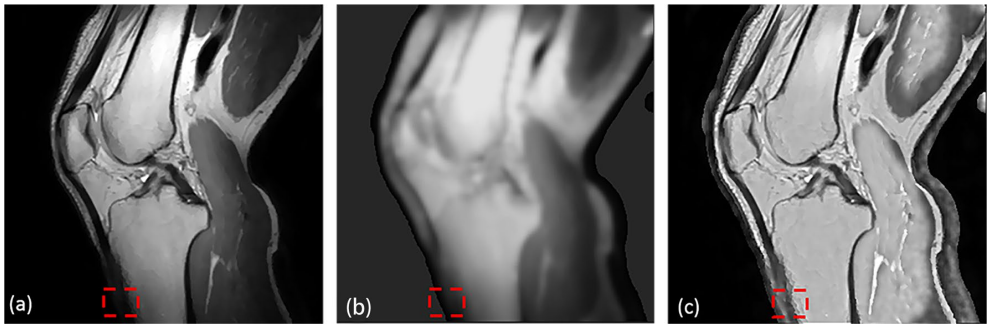


Figure 2.4: Inhomogeneity correction. (a) Original data. (b) Inhomogeneity template. (c) Inhomogeneity corrected data. (Red box represents the region can be optimized to segment).

2.2.4 Generating normal vectors on the rough trabecular surface

Previous steps provide only the rough shape of the trabecular boundary. To obtain the precise trabecular bone boundary, an intensity line is generated along the normal vector of the trabecular surface (slice by slice), as Figure 2.5a shows. To determine the normal vector of each point in the contour, such as green point A in Figure 2.5b, we can apply singular value decomposition (SVD) among its neighbour points (yellow points) and itself (green point). In the case of point A, the coordinates of the points form the matrix M_A,

$$M_A = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix} \tag{2.10}$$

To obtain optimal solution in least squares sense, the first and second rows of M_A are corrected by their respective average x_{mean} and y_{mean}, i.e. $x_n'=x_n - x_{mean}$, and obtain:

$$M_A' = \begin{bmatrix} x_1' & x_2' & \dots & x_n' \\ y_1' & y_2' & \dots & y_n' \end{bmatrix} \tag{2.11}$$

Using SVD, M_A' is then decomposed into three parts, U, Σ and V,

$$M_A' = U\Sigma V^T \tag{2.12}$$

Where U provides the orthonormal vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$. $\mathbf{u}_1$ is the tangent unit vector of point A, while $\mathbf{u}_2$ is the normal unit vector (the vector we use in this study).We refer to [21] for further explanation on U, Σ and V.

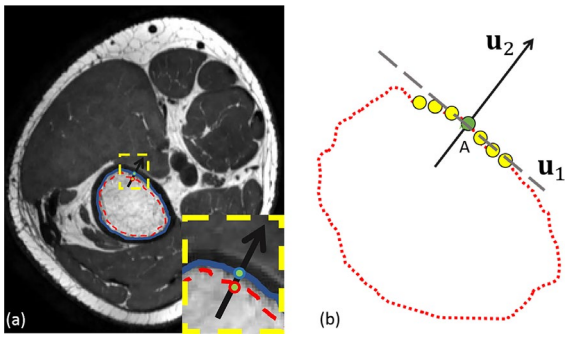


Figure 2.5: Illustration of normal vector calculation. (a) Red dash contour is the result of 3DLICLS, blue solid contour is target boundary and yellow dash rectangle is the enlarged area. (b) black arrow is the normal vector, grey dash line is the tangent line of the contour and yellow circles are the neighbour points of target A. u_1 and $u_2$represent the tangent and normal vectors.

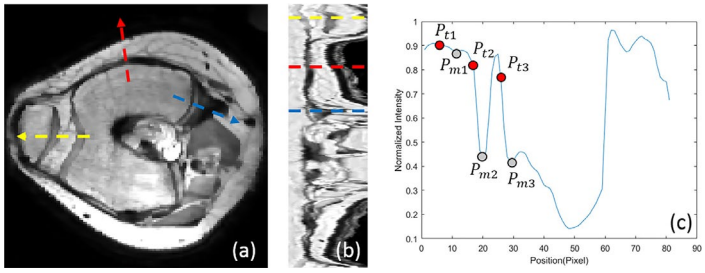


Figure 2.6: (a) Transverse view of femur. The arrow with different colour corresponds to the different coloured dash line in (b). (b) 2D intensity line image with different coloured dash line. (c) Intensity line of blue-dashed line on (b) (grey point represents the local minimum; red point represents trabecular candidate).

In a pilot study, a length of 45 mm for the intensity line along the normal vector (15 mm inward and 30 mm outward) was found to be adequate for robust inclusion of the precise trabecular boundary. Combining all intensity lines around the trabecular bone, an intensity line-based 2D image (IL2DI) is constructed, as Figure 2.6b shows.

2.2.5 Determination of the precise trabecular boundary

From the resulting 2D intensity lines, we now determine the precise trabecular surface slice by slice. Figure 2.6 shows a transverse slice of the femur (a), the complete set of IL2DI (b) and a typical intensity line (c). For each intensity line, trabecular candidate points P_t are defined as the point of maximum decline before a local minimum P_m. A maximum of five candidates P_t are identified per intensity line.

To calculate the precise position of the trabecular boundary, many subsets of boundary candidates are constructed from a set of neighbouring intensity lines (M= 7 in this study). In order to determine the suitable edge point (A or B) for the example of row 23, six permutations are obtained as shown in Figure 2.7.

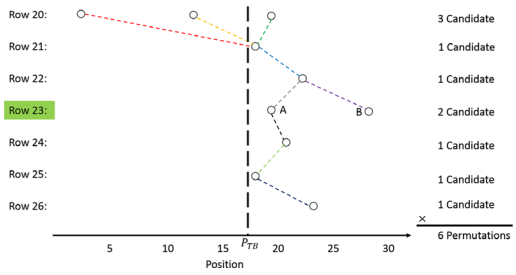


Figure 2.7: Illustration of trabecular boundary determination of row 23 (line connection between candidates (different colour means different line connection), and black-dashed line represents the position of P_{TB}).

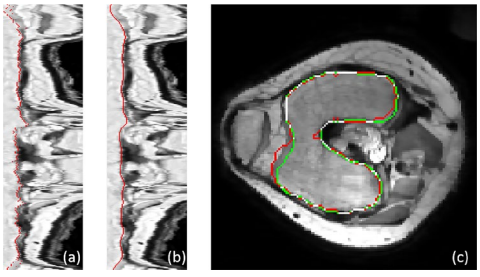


Figure 2.8: (a) Selection of trabecular boundary in IL2DI (red points). (b) Smooth version of (a) (red points). (c) Trabecular boundary in transverse view (red, rough trabecular boundary; green, optimized trabecular boundary; white, overlap of green and red).

The trabecular bone boundary is now determined as the candidates with minimal variance and closest to the rough trabecular boundary as the minimum of the cost function:

$$\min_{n=1:N}\left\{f_{STD}\frac{STD_n}{\max\limits_{n=1:N} STD_n}+f_{DD}\frac{\sum_{m=1}^{M}\mathrm{abs}(P_n^m-P_{TB})}{\max\limits_{n=1:N}\sum_{m=1}^{M}\mathrm{abs}(P_n^m-P_{TB})}\right. \tag{2.13}$$

where f_{STD} and f_{DD} are the weight for standard deviation and distance deviation from the initial trabecular boundary, respectively, and defined as $f_{STD} = f_{DD} = 1$. STD_n is the standard deviation of the given permutation of trabecular candidates. P_n^m represents m row of n permutation and P_{TB} means the position of the rough trabecular boundary. The first term minimizes the distance between candidates among rows, while the second term minimizes the distance between boundary end result and initial 3DLICLS result. The boundary selection of example in Figure 2.7 is A, and the example of selected candidates in IL2DL is shown with red points in Figure 2.8a.

Then, as Figure 2.8b shows, the found trabecular contour is smoothened in the IL2DI view with a Gaussian filter of kernel size 3. The found contours are then mapped back from IL2DI to transverse view and smoothened in the slice direction to ensure a smooth continuous trabecular mask, as Figure 2.8c shows.

2.2.6 Cortical bone boundary detection

The main obstacle to extract the robust cortical edge exists in the weak contrast between cortical bone and ligament tissue. Before solving the obstacle, we make two assumptions:

1.The thickness of cortical bone on femur decreases in inferior direction, while the one on tibia decreases in superior direction [22].

2. From perspective of manual segmentation, the weak boundary is identified based on surrounding tissue among adjacent slices (the assumption is based on discussion with two experts who have segmented over 50 data sets at the orthopaedic lab for the purpose of generating FE models).

According to the assumptions, we propose two steps to solve the challenge of cortical bone determination, especially in the region with a weak edge.

Step 1 Construction of initial cortical boundary

An initial cortical boundary is obtained by searching for the point of maximum incline after the first minimum P_m, starting from the trabecular bone boundary P_t in the IDL2L (see Figure 2.9).

This procedure provides a first guess of the cortical bone. However, there can still be outliers in the area with noise and weak contrast, especially near the ligaments where a ligament boundary can be mistakenly selected for cortical bone (yellow box in Figure 2.10a). For that matter, the actual cortical boundary is iteratively detected based on assumption 2.

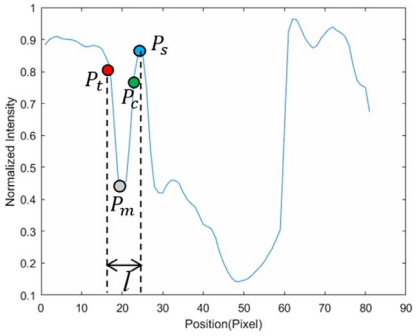


Figure 2.9: Intensity line of blue-dashed line in Figure 2.7b (P_t represents the trabecular boundary, P_m represents the point with local minimum intensity, P_c represents the cortical boundary, P_s represents the maximum search range and l represents the distance between P_t and P_s).

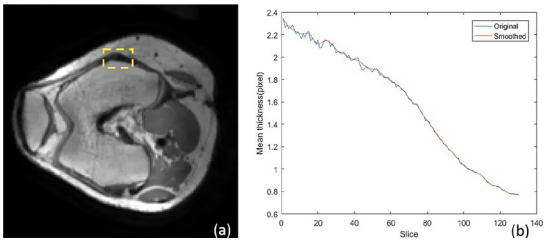


Figure 2.10: (a) Transverse view of femur (yellow box represents the connected area between cortical bone and ligament). (b) Mean cortical thickness variation along inferior direction of femur (blue, original version; red, smoothed version).

Firstly, the average thickness R_{MeanC} along the inferior direction of femur and superior one of tibia in each slice can be calculated by using the cortical area divided by the mean perimeter, C_{TC} (average perimeter of cortical and trabecular boundaries):

$$R_{MeanC} = A_{Cortical}/C_{TC} \tag{2.14}$$

$A_{Cortical}$ means the area of cortical bone, which calculated by the area of total bone (cortical and trabecular) and trabecular bone ($A_{Cortical} = A_{Total} - A_{Trabecular}$). Figure 2.10b shows the variation of the mean cortical thickness in each slice from inferior direction for the femur and the smoothened version.

Similar to trabecular optimization, the cortical boundary is determined as the minimum of the cost function, which consists of the candidates with minimal variance and closest to position of $P_{Cmean} = P_{TB} + R_{MeanC}$:

$$\min_{n=1:N}\left\{f_{STD}\frac{STD_n}{\max\limits_{n=1:N} STD_n} + f_{DD}\frac{\sum_{m=1}^{M}\mathrm{abs}(P_n^m - P_{Cmean})}{\max\limits_{n=1:N}\sum_{m=1}^{M}\mathrm{abs}(P_n^m - P_{Cmean})}\right. \tag{2.15}$$

where f_{STD} and f_{DD} are the weight for standard deviation and distance deviation from P_{Cmean}, respectively, and defined as $f_{STD} = f_{DD} = 1$. STD_n is the standard deviation of the given permutation of cortical candidates. P_n^m represents m row of n permutation. Normally, the candidates of cortical boundary are at most three maximum incline after the trabecular boundary. Nevertheless, if the first incline is larger than the mean thickness of correspondent points of its last three layers, the position P_{Cmean} is added to the candidate set of cortical boundary. To be more exact, this step simulates the assumption 2 and provides an extra option, position of P_{Cmean}, in the area may exist the weak edge. The optimization of cortical boundary is an iterative procedure, which will be updated until convergence of the change of mean cortical thickness. The criterion for convergence in this study was defined as a change in mean cortical thickness between two iterations to be less than 1 pixel.

At last, same as trabecular optimization, the found cortical contour is smoothened and mapped back to transverse view.

2.2.7 Evaluation

All datasets were analysed in MATLAB 2015b. Data analysis was carried out on a conventional laptop with CPU Intel Core I7-4700MQ (2.40 GHz) and 16 GB RAM. The manual segmentation was defined as ground truth for scoring of the automatic segmentation. The manual segmentation was performed in the Mimics software environment. The outcomes were quantified with the Dice sensitivity coefficient (DSC) [13, 14, 23, 24]and the average surface distance (ASD) [12, 13].

2.3 Results

2.3.1 Comparison between 3DLICLS and 2DLICLS

Figure 2.11 shows the initial trabecular result before the erosion operation between the 3DLICLS and 2DLICLS (performed the algorithm slice by slice). The result from 2DLICLS shows more leakage areas than the 3DLICLS.

2.3.2 Segmentation results for trabecular bone

Figure 2.12 shows the final trabecular result for the first data sets in sagittal view at mid-slice position and in transversal view at several key positions. The red contour represents rough trabecular result after 3DLICLS and image morphological operation and the green contour is the trabecular result after optimization in 2.2.5.

Generally, femur and tibia are isolated with 3DLICLS, as Figure 2.12b–h shows. The inhomogeneity problem increases near the boundary of the FOV as can be seen in Figure 2.12h. 3DLICLS however determines the bias field and is still able to segment the trabecular bone robustly in this area.

2.3.3 Segmentation results for cortical bone

Figure 2.13 shows the cortical segmentation results including cortical bone guess (blue) in first maximum incline (2.6 step (1)), final cortical mask within proposed method (yellow) and manual segmented mask (green). Plus, white point represents the overlap between proposed method and manual segmentation.

From (b) to (e) and (h) to (f), the phenomenon of cortical bone thinning towards the femur condyles and top of tibia can be noticed respectively, where the red contour (trabecular boundary) moves more and more towards the yellow and green contours (automatic and manual segmentation of the cortical boundary).

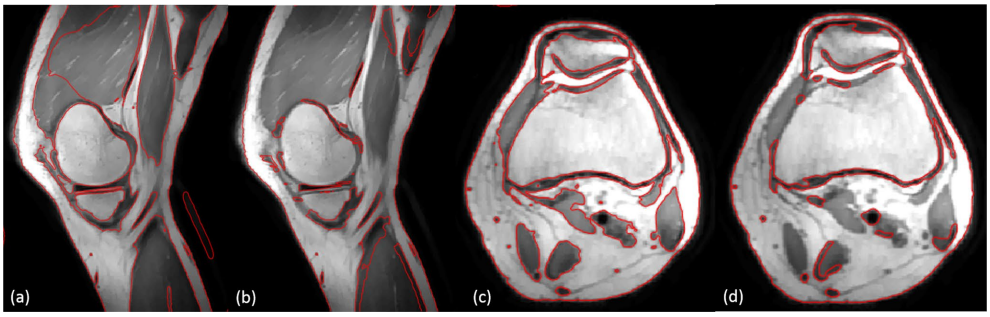


Figure 2.11: Comparison between 3DLICLS and 2DLICLS. Zero level set indicated by red contour. (a), (c) Sagittal and axial result of 3DLICLS. (b), (d) The sagittal and axial result of 2DLICLS.

In the shaft area without inhomogeneity (regions (b) and (c)), there is virtually full agreement between manual segmentation and our method. Difficulties arise in the areas containing transition from cortical bone to cartilage and/or ligament, depicted in (d)–(g). Despite the weak edges between ligament and cortical bone, the automatic segmentation still

displays minimal disagreement with the manual one. Furthermore, the performance in the region with inhomogeneity (h) also displays convincing result.

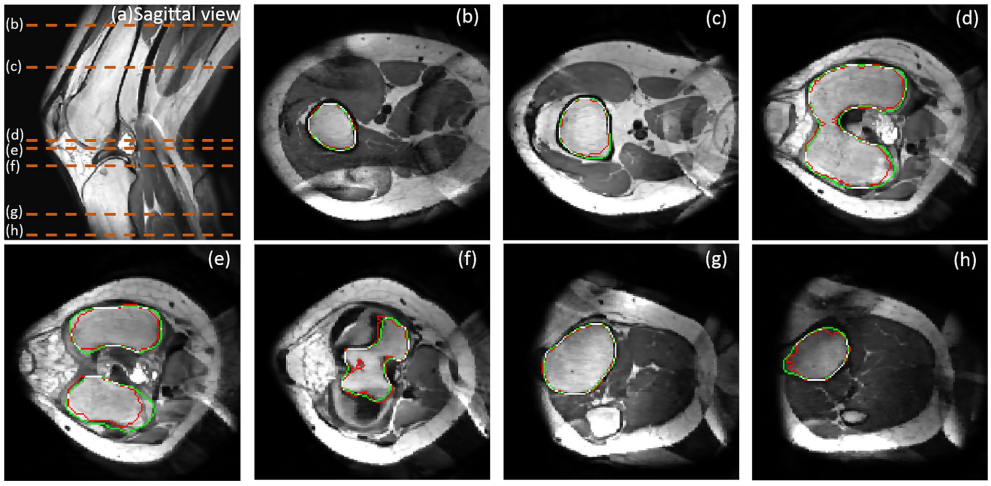


Figure 2.12: (a) Sagittal view from the result of 3DLICLS. (b–h) Transverse view of rough trabecular masks (red contour) and the optimized trabecular masks (green contour). White point means the overlap between rough trabecular mask and optimized mask.

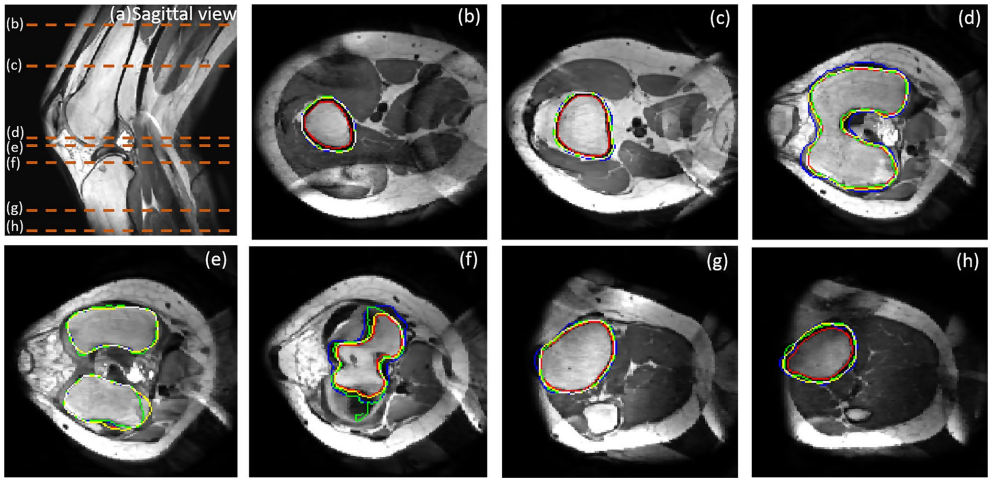


Figure 2.13: (a) Sagittal view to show the position of (b-h). (b–h) Comparison results of cortical boundary from transverse view. Green contour is manual segmentation, yellow contour is cortical boundary of proposed method, blue contour is the initial cortical mask (2.6 step (1)), red contour is the optimized trabecular mask and white point means the overlap between proposed method and manual segmentation.

2.3.4 Convergence of cortical boundary detection

Figure 2.14a and Figure 2.14b show the difference of sum of average cortical thickness along the iteration for femur and tibia among the 18 data, respectively. Generally, the change becomes converged after eight iterations, but the result normally changes little after four iterations.

2.3.5 Segmentation accuracy

As the boxplots in Figure 2.15 show, the average DSC are 0.96 ± 0.01 for the femur and 0.96 ± 0.02 for the tibia. Two typical situations with low DSC score are also shown in Figure 2.15. The average distances to surface between the automatically and manually segmented bones, 0.46 ± 0.14 mm for the femur and 0.47 ± 0.21 mm for tibia, are shown in Figure 2.16a, and a 3D difference for femur and tibia is shown schematically in Figure 2.16b and Figure 2.16c.

2.3.6 Segmentation time

The average time needed to segment one dataset (femur and tibia) with a matrix of $336 \times 336 \times 222$ voxels was around 250 s and 2.5 h for automatic and manual segmentation, respectively. Hence, the prosed method is efficient and promising for assisting segmentation research.

2.4 Discussion

In this study, we proposed an automatic workflow to segment the cortical and trabecular bone of femur and tibia in proton density weighted MRI contrast. A 3D level set-based algorithm was used to segment the rough trabecular boundary and remove any slow varying inhomogeneity. Trabecular and cortical bone boundaries were detected from the intensity profiles along normal vectors generated from the trabecular surface. Upon testing of the method in 18 datasets, the algorithm demonstrated its capability to handle field inhomogeneity and correctly detect trabecular and cortical bone over the full field of view including weak edges near ligament and thinning cortical bone. Scoring of the proposed segmentation method using manual segmentations as a reference yielded DSC over 0.95 and ASD errors less than 0.5 mm for both femur and tibia.

Table 2.1 shows a comparison of evaluation results from similar studies in recent literatures. Although our method's results are well within the range of success rates as reported from literatures, we must stress the difficulty in direct comparison between methods because of the differences in workflow. Shan et al. [14] and Fripp et al. [24] for instance used prior data, whereas we did not. Guo et al. [23] reported scores on trabecular bone

segmentation only, and Pang et al. [12] reported average surface distances for specific slice locations versus over the whole bone surface.

There are several limitations in this study. Firstly, manual segmentations from a trained expert were used as a ground truth for scoring the automatic method. As the exact boundary between cortical bone and ligament is often not completely clear even for orthopaedic surgeons, this ground truth is subject to debate. Hence, the results presented in this study only show the method's capability to simulate the manual evaluation of cortical bone. Secondly, the patients whose data was used in this study were all in relatively good health as the knee is concerned. Patients with pathologies that affect the bone and cartilage (e.g. osteoporosis, osteoarthritis, bone marrow lesions) may require re-tuning of the parameters of the automatic segmentation algorithm. This requires bigger datasets and clinical applications.

This fast and robust segmentation of trabecular and cortical bone boundary of the femur and tibia has the potential of providing a basis for surgical planning and more accurate finite element models of the knee joint. By removing the large workload that is involved in manual segmentation of MRI images, these methods can potentially be introduced in the clinic and in large-scale research projects. In this study, proton density weighted contrast was used, because of its wide availability, short scan time and orthopaedic relevance. The method, however, could also be adjusted to extract bone from other types of contrast, provided there is an overall difference in contrast between trabecular bone, cortical bone and adjacent tissues, and there is enough consistency in the trabecular to cortical bone boundary to correct any weak edges using its surroundings.

Furthermore, as the method contains no substantial assumptions, constraints or premises with respect to the shape of the bone but rather to the contrast, it is feasible to extend this method to the shoulder and elbow joint.

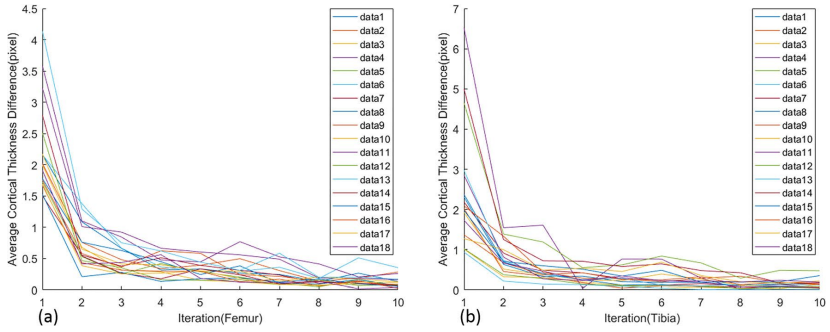


Figure 2.14: Average cortical thickness difference between neighbour iterations. (a) Femur; (b) Tibia.

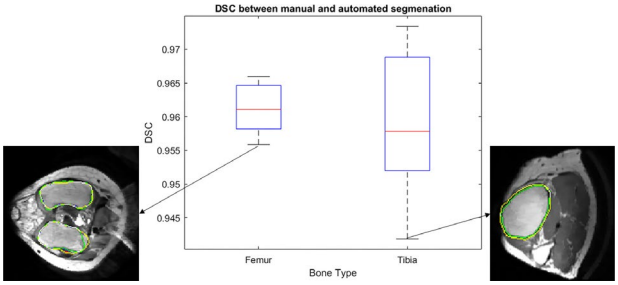


Figure 2.15: Box plot of comparison between automatic and manual segmentation calculated in DSC scores (green contour is cortical boundary of manual segmentation, yellow contour is cortical boundary of proposed method, white point represents the overlap between green and yellow point).

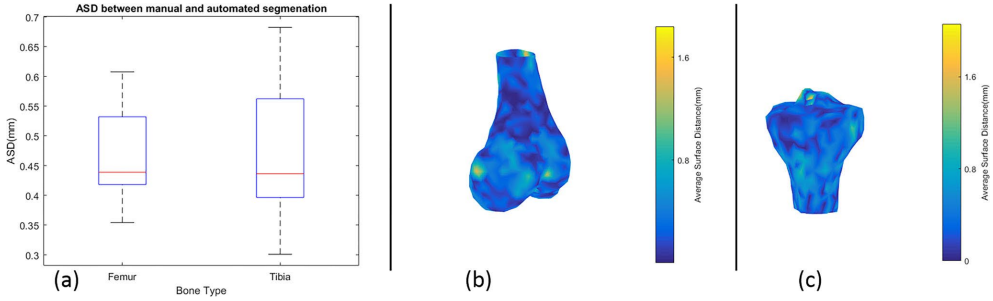


Figure 2.16: (a) Box plot of average surface distance difference between automatic and manual segmentation for femur and tibia. (b), (c) Distance difference in 3D view for femur and tibia, respectively.

Table 2.1: Result comparison between proposed method and previous studies

	DSC		**ASD (mm)**	
	Femur	Tibia	Femur	Tibia
Dodin et al [13]	0.94 ± 0.05	0.92 ± 0.07	0.50 ± 0.12	0.37 ± 0.09
Pang et al. [12]	\	\	0.46 ± 0.18	0.85 ± 0.40
Shan et al. [14]	0.97 ± 0.01	0.97 ± 0.01	\	\
Guo et al. [23]	0.94	0.94	\	\
Fripp et al. [24]	0.96	0.96	\	\
This study	0.96 ± 0.01	0.96 ± 0.02	0.46 ± 0.14	0.47 ± 0.21

2.5 Conclusion

In this paper, we presented and evaluated an automatic workflow to segment the trabecular and cortical bone of femur and tibia with PDW sequence in MRI. Initial results compared with manual segmentation indicate the possibility to provide an automatic segmentation to researchers and clinical doctors to perform further analysis rather than the time-consuming manual segmentation.

Future studies will include the following: an increase in the number of patients of the test group; an extension of the method to determine the knee bone and cartilage; and an automated workflow to provide clinically relevant parameters, such as tibia tubercle-trochlear groove distance (TT-TG) and patella tilt.

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW).

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

[1] P. Teller, H. König, U. Weber, and P. Hertel, "MRI Atlas of Orthopedics and Traumatology of the Knee," 2003, doi: 10.1007/978-3-642-55620-3.

[2] C. Rao, C. K. Fitzpatrick, P. J. Rullkoetter, L. P. Maletsky, R. H. Kim, and P. J. Laz, "A statistical finite element model of the knee accounting for shape and alignment variability," *Medical Engineering and Physics,* vol. 35, pp. 1450-1456, 2013, doi: 10.1016/j.medengphy.2013.03.021.

[3] D. White, K. L. Chelule, and B. B. Seedhom, "Accuracy of MRI vs CT imaging with particular reference to patient specific templates for total knee replacement surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery,* vol. 4, pp. 224-231, 2008.

[4] C. Scholes, V. Sahni, S. Lustig, D. A. Parker, and M. R. J. Coolican, "Patient-specific instrumentation for total knee arthroplasty does not match the pre-operative plan as assessed by intra-operative computer-assisted navigation," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 660-665, 2014.

[5] A. Aprovitola and L. Gallo, "Knee bone segmentation from MRI: A classification and literature review," *Biocybernetics and Biomedical Engineering,* vol. 36, no. 2, pp. 437-449, 2016, doi: 10.1016/j.bbe.2015.12.007.

[6] T. Heimann and B. Morrison, "Segmentation of knee images: A grand challenge," 2010.

[7] T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis,* vol. 13, pp. 543-563, 2009, doi: 10.1016/j.media.2009.05.004.

[8] L. M. Lorigo, O. Faugeras, W. E. L. Grimson, R. Keriven, and R. Kikinis. Segmentation of Bone in Clinical Knee MRI Using Texture-Based Geodesic Active Contours.

[9] H. Yousefi, M. Fatehi, M. Amian, and R. a. Zoroofi, "A fully automated segmentation of radius bone based on active contour in wrist MRI data set," *2013 20th Iranian Conference on Biomedical Engineering (ICBME),* pp. 42-47, 2013, doi: 10.1109/ICBME.2013.6782190.

[10] T. Liu, H. Xu, W. Jin, Z. Liu, Y. Zhao, and W. Tian, "Medical Image Segmentation Based on a Hybrid Region-Based Active Contour Model," *Computational and Mathematical Methods in Medicine,* vol. 2014, pp. 1-10, 2014, doi: 10.1155/2014/890725.

[11] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition,* vol. 43, pp. 1199-1206, 2010, doi: 10.1016/j.patcog.2009.10.010.

[12] J. Pang, J. B. Driban, T. E. McAlindon, J. G. Tamez-Peña, J. Fripp, and E. L. Miller, "On the Use of Coupled Shape Priors for Segmentation of Magnetic Resonance Images of the Knee," *IEEE Journal of Biomedical and Health Informatics,* vol. 19, pp. 1153-1167, 2015, doi: 10.1109/JBHI.2014.2329493.

[13] 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, pp. 1413-1424, 2011, doi: 10.1007/s11517-011-0838-8.

[14] L. Shan, C. Zach, C. Charles, and M. Niethammer, "Automatic atlas-based three-label cartilage segmentation from MR knee images," *Medical Image Analysis,* vol. 18, pp. 1233-1246, 2014, doi: 10.1016/j.media.2014.05.008.

[15] O. Tokuda *et al.*, "MRI of the anatomical structures of the knee: The proton density-weighted fast spin-echo sequence vs the proton density-weighted fast-recovery fast spin-echo sequence," *British Journal of Radiology,* vol. 85, pp. 686-693, 2012, doi: 10.1259/bjr/99570113.

[16] S.-Y. Lee, W.-H. Jee, S. K. Kim, and J.-M. Kim, "Proton density-weighted MR imaging of the knee: fat suppression versus without fat suppression," *Skeletal Radiology,* vol. 40, pp. 189-195, 2011.

[17] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *IEEE Transactions on Image Processing,* vol. 20, pp. 2007-2016, 2011, doi: 10.1109/TIP.2011.2146190.

[18] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 603-619, 2002.

[19] T. F. Chan and L. a. Vese, "Active contours without edges," *IEEE Transactions on Image Processing,* vol. 10, pp. 266-277, 2001, doi: 10.1109/83.902291.

[20] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Transactions on Image Processing,* vol. 19, pp. 3243-3254, 2010, doi: 10.1109/TIP.2010.2069690.

[21] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-9, pp. 698-700, 1987, doi: 10.1109/TPAMI.1987.4767965.

[22] J. H. Gosman, Z. R. Hubbell, C. N. Shaw, and T. M. Ryan, "Development of Cortical Bone Geometry in the Human Femoral and Tibial Diaphysis," *Anatomical Record,* vol. 296, pp. 774-787, 2013, doi: 10.1002/ar.22688.

[23] Y. Guo, J. Jiang, S. Hao, and S. Zhan, "Distribution-Based Active Contour Model for Medical Image Segmentation," *2011 Sixth International Conference on Image and Graphics,* pp. 61-65, 2011, doi: 10.1109/ICIG.2011.11.

[24] J. Fripp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation of the bone and extraction of the bone–cartilage interface from magnetic resonance images of the knee," *Physics in Medicine and Biology,* vol. 52, pp. 1617-1631, 2007, doi: 10.1088/0031-9155/52/6/005.

Chapter 3 Knee bone and cartilage segmentation based on a 3D deep neural network using adversarial loss for prior shape constraint

Abstract:

Introduction: Fast and accurate segmentation of knee bone and cartilage on MRI images is becoming increasingly important in the orthopaedic area, as the segmentation is an essential prerequisite step to a patient-specific diagnosis, optimising implant design, and pre- and intra-operative planning. However, manual segmentation is time-intensive and subject to inter- and intra- observer variation. Hence, in this study, we proposed a 3D deep neural network based method to fully automate the segmentation of knee bone and cartilage.

Methods: A 3D deep neural network using adversarial loss was proposed to segment the knee bone in a resampled image volume in order to enlarging the contextual information and incorporating prior shape constraint. A restoration network was proposed to further improve the bone segmentation accuracy by restoring the bone segmentation back to the original resolution. A conventional U-Net like network was used to segment the cartilage. The ultimate results were the combination of the bone and cartilage outcomes through post-processing. The quality of the proposed method was thoroughly assessed using various measures for the dataset from Grand Challenge SKI10, together with a comparison with a baseline network U-Net.

Results: A fined tuned U-Net like network can achieve state-of-the-art results without any post-processing operations. Our method achieved a total score higher than 76 in terms of SKI10 validation dataset.

Conclusions: Our method showed to be robust to extract bone and cartilage mask from MRI dataset, even for the pathological case.

3.1 Introduction

Quantitative analysis of knee joint structure is a topic of increasing interest as its applications continue to broaden from direct diagnostic purposes to the implant design, and pre- and intra- operative planning. Due to the non-invasive nature and capability to discriminate cartilage from adjacent tissues, magnetic resonance imaging (MRI) is the most effective imaging device to perform knee joint analysis. However, due to the low contrast among different tissues (similar longitudinal and transverse relaxation time), image artefacts and intensity inhomogeneity problems in MRI [1], the accurate segmentation of knee joint is still an open problem, especially in the knee with degenerative disease [2].

To obtain an accurate mask for knee bone and cartilage, fully manual and semi-automatic segmentation approaches were often applied to clinical studies [3-5]. Nonetheless, they were time-consuming, and the reproducibility highly depends on the knowledge of experts. Hence, an automated method to segment the knee joint structure was of great interest in the past decade [6, 7]. The popular methods for this aim can be divided into model-based [8-10], atlas-based [11, 12] and classification-based [1, 2, 13] methods. Although these three types of methods showed promising results to automate the knee structure segmentation, they might perform poorly in the case within high subject variability [2].

Recently, deep convolutional neural network (CNN) based methods have achieved enormous success in biomedical imaging problems, such as classification [14], segmentation [15-18]. Regarding knee joint structure segmentation, Prasoon et al. firstly applied the 2D tri-planar CNNs (axial, coronal and sagittal plane) to classify a pixel label (background or tibial cartilage) by providing local image patches around that pixel [19]. Nevertheless, Ronneberger et al. pointed out there were two drawbacks on above architecture: large redundancy and a trade-off between localisation accuracy and the use of context, and proposed a dense prediction network with skip connection, U-Net [18]. This kind of architecture considered both low-level and high-level features for voxel classification and was applied to the knee joint segmentation by Liu et al. [2], Zhao et al. [20] and Ambellan [21]. In general, the pixel-wise or voxel-wise loss, e.g. cross-entropy loss and dice loss, was utilised as the loss function for U-Net. However, there was no guarantee the spatial consistency of the final output [22], thereafter a further optimisation step was always required to refine the segmentation result such as deformable model [2], conditional random field (CRF) [20], statistical shape model (SSM) [21]. Although the deformable model and CRF considered the relevant spatial information to refine the segmentation, it might cause serious boundary leakage in the low contrast regions [22]. Ambellan et al. proposed to utilise SSM to refine segmentation using the anatomical prior knowledge and achieved the state-of-the-

art result. Nevertheless, the introduction of SSM resulted in a lot of extra calculations and the regulation was limited to the variability of the training dataset. Overall, although deep learning based method have been demonstrated as state-of the-art method in knee joint segmentation, there is still room to improvement.

In this study, we aim to further research a 3D CNN based method to perform knee bone and cartilage segmentation. The contributions in this paper are: (1) introduce the adversarial loss as providing the spatial consistency for the knee bone segmentation; (2) adopt different network and different loss function for bone and cartilage segmentation based on their property; (3) demonstrate the validity of proposed method on a public dataset from MICCAI SKI10 grand challenge and fully compare the performance with the various CNN models.

3.2 Materials and Methods

3.2.1 Data description

The data used in this study was from the Segmentation of Knee Images 2010 (SKI10) competition which was focused on the knee bone and cartilage segmentation [6]. The image data sets were acquired in a sagittal manner with a pixel spacing of 0.4×0.4 mm and a slice thickness of 1 mm. Total number of the knee images used in this study were 100 (60 for training and 40 for testing), and the cases of left and right knees were distributed approximately equally. Among the scans, 90% of the data were acquired in 1.5T, and the rest were at 3T and 1T. The majority of data used T1-weighting, and the rest of them were acquired with T2-weighting. All images were acquired for surgery planning of partial or complete knee replacement, and therefore a high degree of pathological deformations of the knee was included in the dataset.

3.2.2 Automatic workflow for knee bone and cartilage segmentation

In this work, we aim to establish a fully automatic workflow to extract knee joint structure (bone and cartilage) with highly accurate and robust segmentation, including the pathological data. Figure 3.1 depicts the steps of the proposed workflow. Firstly, MRI images were resampled to enlarge the field of view by the networks; Secondly, an image normalization method standardizes the image to a similar intensity range; Thirdly, the bone and cartilage were segmented by the bone network in a resampled resolution; Fourthly, the segmented bone and cartilage masks from the bone network were restored back to the original resolution through a restoration network; Fifthly, the cartilage was segmented through a cartilage network in original resolution; Lastly, the outputs of the cartilage network and the restoration network were post-processed for the final results.

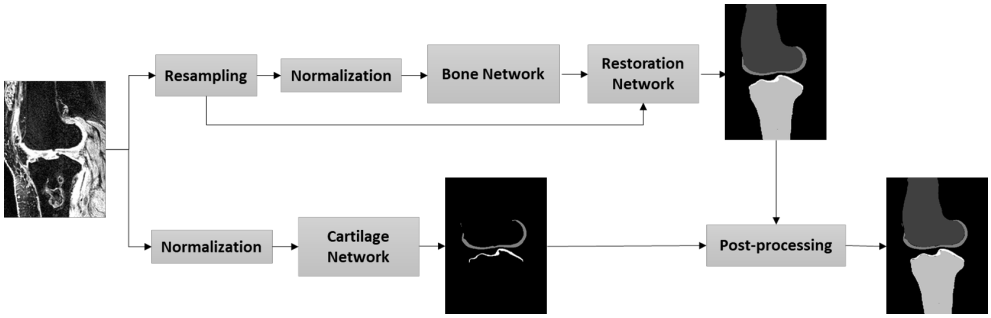


Figure 3.1: Workflow for the proposed workflow of knee bone and cartilage segmentation.

3.2.2.1 Pre-processing

Our pre-processing included pixel size normalization and intensity normalization. The first pre-processing step in this work was volume resampling. One of the main challenges in medical image segmentation using deep learning is the volume size, as it is too large to feed into the networks due to the lack of GPU memory. A patch-wise strategy was an option to solve this issue by breaking down the volume into multiple patches (overlapping or random patches) to fit the GPU memory requirement [23]. Yet, this strategy may result in a higher variance among the patches and lose the contextual information [24], especially for the large target. For the bone segmentation, we downsampled the image volume by a factor of 2 resulting in a new spacing $0.8 \times 0.8 \times 2$. With the resampling step, the input patch can cover more contextual information for the bone segmentation. In contrast, the cartilage segmentation based on CNN is relatively sensitive to the resampling due to its small volume size. Hence, for the cartilage segmentation, we input neural network the image with the original size.

The second step of pre-processing was the intensity normalization. The imaging noise from reconstruction of MRI volume, such as DC spike, results in the extreme intensity of some voxels [25]. A robust intensity cut-off was selected to prevent the long intensity tails effect for both bone and cartilage segmentation [25]. In this study, the minimum and maximum cut-offs were selected as the threshold with first and last 2% cumulative intensity histogram. Then, a following z-score strategy was adopted to normalize the intensity by subtracting the mean and dividing by the standard deviation.

3.2.2.2 Deep neural network for bone and cartilage segmentation

a) The architecture of the networks

Since the advent of U-Net [18], many architecture modifications have been proposed to further improve the performance of the segmentation task. However, Isensee et al.

demonstrated that not all of them were effective and pointed out a typical U-Net architecture can achieve state-of-the-art results with a thorough design of adaptive pre-processing, training scheme and inference strategy [26]. In this study, we extended the idea of nnU-Net [26] by adding the adversarial loss to refine the segmentation and used nnU-Net as a baseline for the segmentation performance comparison. The architecture of the bone network was similar to pix2pix network [27] (Figure 3.2), which consisted a generator trained for mask prediction, and a discriminator trained for discriminate the produced masks ('fake') from ground truth labels ('real') (Figure 3.2). The framework of the generator in this study consisted of an encoding path to encode the valid features and a decoding path to perform a voxel-based classification. The encoding path contained the repeated layers of two convolutions, followed by an instance normalization, leaky rectified linear unit and a max-pooling operation with stride 2 for down-sampling. The up-sampling path also contained the repeated layers of convolution, but a skip connection was adopted by a concatenation of the correspondingly cropped feature from contraction path and the output of the up convolutions from the last layer. At the final layer, a final 1×1×1 convolution was used to map each component feature vectors to the desired number of classes, and a SoftMax calculation was followed at last to output a probability for each class. Both U-Net liked [17] and V-Net like [28] architectures were used for the generator in this study, which might result in some slight variations compared to the above description and the detail of all used networks in this study are summarised in the supplemental material.

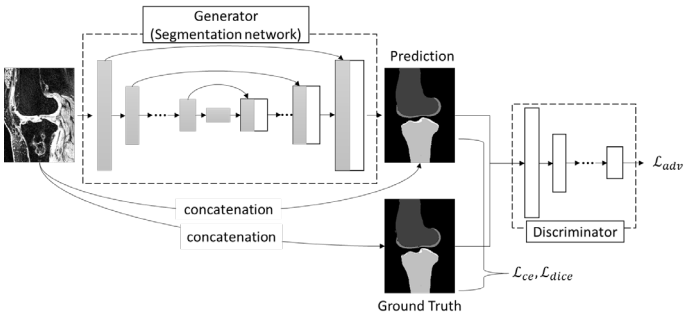


Figure 3.2: Architecture of the bone network.

The architecture of the discriminator of the bone network was a convolutional "PatchGAN" classifier use the module form of convolution-BatchNorm-Relu [27]. The input of the discriminator was the combination of the image patch and the corresponding segmentation patch. The detail of the architecture is provided in the supplemental material.

The input of the restoration network was the concatenation of the resampled image and the segmented mask from the bone network. The architecture of the restoration network

consisted of two convolutional layers, then followed by an upscaled deconvolutional layer and then finally another two convolutional layers to convert the feature maps into desired number of class.

The architecture of the cartilage network was basically nnU-Net 3D full resolution [26]. The input of the cartilage network was in original resolution, with patch size of $160 \times 192 \times 64$.

The details of both cartilage network and restoration network are described in the supplemental material.

b) Loss function

As Figure 3.2 and equation (3.1) illustrate, to test the optimal loss options for a robust knee bone segmentation, the loss function, $\mathcal{L}_{gen}$, used in the generator (bone network) consisted of three parts, category cross-entropy loss ($\mathcal{L}_{cce}$), dice loss ($\mathcal{L}_{dice}$) and adversarial loss ($\mathcal{L}_{adv}$). $\mathcal{L}_{cce}$ and $\mathcal{L}_{dice}$ concern the low-level pixel-wise prediction, while the $\mathcal{L}_{adv}$ preserves the higher-level consistency conditioned on the input.

$$\begin{aligned}\mathcal{L}_{gen}(x, y; \theta_{gen}, \theta_{disc}) \\ &= \lambda_{cce}\mathcal{L}_{cce}(G(x; \theta_{gen}), y) + \lambda_{dice}\mathcal{L}_{dice}(G(x; \theta_{gen}), y) \\ &+ \lambda_{adv}\mathcal{L}_{adv}(G(x; \theta_{gen}), x; \theta_{disc})\end{aligned} \quad (3.1)$$

Where x and y are the input image volume and the corresponding label. λ_{cce}, λ_{dice}, and λ_{adv} are the weights for the corresponding losses, and the loss is ignored if the corresponding weight sets to 0. θ_{gen} and θ_{disc} are the parameters of the networks of the generator and discriminator, respectively. The pixel-wise category cross-entropy loss is formulated as $\mathcal{L}_{cce}(\hat{y}, y) = \frac{1}{whd}\sum_i^{whd}\sum_j^c y_{i,j}\ln(\hat{y}_{i,j})$, where c represents the number of target classes and w, h and d indicates the width, height and depth of the volume patch. The pixel-wise dice loss is formulated as $\mathcal{L}_{dice}(\hat{y}, y) = -\sum_i^c \frac{2\sum_j^{whd} y_{i,j}\ln(\hat{y}_{i,j})}{\sum_j^{whd} y_{i,j}^2 + \sum_j^{whd}\ln(\hat{y}_{i,j})^2}$. For the adversarial loss, we chose the adversarial loss of LSGAN [29] in this study, and therefore, is formulated as

$$\mathcal{L}_{adv}(x; \theta_{gen}, \theta_{disc}) = \mathcal{L}_{MSE}(D(G(x; \theta_{gen}); \theta_{disc}), \mathbf{1}) \quad (3.2)$$

where $\mathcal{L}_{MSE}(\hat{z}, z) = (\hat{z} - z)^2$, x indicates the input patch. The discriminator attempts to learn the differences between the label and prediction distributions by minimising the loss function as,

$$\mathcal{L}_{disc}(G(x; \theta_{gen}), y) = \mathcal{L}_{MSE}(D(G(x; \theta_{gen}); \theta_{disc}), \mathbf{0}) + \mathcal{L}_{MSE}(y, \mathbf{1}) \quad (3.3)$$

Where x and y indicate the input patch and the corresponding annotation, respectively.

For the cartilage network, the loss function is formulated as

$$\mathcal{L}_{cart}(\hat{y}, y; \theta_{cart}) = \lambda_{cce}\mathcal{L}_{cce}(\hat{y}, y) + \lambda_{dice}\mathcal{L}_{dice}(\hat{y}, y) \quad (3.4)$$

Where y and $\hat{y}$ indicate ground truth, and the prediction result of the cartilage network respectively, and θ_{cart} indicates the parameters of cartilage network

For the restoration network, the loss function was as

$$\mathcal{L}_{restore}(\hat{y}, y; \theta_{restore}) = \mathcal{L}_{cce}(\hat{y}, y) \quad (3.5)$$

Where y and $\hat{y}$ indicate ground truth, and the prediction result of the restoration network respectively, and $\theta_{restore}$ indicates the parameters of cartilage network.

c) Training procedure

One common challenge in deep learning training is limited training data. The data augmentation is one of the options to be taken to prevent overfitting and have been generally accepted as an add-in in the deep learning method. The data augmentation adopted in this study was random scaling (0.85-1.15), random elastic deformations, gamma correction augmentation and random mirroring along the frontal axis (simulating the left or right knee joint).

In order to implement a fair comparison among the different architectures, the training strategy was adopted similar to [26]. There are 6000 training batches in an epoch. The Adam optimizer with an initial learning rate of 1×10^{-3} was utilised for both the generator and the discriminator in this study, and the learning rate was reduced by a factor of 5 if the loss was not improved in last 5 epochs and the training stopped if the loss was no improved in the last 20 epochs. The maximum epoch was limited to 500. The proposed deep CNNs were implemented in Python 3.7 using Pytorch with a 3.7 GHz Intel(R) i7 Xeon (R) E5-1620 V2 CPU and a GTX 1080 Ti graphics card with 11GB GPU memory.

d) Inference

In the inference phase, the new input image volume was split into many sub-volume patches and input to the networks. Then, the class of each voxel was determined by the largest probability of the output probability maps from the neural network. At last, we needed to combine all the sub-volume patches back to form a full volume.

3.2.2.3 Post-processing

The main purpose of the post-processing is to combinate the advantages of the bone network and the cartilage network in order to generate a final bone and cartilage masks. Compared to the cartilage mask from the cartilage network, the bone network could provide

less mis-segmented result due to the large contextual information, however, less accurate due to lower resolution. Therefore, the output of the cartilage mask from the bone network after the restoration network was dilated by a $7 \times 7 \times 7$ kernel, after which was used to filter the cartilage mask from cartilage network. Finally, the ultimate output of the proposed workflow was the combination of the bone mask from the restoration network and the filtered cartilage mask of the cartilage network.

3

3.3 Evaluation design

a) Methods designed by SKI10

The evaluation method for knee bone and cartilage were different. Regarding to bone segmentation, average surface distance (AvgD) and root-mean-square symmetric surface distance (RMSD) were proposed [6, 30].

$$\mathrm{AvgD} = \frac{1}{N_S + N_R}\left(\sum_{i=1}^{N_S} \min_{r \in \partial R} \|s_i - r\|_2 + \sum_{i=1}^{N_R} \min_{s \in \partial S} \left\|r_j - s\right\|_2\right) \tag{3.6}$$

$$\mathrm{RMSD} = \sqrt{\frac{1}{N_S + N_R}\left(\sum_{i=1}^{N_S} \min_{r \in \partial R} \|s_i - r\|_2 + \sum_{i=1}^{N_R} \min_{s \in \partial S} \left\|r_j - s\right\|_2\right)} \tag{3.7}$$

Where ∂R and ∂S are the boundary of the automatic segmentation and reference segmentation respectively, N_S and N_R are the number of boundary respectively.

For the cartilage segmentation, volume difference (VD) and volume overlap error (VOE) were proposed [6, 21].

$$\mathrm{VD} = 100 \cdot \frac{|S| - |R|}{R} \tag{3.8}$$

$$\mathrm{VOE} = 1 - \frac{|S \cap R|}{|S \cup R|} \tag{3.9}$$

Where S and R indicate the automatic segmentation and reference segmentation respectively. As indicated by Heimann et al. [6], the cartilage boundaries to the sides were not always accurate, regions of interest (ROIs) for cartilage mask comparison was used in the above calculation.

b) Dice Similarity Coefficient (DSC):

The Dice similarity coefficient (DSC) score is defined as,

$$\text{DSC} = \frac{2T_P}{2T_P + F_P + F_N} \tag{3.10}$$

$$\text{Sensitivity} = \frac{T_P}{T_P + F_N} \tag{3.11}$$

$$\text{Specificity} = \frac{T_N}{T_N + F_P} \tag{3.12}$$

where T_P is the true positive, T_N is the false negative, $F_P$is false positive and $F_N$is false negative. The thickness difference is calculated by the thickness difference from each vertex along the normal vector between automated and manual segmentation masks.

3.4 Experiments and results

Table 3.1 summarizes the results of previous works ([2, 9, 10, 12, 21, 30]), baseline networks (nnU-Net [26, 31], including the 2D version, 3D full resolution version and 3D low resolution version) and the proposed methods for the SKI10 validation dataset in terms of SKI10 metrics [6]. The bone and cartilage segmentation results with proposed networks reached a total score of 76.2±7.6, which was the first time higher than 75 using the validation dataset. Overall, the results of deep learning based methods outperformed the traditional methods (atlas based [12], statistical shape based [9, 10, 30]). The new baseline (nnU-Net) could achieve state-of-the-art results without any post-processing. Still, the proposed method outperformed the baseline.

Table 3.1: Comparison of automatic segmentation methods based on SKI10 validation data (* indicates the deep learning related method; 'res' indicates resolution)

		Femur bone		**Tibia bone**		**Femur cartilage**		**Tibia cartilage**	
Team (reference)	Total score	AvgD (mm)	RMSD (mm)	AvgD (mm)	RMSD (mm)	VOE (%)	VD (%)	VOE (%)	VD (%)
Vincent et al. [10]	52.3 ± 8.6	0.88 ± 0.24	1.49 ± 0.44	0.74 ± 0.21	1.21 ± 0.34	36.3 ± 5.3	−25.2 ± 10.1	34.6 ± 7.9	74.0 ± 7.7
Seim et al. [9]	54.4 ± 8.8	1.02 ± 0.22	1.54 ± 0.30	0.84 ± 0.19	1.24 ± 0.28	34.0 ± 12.7	7.7 ± 19.2	29.2 ± 8.6	−2.7 ± 18.2
Shan et al. [12]	40.0 ± 7.7	-	-	-	-	-	-	-	-
***Liu et al. [2]**	64.1 ± 9.5	0.56 ± 0.12	1.08 ± 0.21	0.50 ± 0.14	1.09 ± 0.28	28.4 ± 6.9	8.1 ± 12.3	33.1 ± 7.1	−1.2 ± 17.4
Dam et al. [30]	67.1 ± 8.0	0.68 ± 0.22	1.25 ± 0.41	0.50 ± 0.18	0.91 ± 0.35	26.9 ± 6.0	0.8 ± 13.5	25.1 ± 6.7	0.41 ± 13.4

***Ambellan et al. [21]**	74.0 ± 7.7	0.43 ± 0.13	0.74 ± 0.27	0.35 ± 0.07	0.59 ± 0.19	20.99 ± 5.08	7.18 ± 10.51	19.06 ± 5.18	4.29 ± 12.34
***nnU-Net 2D [26]**	73.4 ± 10.7	0.37 ± 0.15	0.69 ± 0.35	0.38 ± 0.27	0.80 ± 0.77	21.34 ± 5.59	4.49 ± 11.46	21.43 ± 5.67	5.74 ± 13.41
***nnU-Net 3D full res [26]**	72.5 ± 14.2	0.56 ± 1.00	1.67 ± 2.96	0.44 ± 0.57	1.34 ± 2.46	19.45 ± 5.06	6.79 ± 10.29	18.09 ± 5.09	8.32 ± 11.31
***nnU-Net 3D low res [26]**	75.3 ± 9.3	0.35 ± 0.12	0.65 ± 0.30	0.34 ± 0.23	0.75 ± 1.19	21.72 ± 4.70	3.66 ± 12.14	21.78 ± 5.39	6.58 ± 12.11
***Proposed method**	76.2 ± 7.6	0.38 ± 0.15	0.69 ± 0.37	0.29 ± 0.07	0.52 ± 0.12	19.45 ± 5.06	6.78 ± 10.29	18.09 ± 5.09	8.32 ± 11.31

Moreover, Table 3.2 shows the accuracy evaluation for SKI10 dataset between the baseline networks and the proposed methods in terms of DSC, Sensitivity and Specificity. For the cartilage result, the DSC is only calculated in the defined ROI according to [6]. The DSC scores of the proposed method are 0.98 ± 0.01, 0.98 ± 0.01, 0.89 ± 0.03 and 0.88 ± 0.03 for femur bone, tibia bone, femur cartilage and tibia cartilage, respectively. Overall, the performance of the proposed methods achieved the highest score.

Table 3.2: Segmentation accuracy for the SKI10 validation dataset between baseline networks and the proposed methods (2D: nnU-Net 2D; 3D F: nnU-Net 3D full resolution; 3D L: nnU-Net 3D low resolution; DSC: dice similarity coefficient; Sens: sensitivity; Spec: specificity)

	Femur bone			**Tibia bone**			**Femur cartilage**			**Tibia cartilage**		
	DSC	Sens	Spec	DSC	Sens	Spec	DSC	Sens	Spec	DSC	Sens	Spec
2D	0.98 ± 0.01	0.98 ± 0.02	1.00 ± 0.00	0.98 ± 0.02	0.98 ± 0.03	1.00 ± 0.00	0.88 ± 0.04	0.90 ± 0.04	1.00 ± 0.00	0.86 ± 0.04	0.89 ± 0.06	1.00 ± 0.00
3D F	0.98 ± 0.01	0.98 ± 0.01	1.00 ± 0.00	0.98 ± 0.02	0.98 ± 0.03	1.00 ± 0.00	0.89 ± 0.03	0.92 ± 0.04	1.00 ± 0.00	088 ± 0.03	0.92 ± 0.04	1.00 ± 0.00
3D L	0.98 ± 0.01	098 ± 0.01	1.00 ± 0.00	0.98 ± 0.01	0.98 ± 0.02	1.00 ± 0.00	0.88 ± 0.03	0.89 ± 0.05	1.00 ± 0.00	0.86 ± 0.04	0.89 ± 0.05	1.00 ± 0.00
Proposed	0.98 ± 0.01	0.98 ± 0.01	1.00 ± 0.00	0.98 ± 0.01	098 ± 0.01	1.00 ± 0.00	0.89 ± 0.03	0.92 ± 0.04	1.00 ± 0.00	0.88 ± 0.03	0.92 ± 0.04	1.00 ± 0.00

Some segmentation results on the SKI10 validation set are shown in Figure 3.3-3.5, which compared the baseline networks with the proposed method. The results of nnU-Net 2D might mis-segment the low contrast region (bottom of Figure 3.3c), while the result of nnU-Net 3D full resolution might mis-segment some the unrelated region (left bottom of Figure 3.4c). A segmentation result of knee joint image with specific pathological tissue is depicted in Figure

3.6. All the baseline networks failed to segment it successfully, and the proposed method with the adversarial loss showed a robust result (Figure 3.6f).

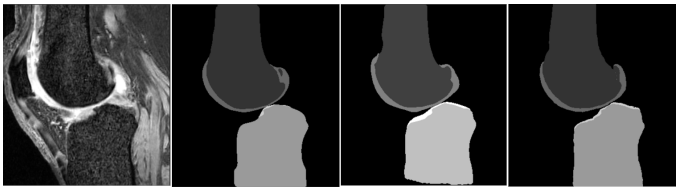


Figure 3.3: Segmentation results based on different schemes: (a) sagittal slice of the image; (b) ground truth; (c) nnU-Net 2D; (d) proposed method.

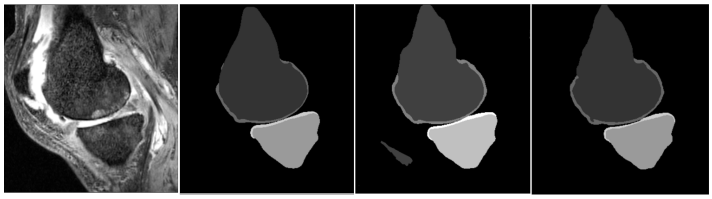


Figure 3.4: Segmentation results based on different schemes: (a) sagittal slice of the image; (b) ground truth; (c) nnU-Net 3D full resolution; (d) proposed method.

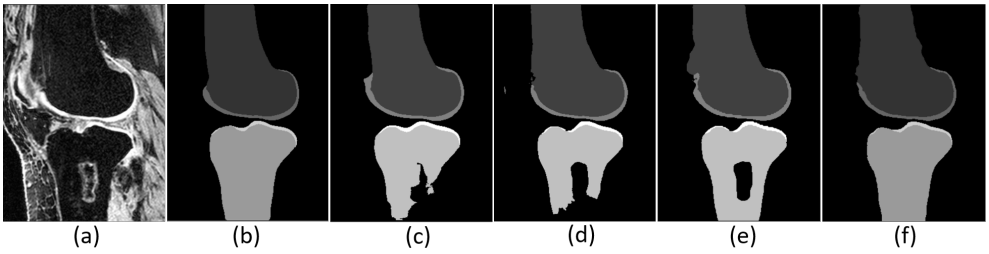


Figure 3.5: Segmentation results based on different schemes: (a) sagittal slice of the image; (b) ground truth; (c) nnU-Net 2D; (d) nnU-Net 3D full; nnU-Net 3D low; (d) proposed method.

Computation time for the whole segmentation pipeline for one subject is measured as around 1 min on a consumer-grade workstation (CPU: Intel Xeon E5 2.3GHz; GPU: GeForce GTX 1080 Ti).

3.5 Discussion

In this work, we presented an end to end deep learning based workflow for knee bone and cartilage segmentation and evaluated the workflow thoroughly on a published dataset, SKI10 [6]. It was the first time the total score greater than 76 achieved on the SKI10 validation dataset, which was comparable to the inter-observer variability of two expert readers [6].

The attempt of applying deep learning-based methods to knee bone and cartilage segmentation was not new and have achieved a lot of state-of-the-art results [2, 21]. Nevertheless, most of the previous attempts added a post-processing step (deformable model [2], conditional random field (CRF) [20], statistical shape model (SSM) [21]) to refine the outcome of the deep learning methods on the area of false segmentation. The main reason behind this is the information of highly patient-specific area might not be derived from the training dataset [21]. To confirm the necessity of the post-processing, a generic UNet architecture with fine-tuned hyperparameter [31] were tested in this study as the baseline. State-of-the-art results can be achieved using the simple nnU-Net architectures (see Table 3.1). Nonetheless, due to the loss of z-axis information and contextual information, the performance of bone segmentation of generic 2D nnU-Net and 3D nnU-Net full resolution might perform poorly in the low contrast region (Figure 3.3c), bone or cartilage liked region (Figure 3.4c) and region with pathological case (Figure 3.6c-d). Table 3.1 has shown a good bone segmentation result using 3D nnU-Net low resolution, but not in cartilage segmentation. This is because the target volume of cartilage is relatively small, which resulted in the loss of the cartilage information, especially in the pathological area. In this sense, whether resampling the image volume to improve the segmentation performance should be considered carefully based on the size of the target and the memory of the GPU.

Moreover, Figure 3.6c-e have shown that all the nnU-Net architectures fail to segment the bone with the specific pathological feature, which demonstrates the necessity of post-processing from previous studies. In this study, we introduced the adversarial loss to serve as a shape regulation penalty to improve the bone segmentation. Although the adversarial loss [32] have been proposed to improve the segmentation performance previously, to the best of our knowledge, it was the first time to serve as a shape consistency term to apply on knee bone MRI image segmentation. Figure 3.6f have shown that the introduction of adversarial loss results in state-of-the-art results for bone segmentation despite with the pathological case.

This study has a number of limitations. First of all, due to the limited memory of Nvidia 1080 Ti, the number of feature channel of the first layer in nnU-Net experiments is 20 rather than 30 stated in [31]. Further experiments with a better GPU should be implemented to investigate the performance influence of the number of feature channel. An additional limitation is there are still a lot of ablation studies can be implemented to discuss the segmentation performance based on different choice of hyper-parameters. Nevertheless, we believe the experiment results are enough to share with the community to help the development of fully automatic segmentation of the knee joint. Moreover, the bone segmentation was segmented in a relatively lower resolution in order to enlarge context information. Isensee et al. proposed a cascaded mode to further improve the low-resolution

segmentations [26]. However, the training data for these two networks should be different; otherwise it will easily result in an overfitted network. As Isensee et al. stated that the cascaded mode was not so much better than the 3d_lowres and 3d_full_res mode in most case [33], we believe the results of 3d_lowres and 3d_full_res are sufficient to be a baseline and we will add the comparison with cascaded mode in the future when more annotated dataset is available.

3.6 Conclusion

To conclude, we presented a robust pipeline to segment the knee bone and cartilage. The result of the proposed method is the first time achieved more than 76 in a well known dataset, SKI10 validation for our best knowledge. The lower resolution strategy and the introduction of adversarial loss improve the shape consistency of the bone segmentation, while a fined-tuned V-Net network was further boosted to achieve a promising result for the cartilage segmentation. Future studies will include segmentation for more knee joint structure, such as ligaments, menisci.

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW).

References

[1] K. Zhang, W. Lu, and P. Marziliano, "Automatic knee cartilage segmentation from multi-contrast MR images using support vector machine classification with spatial dependencies," *Magnetic Resonance Imaging,* vol. 31, pp. 1731-1743, 2013, doi: 10.1016/j.mri.2013.06.005.

[2] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, and R. Kijowski, "Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging," *Magnetic Resonance in Medicine,* vol. 79, pp. 2379-2391, 2018, doi: 10.1002/mrm.26841.

[3] P. A. Yushkevich *et al.*, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage,* vol. 31, pp. 1116-1128, 2006, doi: 10.1016/j.neuroimage.2006.01.015.

[4] H. Shim, S. Chang, C. Tao, J. H. Wang, C. K. Kwoh, and K. T. Bae, "Knee cartilage: efficient and reproducible segmentation on high-spatial-resolution MR images with

the semiautomated graph-cut algorithm method," *Radiology,* vol. 251, no. 2, pp. 548-56, May 2009, doi: 10.1148/radiol.2512081332.

[5] E. J. McWalter *et al.*, "Use of novel interactive input devices for segmentation of articular cartilage from magnetic resonance images," *Osteoarthritis and Cartilage,* vol. 13, pp. 48-53, 2005, doi: 10.1016/j.joca.2004.09.008.

[6] T. Heimann and B. Morrison, "Segmentation of knee images: A grand challenge," 2010.

[7] T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis,* vol. 13, pp. 543-563, 2009, doi: 10.1016/j.media.2009.05.004.

[8] 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, pp. 55-64, 2010, doi: 10.1109/TMI.2009.2024743.

[9] H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, "Model-based auto-segmentation of knee bones and cartilage in MRI data," *MICCAI Workshop Medical Image Analysis for the Cinic,* pp. 215-223, 2010.

[10] G. Vincent, C. Wolstenholme, I. Scott, and M. Bowes. Fully Automatic Segmentation of the Knee Joint using Active Appearance Models.

[11] 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, p. 092303, 2014, doi: 10.1118/1.4893533.

[12] L. Shan, C. Zach, C. Charles, and M. Niethammer, "Automatic atlas-based three-label cartilage segmentation from MR knee images," *Medical Image Analysis,* vol. 18, pp. 1233-1246, 2014, doi: 10.1016/j.media.2014.05.008.

[13] J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen, "Segmenting Articular Cartilage Automatically Using a Voxel Classification Approach," *IEEE Transactions on Medical Imaging,* vol. 26, pp. 106-115, 2007, doi: 10.1109/TMI.2006.886808.

[14] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Transactions on Medical Imaging,* vol. 35, pp. 1299-1312, 2016, doi: 10.1109/TMI.2016.2535302.

[15] K. Kamnitsas, C. Ledig, and V. F. J. Newcombe, "Efficient Multi-Scale 3D CNN with fully connected CRF for Accurate Brain Lesion Segmentation," vol. 36, pp. 61-78, 2017.

[16] D. M. Vigneault, W. Xie, C. Y. Ho, D. A. Bluemke, and J. A. Noble, "Ω-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks," *Medical Image Analysis,* vol. 48, pp. 95-106, 2018, doi: 10.1016/j.media.2018.05.008.

[17] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation.

[18] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation.

[19] A. Prasoon and E. Al., "Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolution Neural Network," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture*

Notes in Bioinformatics), vol. 8150, pp. 599-606, 2013, doi: 10.1007/978-3-642-40763-5.

[20] Z. Zhou, G. Zhao, R. Kijowski, and F. Liu, "Deep convolutional neural network for segmentation of knee joint anatomy," *Magnetic Resonance in Medicine,* vol. 80, pp. 2759-2770, 2018, doi: 10.1002/mrm.27229.

[21] F. Ambellan, A. Tack, M. Ehlke, and S. Zachow, "Automated Segmentation of Knee Bone and Cartilage combining Statistical Shape Knowledge and Convolutional Neural Networks: Data from the Osteoarthritis Initiative," *Medical Image Analysis,* vol. 52, pp. 109-118, 2018, doi: 10.1016/j.media.2018.11.009.

[22] X. Yi, E. Walia, and P. Babyn, "Generative Adversarial Network in Medical Imaging: A Review," 2018.

[23] K. Kamnitsas, L.-C. Chen, C. Ledig, D. Rueckert, and B. Glocker, "Multi-scale 3D convolutional neural networks for lesion segmentation in brain MRI," *Proceedings of MICCAI-ISLES 2015,* pp. 13-16, 2015.

[24] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges," *Journal of Digital Imaging,* vol. 3, pp. 582-596, 2019, doi: 10.1007/s10278-019-00227-x.

[25] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping,* vol. 17, pp. 143-155, 2002, doi: 10.1002/hbm.10062.

[26] F. Isensee *et al.* nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation.

[27] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017,* vol. 2017-Janua, pp. 5967-5976, 2017, doi: 10.1109/CVPR.2017.632.

[28] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," pp. 1-11, 2016, doi: 10.1109/3DV.2016.79.

[29] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least Squares Generative Adversarial Networks," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv161104076M

[30] E. B. Dam, M. Lillholm, J. Marques, and M. Nielsen, "Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative," *Journal of Medical Imaging,* vol. 2, p. 024001, 2015, doi: 10.1117/1.JMI.2.2.024001.

[31] F. Isensee, J. Petersen, S. A. A. Kohl, P. F. Jäger, and K. H. Maier-Hein. nnU-Net: Breaking the Spell on Successful Medical Image Segmentation.

[32] D. Yang *et al.*, "Automatic liver segmentation using an adversarial image-to-image network," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* vol. 10435 LNCS, pp. 507-515, 2017, doi: 10.1007/978-3-319-66179-7_58.

[33] F. Isensee. "Problem with Cascade Training." https://github.com/MIC-DKFZ/nnUNet/issues/33 (accessed 2019.09.01, 2019).

Supplemental Material

a) The detail of bone network:

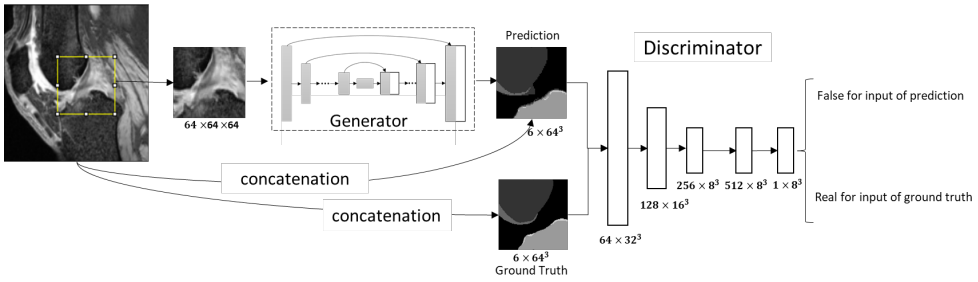


Figure 3.6: The architecture of the bone network (the architecture of the generator is depicted in Figure 3.7; In the discriminator, first two layers of discriminator follow the module form of Conv3-LeakyRelu, the third and fourth layers follow module form of Conv3-InstanceNorm-LeakyRelu, and the last layer is with only Conv3).

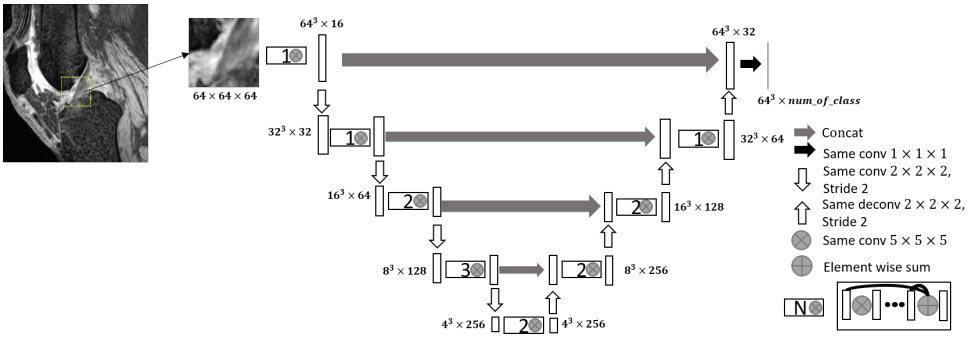


Figure 3.7: The detail architecture of the generator of bone network.

b) The detail of cartilage network:

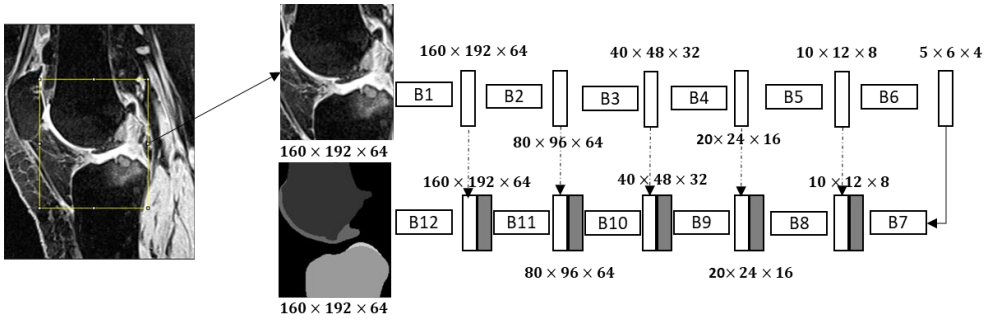


Figure 3.8: The architecture of the cartilage network (The layers of the block follow the module form of Conv3-InstanceNorm-LeakyRelu).

The architecture of the cartilage network is depicted in Figure 3.8, and the detail of each block is illustrated in Table 1.

Table 3.3: Parametric setting of blocks in neural network

Block	Kernel	Stride	Padding	Feature	Block	Kernel	Stride	Padding	Feature
B1	Conv (3,3,1)	(1,1,1)	(1,1,0)	1-20	B8	Conv (3,3,3)	(1,1,1)	(1,1,1)	640-320
	Conv (3,3,1)	(1,1,1)	(1,1,0)	20-20		Conv (3,3,3)	(1,1,1)	(1,1,1)	320-320
B2	Conv (3,3,3)	(2,2,1)	(1,1,1)	20-40		Deconv (2,2,2)	(2,2,2)	-	320-160
	Conv (3,3,3)	(1,1,1)	(1,1,1)	40-40	B9	Conv (3,3,3)	(1,1,1)	(1,1,1)	320-160
B3	Conv (3,3,3)	(2,2,2)	(1,1,1)	40-80		Conv (3,3,3)	(1,1,1)	(1,1,1)	160-160
	Conv (3,3,3)	(1,1,1)	(1,1,1)	80-80		Deconv (2,2,2)	(2,2,2)	-	160-80
B4	Conv (3,3,3)	(2,2,2)	(1,1,1)	80-160	B10	Conv (3,3,3)	(1,1,1)	(1,1,1)	160-80
	Conv (3,3,3)	(1,1,1)	(1,1,1)	160-160		Conv (3,3,3)	(1,1,1)	(1,1,1)	80-80
B5	Conv (3,3,3)	(2,2,2)	(1,1,1)	160-320		Deconv (2,2,2)	(2,2,2)	-	80-40
	Conv (3,3,3)	(1,1,1)	(1,1,1)	320-320	B11	Conv (3,3,3)	(1,1,1)	(1,1,1)	80-40
B6	Conv (3,3,3)	(2,2,2)	(1,1,1)	320-320		Conv (3,3,3)	(1,1,1)	(1,1,1)	40-40
	Conv (3,3,3)	(1,1,1)	(1,1,1)	320-320		Deconv (2,2,1)	(2,2,1)	-	40-20
B7	Deconv (2,2,2)	(2,2,2)	-	320-320	B12	Conv (3,3,3)	(1,1,1)	(1,1,1)	40-20
						Conv (3,3,3)	(1,1,1)	(1,1,1)	20-20
						Conv (3,3,3)	(1,1,1)	(1,1,1)	20-5

c) The detail of restoration network:

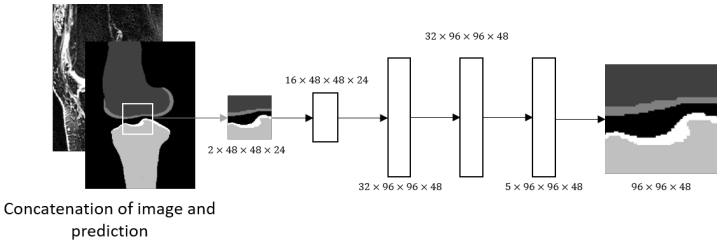


Figure 3.9: The architecture of the restoration network (the first, third and fourth layers follow the module form of Conv3-BatchNorm-RELU; the second layer follows the module form of Deconv3-BatchNorm-RELU).

Chapter 4 Knee cartilage estimation based on knee bone geometry using posterior shape model: a feasibility study

Abstract

Introduction: Computed tomography (CT) is an often-preferred image modality for knee joint related disease. Although the bone geometry can be clearly identified by the CT image, cartilage is hardly visible, which prevents providing the diagnostic potential of cartilage status and construction of 3D knee joint models from CT images. In this study, we propose to use posterior shape model to estimate the cartilage thickness only based on bone geometry with a public dataset.

Methods: Bone shape model and bone-cartilage shape model were constructed by 99 fold dataset from a public dataset SKI10 in a leave-one-out experiment. The cartilage information in the tibiofemoral contact area and the surgery related area was estimated based on bone geometry only information using posterior shape model. State-of-the-art methods were reimplemented and compared with the proposed method.

Results: The experiments showed that the cartilage thickness was correlated across the articular surface and can be predictable if part of cartilage thickness was known. Bone gap information at tibiofemoral contact area can be used as bone geometry for cartilage thickness prediction, with a root mean square error of 0.64 mm for femur and 0.58 mm for tibia.

Conclusions: The initial results in this study showed that the proposed method can provide a reasonable cartilage information in the situation that the image information of the cartilage was absent.

4.1 Introduction

Construction of the knee cartilage model is a topic of increasing interest as its application continue to broaden from knee joint pain analysis, to pre- and intra- operative planning. In clinical practice, computed tomography (CT) is an important technique for knee replacement surgery. However, as the cartilage information is hardly visible from the X-ray based imaging [1], the value of the CT image for the construction of cartilage model is limited. Moreover, a lot of patients with knee pain accompanied with a defect of cartilage, and thus it is of great benefit to the surgeon that the original shape of the cartilage could be estimated given the healthy part of the cartilage. Therefore, a method that could estimate the cartilage model based on subject-specific information, e.g. bone geometry, cartilage thickness at the healthy part, etc., is therefore an unmet need.

A lot of studies have shown that the characteristic patterns of the cartilage thickness can be found on the different region of interest (ROI) of articular surfaces of the knee [2-4]. Hence, a mean cartilage thickness and its standard deviation from statistical shape model (SSM) are useful representatives of the knee cartilage model [2]. Although the cartilage thickness distribution follows characteristic patterns across the articular surface, the individualistic relationship between bone geometry and cartilage thickness were still under debate [1, 5]. Koo et al. reported that cartilage thickness was not correlated to the individual variation in contact geometry, but mechanical stress due to the geometries of the contact surfaces appeared to influence the general thickness distribution in the medial and lateral compartments of the knee [5]. However, in contrast to Koo's studies, Van Dijck et al. recently reported the knee cartilage model could be predicted solely by the bone shape information with sufficient accuracy [1].

From a technical perspective, a lot of researches attempted to validate the feasibility of predicting the cartilage information given the bone geometry. Although the mean cartilage thickness distribution model provided a good guess of the cartilage shape, this was not patient-specific [1, 6]. Johnson proposed two steps to optimize the mean cartilage thickness distribution model [6]: first, the cartilage thickness was estimated based on the mean thickness of cartilage SSM; second, the overlap area between the cartilage models was identified and a reduced cartilage thickness was updated until no overlap was found. Recently, Van Dijck et al. constructed the bone SSM and bone-cartilage SSM separately, and assumed the two SSMs shared the same coefficient matrix [1]. From this assumption, once the coefficients of the bone model were identified by the new input bone shape model, the bone-cartilage model could be also calculated based on the coefficients of bone shape model and the bone-cartilage SSM.

In this study, we aim to use posterior shape model [7] to predict tibiofemoral cartilage shape based on the bone geometry using a public MRI dataset [8]. We hypothesize that: a) the individual cartilage thickness across the articular surface is correlated and can be predicted by posterior shape model given part of true cartilage thickness; (b) the bone gap distance in tibiofemoral contact area could be used as bone geometry for the cartilage prediction.

The contributions in this paper are:

(1) we propose to directly estimate the cartilage thickness solely by bone geometry using posterior shape model;

(2) we propose to include the bone gap information as an information of bone geometry to estimate the cartilage;

(3) we construct and test the proposed model based on a public dataset;

(4) we compare the proposed model with state-of-art methods.

4.2 Materials and methods

4.2.1 Data description

The data used in this study was from the Segmentation of Knee Images 2010 (SKI10) competition which was focus on the knee bone and cartilage segmentation [8]. The image dataset were acquired in a sagittal manner with a pixel spacing of 0.4×0.4 mm and a slice thickness of 1 mm. The total number of the knee images used in this study were 100 (Training set A, B and Testing set C) and the cases of left and right knees are distributed approximately equally.

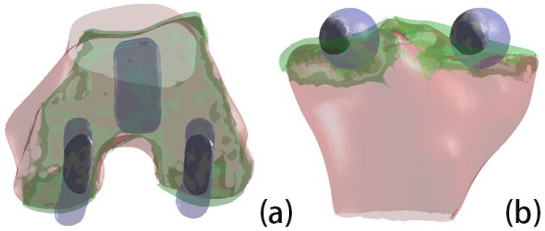


Figure 4.1: Definition of regions of interest (ROI) of femoral (a) and tibial (b) cartilage (red: bone; green: cartilage; purple: surgery related ROIs defined by the experts; black: tibiofemoral contact ROI).

As the dataset was used for the surgery planning, Heimann et al. pointed out that the cartilage mask was accurate in the region of implant position guides should be placed, and thus, the surgery related region of interest (ROI) was defined as the purple regions for femur

and tibia in Figure 4.1a-b [8]. Moreover, we defined the tibiofemoral contact area as the overlap region between the surgery ROIs of femoral and tibial (black regions Figure 4.1a-b).

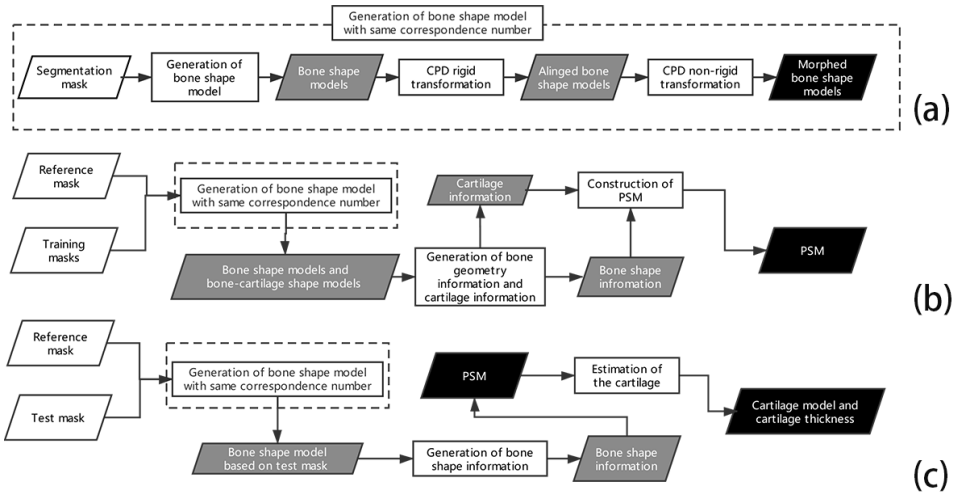


Figure 4.2: Workflow for the proposed cartilage estimation: (a) generation of bone shape model with same correspondence number; (b) stage of model construction; (c) stage of inference (white, grey and black parallelogram indicate input data, medium result, and final result; white box indicates method; CPD: coherent point drift [9]; PSM: point shape model [7]).

4.2.2 Methods

As Figure 4.2 illustrates, the proposed workflow consisted of two stages (model construction stage (b) and inference stage (c)). The first step for both stages was to generate the bone shape models with the same correspondences (Figure 4.2a). In the phase of generation of bone shape model, the segmented mask was firstly converted to bone shape models. Next, a rigid transformation was performed to align the target bone shape model to reference bone shape model. Then, a non-rigid transform was used to morph the reference shape model to the target shape model in order to keep the same correspondences among shape models. The cartilage thickness and bone geometry information were then calculated based on the bone shape models and the bone-cartilage models (the shape model consists of bone and cartilage). With the bone geometry information and their corresponding cartilage information, a posterior shape model (PSM) was established. Finally, the cartilage information for a new subject could be estimated based on the new input bone shape information using PSM.

4.2.2.1 Data Pre-processing

a) Generation of shape models

Firstly, we mirrored all the right knee image in the dataset, which means we regarded the side of all knee bones image as left. Then, the shape model was generated from the mask (Figure 4.3a to Figure 4.3b), followed by the smoothing method using [10] to avoid the discontinuity caused by manually slice by slice segmentation (Figure 4.3b to Figure 4.3c).

Secondly, we created the bone-cartilage shape model by converting the bone and cartilage masks together to the shape model (Figure 4.3d to Figure 4.3e and Figure 4.3f). As Figure 4.3e and Figure 4.3f illustrate, the bone and bone-cartilage shape model were intersected in the bone area, while the bone-cartilage shape model covered the bone shape model in the cartilage area.

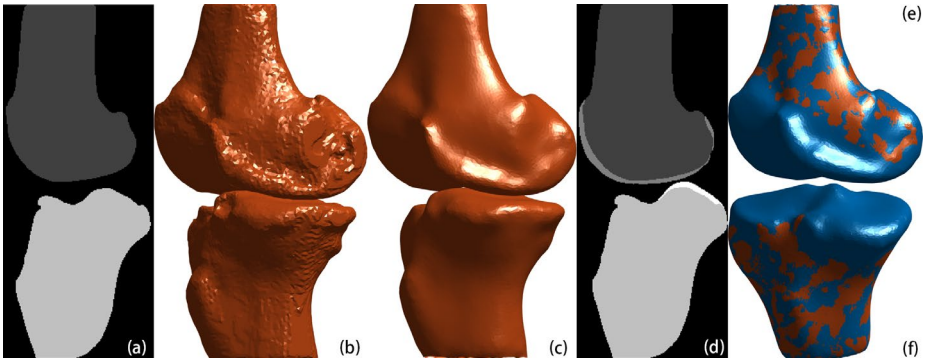


Figure 4.3: Generation of the shape models: (a) femur and tibia masks; (b) bone shape model; (c) smoothed bone shape model; (d) femur, femur cartilage, tibia and tibia cartilage masks; combination of bone (orange) and cartilage (blue) shape models ((e) femur, (f) tibia)

b) Rigid and non-rigid registration among the bone shape models

The first step to construct the statistical shape model was to establish the correspondences among the shape models [11]. All the registration used in this study were based on coherent point drift (CPD) framework [9]. The key idea of CPD was to force the Gaussian mixture model (GMM) centroids (representing the first point set) transform coherently (preserving the topological shape) to the data (the second point set) by maximizing the likelihood. We recommend [9] for the further details.

Reference was randomly chosen from dataset and the correspondences among the bone shape models were determined in two steps. First, a rigid registration was applied from the target bone shape to the reference bone shape. To avoid the low quality of rigid registration caused by the different bone size, both target bone and reference bone were normalized into same scale (Figure 4.4a to Figure 4.4b). After the rigid registration, the registered target bone was de-normalized back to original shape (Figure 4.4c). Next, as Figure 4.4d illustrates, the non-rigid registration is applied to deform the reference bone shape to the target bone shape.

In other words, the target bone shape now could be represented by the point set of the reference bone shape. Once the correspondences were confirmed, the bone shapes could be represented by a linear combination of the training shapes [7, 11].

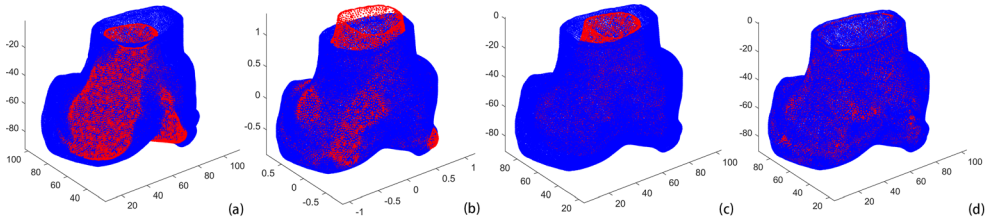


Figure 4.4: Illustration of the registrations (red point cloud: reference object; blue: target object): (a) before registration; (b) normalized rigid registration; (c) de-normalized back to original size; (d) non-rigid registration

c) Measurement of cartilage thickness

In this study, the bone-cartilage model was created from the bone and cartilage mask together (Figure 4.3e and Figure 4.3f). The cartilage thickness of each vertex of bone shape model was estimated based on the intersect point between the bone-cartilage surface and its normal vector. Figure 4.5a and Figure 4.5b illustrate the normal vector of the bone surface with its corresponding thickness as magnitude for femur and tibia. Figure 4.5c and Figure 4.5d illustrate the cartilage thickness map of femur and tibia respectively.

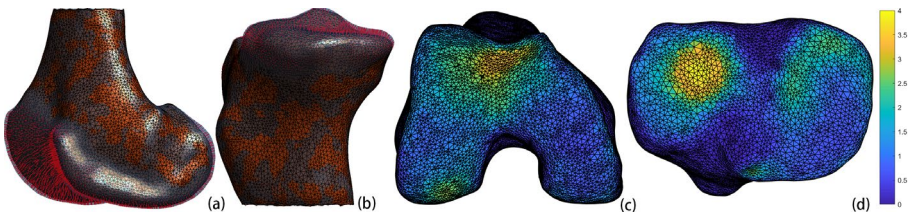


Figure 4.5: (a)-(b) orange and blue indicate the bone and cartilage shape model respectively, red arrow indicates the cartilage thickness with different magnitude; (c)-(d) bone shape model with cartilage thickness with different colours.

d) Definition of bone gap information

The bone gap information in the tibiofemoral contact ROI was introduced in this study as one of the bone geometries. As Figure 4.6 illustrates, femur bone gap information was defined as the closest distance from vertices in the contact region of femur to the tibia surface (black line), same as the definition of tibia bone gap information.

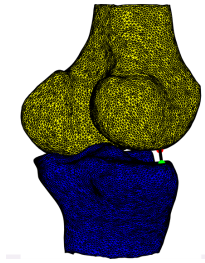


Figure 4.6: Illustration of bone gap information (black line), the closest point from vertex (red point) of femur bone to tibia surface (green point)

4.2.2.2 Posterior shape model

a) Construction of shape vectors for PSM

Instead of building the SSM based on the coordinates of shape models [12], we encoded the model with the prior knowledge part and the estimated part as

$$\boldsymbol{s}_i = \left(v_{p,1}, v_{p,2}, \dots, v_{p,t}, v_{e,1}, v_{e,2}, \dots v_{e,m}\right), i = 1,2, \dots n \tag{4.1}$$

Where $v_{p,t}$ indicates the t[th] prior factor, which can be the 3D coordinates of bone shape model, bone gap information, etc.; and $v_{e,m}$ indicates the m[th] information needed to be estimated, which can be the cartilage thickness or the 3D coordinates of the cartilage shape model.

b) PSM for estimation of cartilage

Traditionally, the training shape vectors can be represented with a compact form $\boldsymbol{s}_i = \boldsymbol{\mu} + \mathbf{UD}\boldsymbol{\alpha}$ by coefficient vector α, sample mean vector $\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{s}_i$ and the result of principal component analysis (PCA) on the covariance matrix $\boldsymbol{\Sigma} = \mathbf{UD^2U}^T$, where $\mathbf{U}$ contains the orthonormal eigenvectors or "principal components" of $\boldsymbol{\Sigma}$. Although the PCA models are enough to represent the shape variance, it cannot perform like a posterior model, such as predict the cartilage information from bone geometry. Albrecht et al. discussed how to predict the shape related information by a posterior shape model using posterior PCA model, probabilistic PCA model, and Gaussian process model and demonstrated their relationship in previous study [7]. In this study, we made use of the probabilistic PCA model to find the relationship between the bone shape information and cartilage shape information. The cartilage information was predicted as the posterior mean,

$$\boldsymbol{\mu}_c = \boldsymbol{\mu} + \boldsymbol{Q}_{ML}\left(\boldsymbol{Q}_g^T\boldsymbol{Q}_g + \sigma^2 I_n\right)^{-1}\boldsymbol{Q}_g^T(\boldsymbol{s}_g - \boldsymbol{\mu}_g) \tag{4.2}$$

$$\boldsymbol{Q}_{ML} = \boldsymbol{U}(\boldsymbol{D}^2 - \rho^2 \boldsymbol{I}_p)^{\frac{1}{2}} \tag{4.3}$$

$$\sigma = \sqrt{\frac{1}{t}\left\|\boldsymbol{s}_g - \hat{\boldsymbol{s}}_g\right\|_2} \tag{4.4}$$

$$\rho = \sqrt{\frac{1}{n(t+m)}\sum_{i=1}^{n}\|\boldsymbol{s}_i - \hat{\boldsymbol{s}}_i\|_2^2} \tag{4.5}$$

Where $\boldsymbol{Q}_g = \boldsymbol{U}_g\boldsymbol{D}_g$, which $\boldsymbol{U}_g$ and $\boldsymbol{D}_g$ are selecting entries and rows of prior knowledge part from the model's full $\boldsymbol{U}$ and $\boldsymbol{D}$. $\boldsymbol{s}_g$ is target vectors of the prior knowledge, and $\boldsymbol{\mu}_g$is the mean vector of the training dataset of prior knowledge. $\hat{\boldsymbol{s}}_g$ is the best approximation of the target information based on the model. t and m are the number of dimension of the prior knowledge part and the estimated part, and n is the number of the training samples. The predicted part of information is the subset of $\boldsymbol{\mu}_c$, in other word $v_{e,m}$. We recommend [7, 13], for further details of derivation.

4.3 Experiments

In this study, the number of vertices to construct the bone shape model and bone-cartilage model was 11255 for the femur and 8855 for the tibia. The number of bone vertices in the surgery ROI was 2065 for the femur and for 645 for the tibia. The number of bone vertices in the tibiofemoral contact ROI was 455 for the femur and 393 for the tibia.

The posterior shape model (PSM) were constructed using leave-one-out experiments, where one case was excluded from the training dataset and the PSM was constructed based on the rest of data.

4.3.1 Proposed experiments

The baseline accuracy of the cartilage estimation was based on the mean cartilage thickness distribution, i.e. the cartilage thickness of each vertices of the shape model was identified by an averaged cartilage thickness of the corresponding vertices among all the training bone shape models. Moreover, we re-implemented Van Dijck's estimation method [1] for the further comparison of the proposed method. To testify the effectiveness of the proposed method, we implemented three experiments as:

Firstly, we testified whether we could estimate part of cartilage thickness by providing part of true cartilage thickness, i.e. estimation of the cartilage in surgery related ROI (purple

area but not including the black area in Figure 4.2a-b) by providing the cartilage information on tibiofemoral contact ROI (black area in Figure 4.2a-b).

Secondly, we testified whether we could estimate the cartilage thickness on the tibiofemoral contact ROI by providing the bone geometry from the tibiofemoral contact ROI, which included three situation: 1) the bone gap information in tibiofemoral contact was provided as prior information to predict the cartilage thickness from tibiofemoral contact ROI; 2) the bone shape coordinates in tibiofemoral contact ROI was provided as prior information to predict the cartilage thickness from tibiofemoral contact ROI; 3) the bone shape coordinates from tibiofemoral contact ROI was used to predict the cartilage coordinates from tibiofemoral contact ROI; 4) the bone shape model was used to predict the bone-cartilage shape model with Van Dijck's explicit model [1] on tibiofemoral contact ROI.

Thirdly, we testified whether we could estimate the cartilage thickness on the surgery related ROI by providing the bone geometry from surgery related ROI. The experiments were similar to the second one, but just replacing the prior information from tibiofemoral contact ROI to surgery related ROI.

4.3.2 Evaluation metrics

To ensure a fair comparison, we adopted the similar evaluation metrics from the previous study [1], which includes the compactness, generalization ability and mean root mean square error.

a) Compactness

Compactness is a measure of the quality of the data reduction process, which is calculated as

$$\mathrm{C}(m) = \frac{\sum_{n=1}^{m} \lambda_n}{\sum_{n=1}^{d} \lambda_n} \tag{4.6}$$

Where m indicates the number of the mode, λ_n indicates the eigen value of the covariance and d is the number of the training samples.

b) generalization ability

Generalization ability was performed using leave-one-out (LOO) reconstruction experiments and reflects the ability of the model to generate new instances. For each left-out instance, the best approximation was computed from the remaining training dataset. The generalization ability of the model is calculated as

$$G(s) = \frac{1}{d}\|s - \hat{s}\|_2^2 \tag{4.7}$$

Where s is the target shape model, $\hat{s}$ is the best approximation of this data set within the model span based on SSM, and d is the dimension of the model.

(c) Mean root mean square error

The accuracy of cartilage thickness is calculated by the point-to-surface root mean squared error (RMSE) from the cartilage surfaces and averaged across the region of interest. Furthermore, the bootstrapped 95% confidence intervals are also included.

$$\text{RMSD} = \frac{\sqrt{\sum_{n=1}^{N} d_n^2}}{N} \tag{4.8}$$

All the evaluation in this study were performed using leave one out (LOO) experiment. The compactness and generalization ability were used to evaluate the goodness of the built model, while the root mean square error was used to quantify the accuracy of the prediction.

4.4 Results

The compactness and the generalization ability of the bone and bone-cartilage model based on the different variation modes of the model, i.e. the different amount of eigen vectors used to represent the model, is depicted on Figure 4.7. The average compactness achieved higher than 0.99 after 50 eigen vectors were used for both femur and tibia. The average generalization ability reduced to around 0.2 mm after 40 eigen vectors were used. The generalization ability in PSM at mode of 98 on bone model were 0.18 and 0.19 for femur and tibia respectively. The generalization ability at mode of 98 on bone-cartilage model were 0.18 and 0.18 for femur and tibia, respectively.

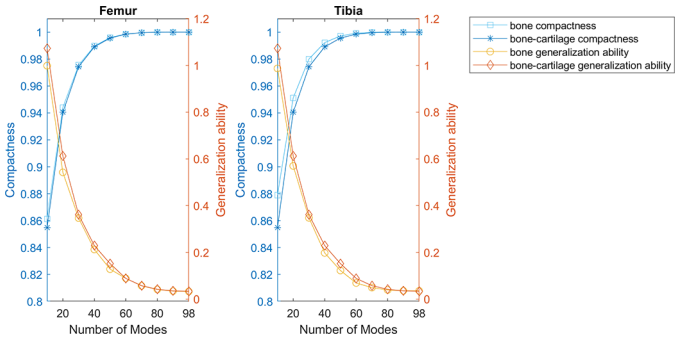


Figure 4.7: Compactness and generalization ability of the bone model and bone-cartilage model

The results of estimation of cartilage thickness in surgery related ROI (not including the tibiofemoral contact ROI) by providing the true cartilage thickness in tibiofemoral contact ROI are depicted in Figure 4.8. The application of mean cartilage thickness distribution method resulted in a RMSE of 0.77 mm for femur and 0.74 for tibia. Application of PSM to the cartilage thickness estimation resulted in an RMSE from 0.91 mm to 0.64 mm for femur and from 0.48 mm to 0.47 mm to for tibia, respectively, when the number of modes was raised from 10 to 98. The estimated results for femur converged after 40 variation modes, while the results for tibia remained a higher accuracy with all variation modes.

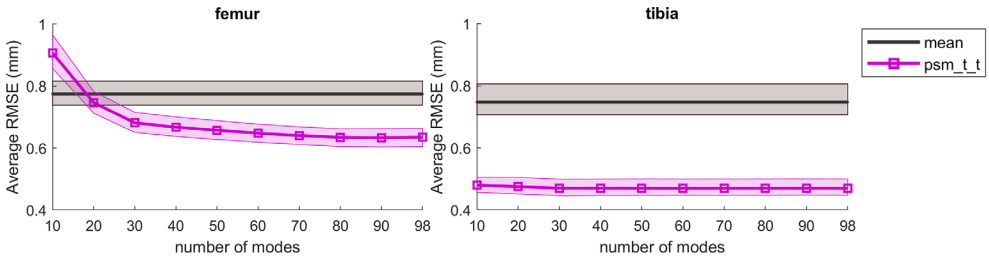


Figure 4.8: Average RMSE of cartilage thickness estimation of surgery related ROI based on ground true input from tibiofemoral contact ROI for femur (left) and tibia (right) (average as line and shaded area as bootstrapped 95% confidence interval, same as Figure 4.9 and Figure 4.10).

The results of estimation of cartilage thickness in tibiofemoral contact ROI by providing the bone geometry in tibiofemoral contact ROI as prior information are depicted in Figure 4.9. The mean cartilage thickness distribution method resulted in an RMSE of 0.70 mm for femur and 0.77 for tibia. The application of PSM to the cartilage thickness estimation by providing the bone gap information in tibiofemoral contact ROI resulted in an average RMSE from 0.79 mm to 0.64 mm for femur and from 0.80 mm to 0.59 mm to for tibia, respectively. The prediction of cartilage thickness based on bone coordinates in tibiofemoral contact ROI resulted in an average RMSE from 0.99 mm to 0.71 mm for femur and from 1.10 mm to 0.80 mm to for tibia. The prediction of cartilage coordinates based on bone coordinates in tibiofemoral contact ROI resulted in an average RMSE from 0.79 mm to 0.70 mm for femur and from 0.90 mm to 0.76 mm to for tibia. The difference of the output between the cartilage thickness and cartilage coordinate based on prior information of bone coordinates were small, but the mode of cartilage coordinates achieved a faster convergency. Moreover, the average RMSE of best prediction based on the explicit model from [1] was 0.72 mm for femur and 0.79 for tibia in the tibiofemoral contact ROI. The results of Van Dijck's method appeared the overfitting and achieved the optimal point around mode 20 - 30.

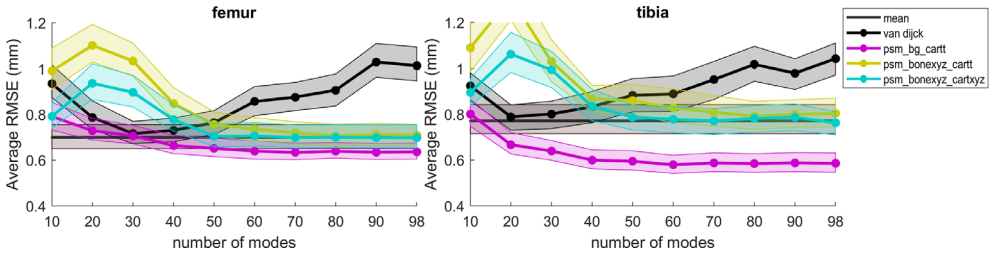


Figure 4.9: Average RMSE of cartilage thickness estimation in tibiofemoral contact ROI based on mean thickness distribution method (square black line), Van Dijck's method (black circle), proposed PSM (using bone gap information of tibiofemoral contact ROI as prior information for cartilage thickness estimation (purple), using bone coordinates of tibiofemoral contact ROI as prior information for cartilage thickness estimation (yellow), and using bone coordinates of tibiofemoral contact ROI as prior information for cartilage coordinates estimation (cyan); bg_cartt indicates using bone gap information to estimate cartilage thickness, bonexyz_cartt indicates using bone coordinates to estimate cartilage thickness, bonexyz_cartxyz indicates using bone coordinates to estimate cartilage coordinates)

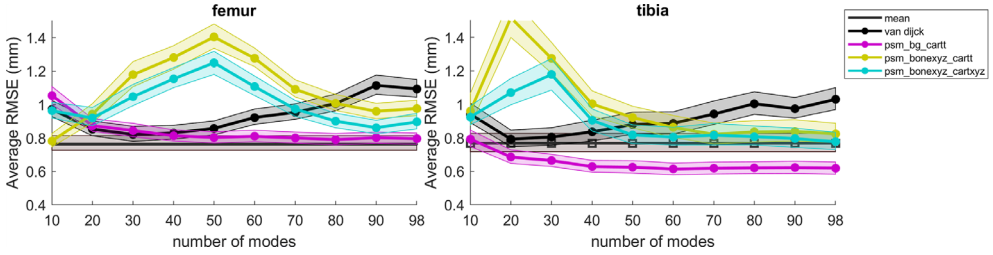


Figure 4.10: Average RMSE of cartilage thickness estimation in surgery related ROI based on mean thickness distribution method (square black line), Van Dijck's method (black circle), proposed PSM (using bone gap information in surgery related ROI as prior information for cartilage thickness estimation (purple), using bone coordinates in surgery related ROI for cartilage thickness estimation (yellow), and using bone coordinates in surgery related ROI for cartilage coordinates estimation (cyan); bg_cartt indicates using bone gap information to estimate cartilage thickness, bonexyz_cartt indicates using bone coordinates to estimate cartilage thickness, bonexyz_cartxyz indicates using bone coordinates to estimate cartilage coordinates)

The results of estimation of cartilage thickness in surgery related ROI by providing the bone geometry in surgery related ROI are depicted in Figure 4.10. The mean cartilage thickness distribution method resulted in an RMSE of 0.76 mm for femur and 0.77 for tibia. The application of PSM to the cartilage thickness estimation in surgery ROI by providing the bone gap information on surgery related ROI resulted in an average RMSE from 1.05 mm to 0.80 mm for femur and from 0.79 mm to 0.62 mm to for tibia, respectively. The prediction of cartilage thickness based on bone coordinates resulted in an average RMSE from 0.78 mm to 0.98 mm for femur and from 0.96 mm to 0.82 mm to for tibia. The prediction of cartilage coordinates based on bone coordinates resulted in an average RMSE from 0.96 mm to 0.90

mm for femur and from 0.92 mm to 0.78 mm to for tibia. The estimation of cartilage thickness based on output of cartilage coordinates outperformed the directed estimated output of the cartilage thickness on femur experiment, but the result was similar in the tibia experiment. The mode of cartilage coordinates output achieved a faster convergency. The results of Van Dijck's method appeared the overfitting and achieves the optimal point around mode 20 - 30.

Furthermore, additional statistics are provided in Table 4.1 for the best modes of each method. The proposed method with bong gap information achieved the best results on statistics of mean, standard deviation, and 5% percentiles and 95% percentiles in most of the cases.

Table 4.1: Summary of the performance of cartilage thickness estimation for femur and tibia in the optimal number of modes according to different methods.

	Bone	**Method**	**Mode**	**Mean(mm)**	**Standard deviation (mm)**	**5%(mm)**	**95%(mm)**
Tibiofemoral contact ROI	Femur	Average	-	0.70	0.26	0.40	1.19
		Van Dijck	30	0.72	0.24	0.39	1.15
		Proposed	90	0.64	0.18	0.36	0.95
	Tibia	Average	-	0.77	0.32	0.33	1.27
		Van Dijck	20	0.79	0.29	0.41	1.38
		PSM	60	0.58	0.21	0.29	1.00
Surgery related ROI	Femur	Average	-	0.76	0.21	0.50	1.09
		Van Dijck	30	0.82	0.22	0.53	1.27
		Proposed	80	0.79	0.18	0.50	1.12
	Tibia	Average	-	0.77	0.28	0.41	1.24
		Van Dijck	20	0.79	0.26	0.48	1.33
		Proposed	60	0.61	0.18	0.37	0.95

4.5 DISCUSSION

In biomechanical analysis of the knee joint movement, the cartilage information is important for the modelling. However, the cartilage information is not always accessible in medical images, such as the CT images or pathological cartilage images. In this study, we aimed to use posterior shape model to estimate the cartilage thickness using solely the bone geometry based on a public dataset and compared our method to state-of-the-art methods.

As the compactness and generalization ability shown in Figure. 4.7, the proposed workflow effectively constructed a statistical shape model that could represent the mean shape and variation of the bone shape model and cartilage shape model with a public SKI10 dataset, which was in agreement with the statistical shape model introduced in [1, 7].

We found that the baseline accuracy of the mean cartilage thickness distribution method was lower than the previous study (Figure 4.8-4.10) [1]. The reason might be the statistics of the cartilage thickness estimation in this study were based on the surgery related ROI and the tibiofemoral contact ROI, which was in agreement with the finding in [1] that the accuracy of cartilage thickness estimation at anatomical knee joint contact area were relatively lower than the edge part of the knee cartilage.

Moreover, we showed that using part of real cartilage thickness to estimate the unseen part of cartilage thickness with the posterior shape model outperformed the mean thickness distribution method in the order of 0.64 mm vs 0.77 mm for femur and 0.47 mm vs 0.74 mm for tibia. The RMSE was smaller compared with reported inter-observer MR segmentation variability of 0.77 and 0.62 mm for femur and tibia respectively [8]. Hence, the first hypothesis was supported by the results that the cartilage structure across the articular surface is correlated and could be predictable using PSM, which was in agreement with the previous study [2] that the characteristic patterns of cartilage structure were found in different part of cartilage. From a clinical perspective, the proposed method might be able to provide meaningful suggestion of the cartilage thickness in the case of defected cartilage by providing the thickness from the healthy part, which is benefit to the surgery planning or biomechanical model building.

Furthermore, the average RMSE illustrated in Figure 4.9 reflected the bone gap information could be used to predict the cartilage information at tibiofemoral contact ROI within enough accuracy (0.64 mm for femur and 0.59 mm for tibia). Also, the standard deviation of proposed method within the bone gap information in Table 1 showed its robustness over other combinations. However, as Figure 4.10 illustrates, the predicted ability from bone gap information at the tibiofemoral contact ROI to the surgery related ROI was

not as good as the prediction only at contact ROI. Hence, it is suggested to add the bone gap information between femur and patella to increase the accuracy of the prediction.

We also compared the relationship between bone shape coordinates and the cartilage information. As Figure 4.9 and Figure 4.10 illustrate, the results in this study were not as good as Van Dijck's results. The lower performance might be caused by the amount of training samples, where there were 522 samples in Van Dijck's experiments. Another possible reason is the training dataset included the patients with different gender and different KL scores of the knee joint. Although Van Dijck et al. claimed that there was no bias between the gender groups [1], several previous studies reported the gender difference in knee joint cartilage thickness, volume and articular surface [14, 15]. Plus, Favre et al. showed the characteristic patterns of cartilage distribution in the patient group with different KL score [2]. Hence, we suggest constructing the PSM based on a gender separated group with different KL score in the future researches.

Although this study suggested a promising result for cartilage estimation using solely bone geometry, there are several limitations in this study: (1) the amount of the samples to construct the posterior shape model was less than previous reported study [1], however, for a proof-of-concept study, we think it is sufficient for this limited conceptual purpose; (2) the cartilage edge information was not accurate in this public dataset, which limited to the proposed experiments were only focus on the surgery related ROI and the tibiofemoral contact ROI. However, the surgery related ROI are the most concerned region for the surgeons and clinicians, and thus, the outcome of this study are still benefit to them from a clinical perspective; (3) the patella was not included in this study based on the constraint by the public dataset, and we would consider to include the patella information in the future study.

4.6 Conclusion

In this study, we proposed to use posterior shape model to estimate the cartilage thickness solely based on bone geometry using a public dataset. We showed that the cartilage structure across the articular surface was correlated at the knee joint surgery region of interest. We also showed the bone gap information could be used to predict the cartilage thickness with enough accuracy using the posterior shape model. As a feasibility study, further researches on a larger number of healthy subjects and orthopaedic patients are still needed to assess the clinical potential of this technique.

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW). Besides, this work is supported by Google Cloud Platform (GCP) research credits program.

References

[1] C. Van Dijck, R. Wirix-Speetjens, I. Jonkers, and J. Vander Sloten, "Statistical shape model-based prediction of tibiofemoral cartilage," *Computer Methods in Biomechanics and Biomedical Engineering,* vol. 21, pp. 568-578, 2018, doi: 10.1080/10255842.2018.1495711.

[2] J. Favre, J. C. Erhart-Hledik, K. Blazek, B. Fasel, G. E. Gold, and T. P. Andriacchi, "Anatomically Standardized Maps Reveal Distinct Patterns of Cartilage Thickness With Increasing Severity of Medial Compartment Knee Osteoarthritis," *Journal of Orthopaedic Research,* vol. 35, pp. 2442-2451, 2017, doi: 10.1002/jor.23548.

[3] 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, pp. 737-744, 2008, doi: 10.1109/TMI.2007.907323.

[4] M. P. Hellio Le Graverand *et al.*, "Subregional femorotibial cartilage morphology in women - comparison between healthy controls and participants with different grades of radiographic knee osteoarthritis," *Osteoarthritis and Cartilage,* vol. 17, pp. 1177-1185, 2009, doi: 10.1016/j.joca.2009.03.008.

[5] S. Koo and T. P. Andriacchi, "A comparison of the influence of global functional loads vs. local contact anatomy on articular cartilage thickness at the knee," *Journal of Biomechanics,* vol. 40, pp. 2961-2966, 2007, doi: 10.1016/j.jbiomech.2007.02.005.

[6] J. M. Johnson, "Analysis , Segmentation and Prediction of Knee Cartilage using Statistical Shape Models," ed, 2013.

[7] T. Albrecht, M. Lüthi, T. Gerig, and T. Vetter, "Posterior shape models," *Medical Image Analysis,* vol. 17, pp. 959-973, 2013, doi: 10.1016/j.media.2013.05.010.

[8] T. Heimann and B. Morrison, "Segmentation of knee images: A grand challenge," 2010.

[9] A. Myronenko and X. Song, "Point Set Registration: Coherent Point Drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 2262-2275, 2010, doi: 10.1109/TPAMI.2010.46.

[10] M. Meyer, P. Schr, and A. H. Barr. Implicit Faring of Irregular Meshes using Diffusion and Curvature Flow.

[11] T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis,* vol. 13, pp. 543-563, 2009, doi: 10.1016/j.media.2009.05.004.

[12] G. Vincent, C. Wolstenholme, I. Scott, and M. Bowes. Fully Automatic Segmentation of the Knee Joint using Active Appearance Models.

[13] C. E. Rasmussen and C. K. I. Williams, "Gaussian processes for machine learning," *International Journal of Neural Systems,* 2006.

[14] C. Ding, F. Cicuttini, F. Scott, M. Glisson, and G. Jones, "Sex differences in knee cartilage volume in adults: Role of body and bone size, age and physical activity," *Rheumatology,* vol. 42, pp. 1317-1323, 2003, doi: 10.1093/rheumatology/keg374.

[15] S. C. Faber *et al.*, "Gender differences in knee joint cartilage thickness, volume and articular surface areas: Assessment with quantitative three-dimensional MR imaging," *Skeletal Radiology,* vol. 30, pp. 144-150, 2001, doi: 10.1007/s002560000320.

Chapter 5 Axial orientation in CT scanner is of large influence on tibial tubercle-trochlear groove distance

Abstract

Introduction: The tibial tubercle - trochlear groove (TT-TG) distance is a clinical measure for lateralization of the tibial tubercle. The standard healthy range has been subject to debate, and the causes of this variation have not yet been fully substantiated. The aim of this study is to quantify the misalignment in axial orientation of the knee in CT and its effect on TT-TG measurement for patients with patellofemoral instability.

Methods: Static axial CT scans of both knees in extended position were acquired from eight patients with patellofemoral instability. The axial orientation angle between mechanical axis and craniocaudal axis of the CT scanner was determined in each patient and varied in computer simulation from 1° to 7° leg adduction and abduction. Effect on TT-TG distance was determined in manual and automated fashion. Finally, TT-TG results were corrected for misalignment and then re-compared.

Results: The average misalignment between the leg axis and craniocaudal axis of the CT scanner based on patient data was 1.4° with a standard deviation of 4.8°. Intra and inter observer variability in TT-TG distance were 1.2 ± 1.0 mm and 2.3 ± 1.8 mm. A mean increase of 1.0 ± 0.2 mm on TT-TG distance per degree of abduction/adduction was found. The correction method reduced TT-TG variation to -0.24 ± 0.24 mm.

Conclusions: The TT-TG distance is sensitive to small variations in axial slice direction of CT with around 1 mm per degree of misalignment. Clinicians should be aware of the important influence of axial slice direction in relation to the TT-TG distance.

5.1 Introduction

Patellar instability is one of the most frequent types of knee complaints in adolescents and young adults. Recurrent patellar instability has a multifactorial origin, including trochlear dysplasia, medial patellofemoral ligament (MPFL) insufficiency, patella alta and a lateral position of the tibial tubercle [1]. In order to determine the mediolateral position of the tibial tubercle, the measurement of the tibial tubercle – trochlear groove (TT-TG) distance is used in clinical practices [2].

The TT-TG distance is an absolute measure and a distance of 20 mm is clinically accepted as the threshold to consider medializing tibial tubercle osteotomy [3]. However, the healthy range of TT-TG distance is still under debate [4-8]. Pennock et al. [4] attempted to analyse the variation of TT-TG distance based on age, sex, height and patellar instability, and found each centimetre increment in height accompanies with 0.12 mm increment in TT-TG distance. Tensho et al. [7] and Izadpanah et al. [8] reported that the TT-TG distance was affected more by knee rotation than tubercle malposition. Tanaka et al. also showed a mean TT-TG distance at 5° of knee flexion being 1.5 times greater than at 30° of knee flexion [9].

Despite the widely reported variation of TT-TG distance in clinical research [4, 10, 11], only few reports mentioned the influence of the axial orientation misalignment between the leg and the CT scanner, which could easily occurs since this orientation is set by the CT operator and determined by the position of the patient in the CT scanner. Yao et al. reported how a five degree of simulated femur abduction led to an average 3.4 mm change of TT-TG distance on MRI data and proposed a correction method [12]. However, the results cannot be directly applied to CT, as a standard knee CT scan is acquired at zero knee flexion while MRI is acquired at light flexion depending on coil shape and size [11, 13]. Also, many researchers has demonstrated that the TT-TG distance cannot be used interchangeably between CT and MRI [6, 14]. Furthermore, in Yao's study, all landmarks were re-selected manually in the reconstructed slices, which may introduce error [15].

The aim of this study is therefore to quantify the misalignment in axial orientation of the knee on CT scan under standard scanning conditions and to quantitatively assess the offset of TT-TG distance caused by misalignment of the leg axis and the craniocaudal axis of the CT scan in a comprehensive manner for patients with patellofemoral instability. In this study, we hypothesized that the abduction and adduction of extended knee leads to an unneglectable change in TT-TG distance in CT scans and the change is primarily caused by the reference line selection on TT-TG measurement. We also hypothesized that an analytic correction model [12] (analysed in MRI scan) can be also applied to CT data to reduce the misalignment influence.

5.2 Methods and materials

5.2.1 Data description

Informed consent was obtained from all subjects in accordance with the institutional review board approved study protocol. Eight patients with knee pain complain underwent CT scans (4 men and 4 women, mean age of 23.5 (18-33) years old) of both knees in full extension, i.e. maximal extension and to have the knee in a locked position. The obtained volume had a fixed matrix size of $512 \times 512 \times 626$. The in-plane pixel spacing varied from 0.511×0.511 mm to 0.824×0.824 mm. Slice thickness was 1 mm with an increment of 0.8 mm. The bone kernel filter was applied for reconstruction.

Figure 5.1: (a) CT image in coronal view with the target to align the patient perfectly within the CT scanner. The black line indicates the anatomical plane of the knee joint, showing a misalignment of 3° compared to the CT axial plane, the black dash line indicates the leg axis / mechanical axis / tibia proximodistal axis; (b) the CT scanner with corresponding coordinate frame (x, y and z)

5.2.2 Introduction of related coordinates

The mechanical axis is defined as the line between the centre of the femur head and centre of the ankle joint [16]. As Figure 5.1a illustrates, these two joints are not in the field of view (FOV) of the data set used in this study, so we approximated the mechanical axis with the anatomical proximodistal axis of the tibia [16, 17], which is highly comparable to the mechanical axis of the full leg in zero flexion. As Figure 5.1b illustrates, the mechanical axis has the same orientation as the z-axis of the scanner. A positive angle between the tibia proximodistal axis and z-axis of scanner is defined as the angle of abduction.

5.2.3 Measurement of TT-TG distance

As depicted in Figure 5.2, all manual measurements were performed on an independent workstation using Lyon's protocol (Dejour et al., Chapter 6) [1]. TT-TG distance was determined in the right and left legs of 8 patients, i.e. 16 legs by determining three geometrical measures: the TT point, TG point and posterior condylar reference line (PCRL). The trochlear groove (TG) and PCRL are measured in the same axial slice, which includes the most proximal part of the TG and slight subchondral condensation seen in the trochlea [1]. The TG point is defined as the deepest part of the trochlear groove. The PCRL is defined as the tangent line to both posterior condyles. The tibial tubercle (TT) point is defined in the slice existing the most anterior part of tibial tubercle. The measurement of TT-TG distance for each patient was performed three times (one-week interval) by an experienced radiologist for intra-observer variability. Two experienced radiologists and one experienced orthopaedic surgeon were included for an inter-observation measurement.

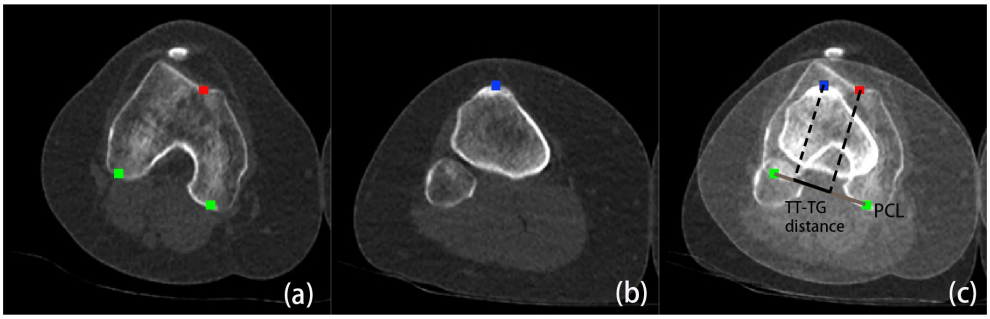


Figure 5.2: Manual measurement of TT-TG distance within the CT image ((a) red box indicates the TG, green boxes indicate the landmarks of PCRL; (b) blue box indicates the TT; (c) overlap of (a) and (b), and black line is the TT-TG distance).

5.2.4 Measurement of deviation between leg axis and z-axis of CT scanner

The optimal orientation of the axial slices of CT image, in which the TT-TG distance can be determined, should be 90° to the mechanical axis. The angle of deviation between the anatomical axial plane of the leg and the axial plane of the CT scanner frame is defined as the angle between the mechanical axis and z-axis of the CT scanner frame. Misalignment of the anatomical axial plane of the leg with respect to the CT scanner frame is thus defined in this paper as the angle between the proximodistal axis of the tibia and the z-axis of the CT scanner frame. Misalignment of the leg was determined in all original data sets i.e. for 16 legs.

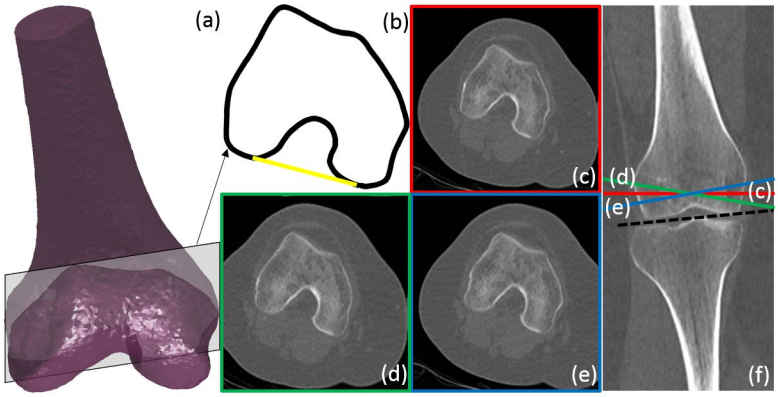


Figure 5.3: (a) Mesh representation of patient 2 (purple mesh indicates the femur, light grey plane indicates the cut plane through the most proximal part of the trochlear groove); (b) black curve represents the bone contour at slice location indicated with grey cut plane in (a), yellow line indicates the reference posterior condylar reference line; (c-e) digitally simulated abduction to adduction of the leg from -7° to 0° to 7° with respect to the CT scanner frame. Artificial deviation angles were created by changing the slice direction of the images based on the y-axis of CT scanner and re-interpolating; (f) the coronal view of the CT image, the red, green and blue line indicate the corresponding plane of (c-e), and black dash line represents the anatomical plane of the knee joint based on the anatomical proximodistal/mechanical axis of the tibia.

5.2.5 Synthesizing data with varying adduction-abduction

Previously, Yao et al. reported a misalignment of the anatomic axial plane from routine axial plane on MRI scanner was 5°±2.3° abduction [12]. In this study, we simulated the angle of abduction and adduction up to 7° based on CT data. As Figure 5.3c–e show, varying abduction and adduction was simulated for all data sets in a range of abducted 7° to adducted 7° in steps of 1° by gradually changing the slice direction of the images based on the anterior-posterior axis of tibia and re-interpolating for each incremental slice direction. The location of the TT and TG point were kept constant, i.e. the manually selected locations in the original data were used for all synthesized data sets. For each angle of abduction/adduction the PCRL was now detected based on an automatic algorithm, as described in Chapter 6 (Figure 6.3a-b).

5.2.6 Outcome parameters

Firstly, to assess the misalignment between transverse plane of the knee and the x-y plane of CT scanner in standard scanning conditions, we calculated the angle difference between the z-axis of CT scanner and tibia proximodistal axis of the leg. Secondly, to guarantee the robustness of automatic calculated PCRL, inter-observation and intra-observation of the PCRL were compared with the automatic method. Thirdly, the TT-TG distance and corrected

TT-TG distance based on Yao's method [12] was re-quantified for all synthesized data sets and plotted as a function of leg abduction/adduction in degrees. Fourthly, an averaged offset of TT-TG distance and corrected TT-TG distance based on Yao's method [12] over all patients as a function of adduction to abduction angle was calculated.

5.3 Results

Table 5.1 shows the misalignment between tibia proximodistal axis and z-axis of the CT scanner for all original data sets. The average deviation was 1.4° (range: -5.5° – 12.6°; SD=4.8°). The highest mismatch (12.6°, subject 5) was found in a patient with a serious valgus deformity of the leg.

Table 5.1: Misalignment of the leg and z-axis of the CT scanner (positive indicate abduction)

	Subject1	**Subject2**	**Subject3**	**Subject4**	**Subject5**	**Subject6**	**Subject7**	**Subject8**
Left	-2.1	-3.8	2.9	-1.6	5.5	7.9	-2.9	-1.4
Right	-1.6	-5.5	1.9	2.2	12.6	6.3	3.3	-0.5

As Figure 5.4 shows, the intra-observer variation of TT-TG distance in original data sets was associated with a mean difference of 1.2 mm (SD = 1.0), while the inter-observer variation was 2.3 mm (SD=1.8 mm). However, the difference of TT-TG distance between manual measured PCRL and automatic detected PCRL was 0.1 mm (SD=0.2 mm). The TT-TG distance based on the automatic detected PCRL was highly consistent to the manual detected one.

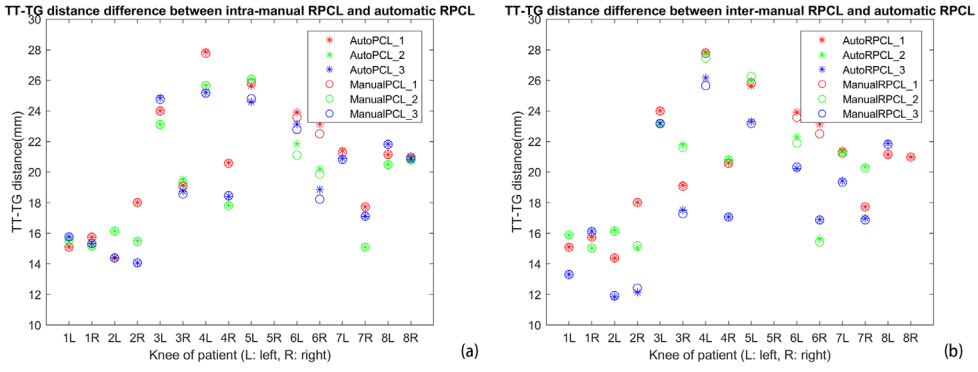


Figure 5.4: The comparison of the TT-TG distance among the intra-measured PCRL (a) / inter-measured PCRL (b) and the automatic detected PCRL (Chapter 6) for both legs of eight

patients (star and circle indicate the TT-TG distance based on automatic and manual detected PCRL, respectively).

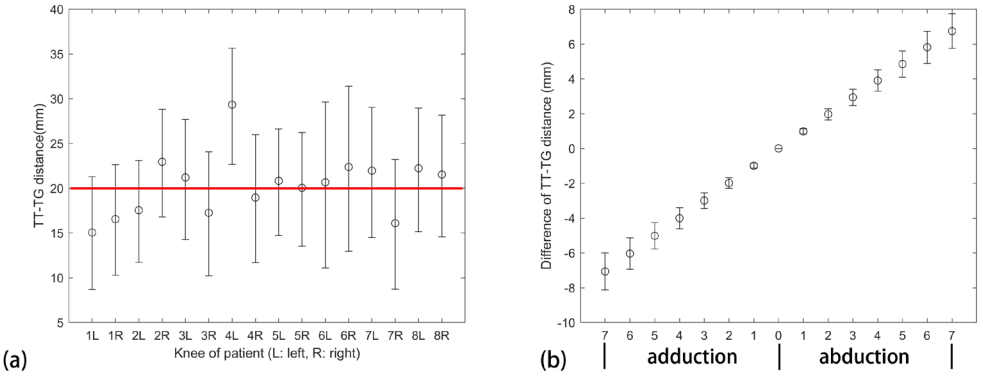


Figure 5.5: TT-TG distance range from adduction 7° to abduction 7° (red line indicates clinical standard of 20 mm for considering surgical treatment); (a) circle indicates the TT-TG distance on optimal CT axial scan (i.e. anatomical axial plane of tibia), and the error bar represent the range of measured TT-TG distance among the simulated data; (b) averaged TT-TG offset over all patients as a function of adduction to abduction angle.

Figure 5.5a illustrates the variation of averaged TT-TG distance due to simulated abduction/adduction of -7 to 7 degrees per leg. The red line in Figure 5.5a indicates the surgical threshold criterion in patients with knee pain complaints. Fifteen out of sixteen datasets have TT-TG distance's above and beneath this critical distance with the imposed abduction/adduction misalignment. Figure 5.5b illustrates averaged TT-TG offset over all patients as a function of adduction to abduction angle. The average offset in TT-TG distance as function of the abduction/adduction angle was calculated at 1.0 ± 0.2 mm per degree of abduction/adduction. Figure 5.6 shows the variation of averaged corrected TT-TG distance based on Yao's method [12] due to simulated abduction/adduction of -7 to 7 degrees per leg and the difference of TT-TG distance is reduced to -0.24 ± 0.24 mm.

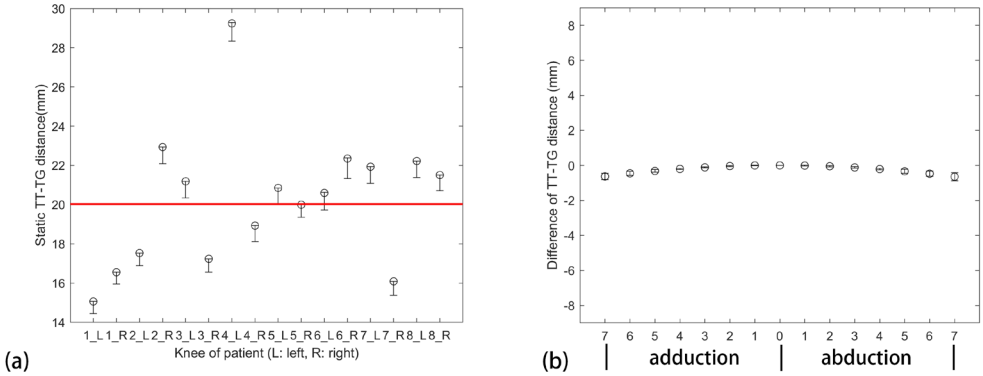


Figure 5.6: Corrected TT-TG distance range from adduction 7° to abduction 7° (red line indicates clinical standard of 20 mm for considering surgical treatment); (a) circle indicates the TT-TG distance on optimal CT axial scan (i.e. anatomical axial plane of tibia), and the error bar represent the range of measured TT-TG distance among the simulated data; (b) averaged TT-TG offset over all patients as a function of adduction to abduction angle.

5.4 Discussion

In this study, we showed that the difference in TT-TG distance due to a slightly adducted or abducted leg was large and a correction method to the measurement was needed. Clinically, the axial CT image are acquired perpendicular to the longitudinal axis of the scanner by the scanning operator. Unlike with MRI imaging, the reconstruction direction of CT images cannot be tuned before imaging and depends on the position and orientation of the patient in the scanner. Although the image volume can be interpolated to adjust the axial plane of the knee within post-processing of the image, this is seldomly performed in real practice. Our study showed a mismatch between the optimal orientation of the leg and the scanned orientation being about 1.44° (range: -5.5° to +12.6°; SD=4.8°). This mismatch is hardly judged by the radiologist, unless the full leg is scanned to confirm the mechanical axis. For that matter, an automatic determination of anatomical coordinate frame is preferred in the measurement.

In this study, we used an automatic detected PCRL instead of the manual detection. Although some subjects showed varied TT-TG distance between use of manual and automatic detected PCRL, it was relatively small compared to the variability of inter- / intra-observed TT-TG distance. The variation in TT-TG distance between manual and automatic PCRL detection is caused by variation in landmarks selection of the PCRL. The larger inter- and intra-observer variability in some patients can also be explained by the fact that they have trochlea dysplasia which can make it difficult to determine the deepest point of the trochlea groove.

Our study confirmed the hypothesis that minor changes in leg orientation had a clinically significant influence on the TT-TG distance in a CT scan. In this study, an average 1.0 ± 0.2 mm increment of TT-TG distance per degree of adduction/ abduction was found. The estimation of the difference on the TT-TG distance due to misalignment between the anatomical axial plane of knee and axial plane has also been assessed for MRI acquisition. Yao et al. reported how 5 degree of simulated femur abduction leads to average of 3.4 mm change in TT-TG distance [12] in MRI images. The trend of the variation from abduction to adduction was the same in this study. However, the average variation of the TT-TG distance was different. We supposed the difference was caused by three reasons: 1) the TT-TG distance are not equivalent between CT and MRI, as the reported studies show a 2.2 ± 3.8

mm underestimation in MRI as compared to measurements based on CT [6] ; 2) Yao et al. estimated the TT-TG distance in an interpolated 2D MRI image, and a thickness factor was included in the measured landmark for TT-TG distance calculation; 3) the small variation in location of TT and TG point in each measurement (abduction 5°, 0°, adduction 5°). In this study, we kept the TT and TG fixed and analysed the influence of axial orientation by the single variable of change of the PCRL.

The TT-TG distance has been reported as an important measure in determining the appropriate surgery for patients with patellofemoral instability [1, 12, 18]. In previous studies, an average of 13 mm for the TT-TG distance was found on subjects without history of patellar instability, while an average of more than 20 mm of TT-TG distance has been found in patients with patellar instability [19]. In this sense, a variance of 7 mm or less of TT-TG distance may be clinically significant and affect the surgical decision. Hence, the potential system error of minor leg abduction or adduction should be considered when determining and interpreting the TT-TG distance in a CT image. In this study, after applying the correction method proposed by Yao et al.[12], the difference of TT-TG distance due to the misalignment was reduced to -0.24 ± 0.24 mm, which showed the correction method can be also applied to CT scan.

There are several limitations to this study. For ethical consideration we did not perform multiple scans on each patient but simulated the misalignment. However, we think the difference of the results between these two procedures are minor, but merely allowed us to assess a sufficient range of misalignment angles systematically without exposing the patients to additional radiation. Another limitation is that we used a standard clinical scanning protocol that led to the misalignments as reported in this study. One could think of additional alignment features to reduce the errors in misalignment, but these were not utilized in this study as we wanted to remain close to clinical practice. Also, the sample of the patients in this study were eight. In fact, both knee of the patient were analysed separately, so the sample of target in this study was 16. Besides, this study is designed to test a potential system error of TT-TG distance in CT scan and suggested a correction method on CT data. Hence, the small study group is sufficient for this limited purpose. Furthermore, we only used the patient data with patellofemoral instability. Hence, the outcome of misalignment of bone axis and the craniocaudal axis of CT scanner may not directly apply to the healthy subject. However, this is unlikely since we also scanned patients with “normal” TT-TG values.

5.5 Conclusions

In summary, this study showed that a minor misalignment of the mechanical axis of the lower extremity and the craniocaudal axis of the CT scanner could cause a significant change

of TT-TG distance measurement from CT images. A systematic assessment of this misalignment between the mechanical axis and CT axis from abduction (-7°) to adduction (7°) showed an average offset in TT-TG distance of 1.0 ± 0.2 mm per degree. In clinical practice, the mechanical axis of the leg was not always aligned to the longitudinal axis of the scanner, and thus, the TT-TG distance measurement would easily include a systematic error, which might be one of the reasons for the different healthy ranges of TT-TG distances published in previous studies. For a more robust and reproducible TT-TG distance, the offset in the obtained TT-TG could be corrected once the abduction/adduction is known or may be included in an adjustment of the axial plane before a TT-TG distance measurement is performed.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW). Also, we greatly thank Léon Driessen for the bone shape models and Marieke Ploegmakers for the manual assessments of TT-TG distance.

References

[1] W. N. Scott, "Surgery of the Knee," *Insall & Scott Surgery of the Knee,* pp. 1-1504, 2014, doi: 9781455727896.

[2] P. B. Schoettle, M. Zanetti, B. Seifert, C. W. A. Pfirrmann, S. F. Fucentese, and J. Romero, "The tibial tuberosity-trochlear groove distance; a comparative study between CT and MRI scanning," *Knee,* vol. 13, pp. 26-31, 2006, doi: 10.1016/j.knee.2005.06.003.

[3] A. C. Colvin and R. V. West, "Patellar Instability," *The Journal of Bone and Joint Surgery-American Volume,* vol. 90, pp. 2751-2762, 2008, doi: 10.2106/JBJS.H.00211.

[4] A. T. Pennock, M. Alam, and T. Bastrom, "Variation in tibial tubercle-trochlear groove measurement as a function of age, sex, size, and patellar instability," *American Journal of Sports Medicine,* vol. 42, pp. 389-393, 2014, doi: 10.1177/0363546513509058.

[5] P. P. Koch, C. W. A. Pfirrmann, S. F. Fucentese, M. Betz, and T. J. Dietrich, "End-stage extension of the knee and its influence on tibial tuberosity-trochlear groove distance (TTTG) in asymptomatic volunteers," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 214-218, 2012, doi: 10.1007/s00167-012-2357-z.

[6] C. L. Camp *et al.*, "CT and MRI measurements of tibial tubercle-trochlear groove distances are not equivalent in patients with patellar instability," *American Journal of Sports Medicine,* vol. 41, pp. 1835-1840, 2013, doi: 10.1177/0363546513484895.

[7] K. Tensho *et al.*, "What Components Comprise the Measurement of the Tibial Tuberosity-Trochlear Groove Distance," *J Bone Joint Surg Am,* vol. 97, pp. 1441-1448, 2015, doi: http://dx.doi.org/10.2106/JBJS.N.01313.

[8] K. Izadpanah *et al.*, "Influence of knee flexion angle and weight bearing on the Tibial Tuberosity-Trochlear Groove (TTTG) distance for evaluation of patellofemoral alignment," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2655-2661, 2014, doi: 10.1007/s00167-013-2537-5.

[9] M. J. Tanaka, J. J. Elias, A. A. Williams, J. A. Carrino, and A. J. Cosgarea, "Correlation Between Changes in Tibial Tuberosity-Trochlear Groove Distance and Patellar Position During Active Knee Extension on Dynamic Kinematic Computed Tomographic Imaging," *Arthroscopy : the journal of arthroscopic & related surgery,* vol. 31, pp. 1748-1755, 2015, doi: 10.1016/j.arthro.2015.03.015.

[10] J. H. Sojka, J. S. Everhart, J. C. Kirven, M. D. Beal, and D. C. Flanigan, "Variation in tibial tuberosity lateralization and distance from the tibiofemoral joint line: An anatomic study," *Knee,* 2018, doi: 10.1016/j.knee.2018.03.006.

[11] A. Marquez-Lara, J. Andersen, L. Lenchik, C. M. Ferguson, and P. Gupta, "Variability in patellofemoral alignment measurements on MRI: Influence of knee position," *American Journal of Roentgenology,* vol. 208, pp. 1097-1102, 2017, doi: 10.2214/AJR.16.17007.

[12] L. Yao, N. Gai, and R. D. Boutin, "Axial scan orientation and the tibial tubercle-trochlear groove distance: Error analysis and correction," *American Journal of Roentgenology,* vol. 202, pp. 1291-1296, 2014, doi: 10.2214/AJR.13.11488.

[13] B. B. Hinckel *et al.*, "Are the osseous and tendinous-cartilaginous tibial tuberosity-trochlear groove distances the same on CT and MRI?," *Skeletal Radiology,* vol. 44, pp. 1085-1093, 2015, doi: 10.1007/s00256-015-2118-4.

[14] C. M. Anley, G. V. Morris, A. Saithna, S. L. James, and M. Snow, "Defining the Role of the Tibial Tubercle-Trochlear Groove and Tibial Tubercle-Posterior Cruciate Ligament Distances in the Work-up of Patients With Patellofemoral Disorders," *American Journal of Sports Medicine,* vol. 43, pp. 1348-1353, 2015, doi: 10.1177/0363546515576128.

[15] S. K. Aoki, J. J. Wilcox, R. T. Burks, B. J. Snow, and M. Hung, "Does Landmark Selection Affect the Reliability of Tibial Tubercle–Trochlear Groove Measurements Using MRI?," *Clinical Orthopaedics and Related Research®,* vol. 470, pp. 2253-2260, 2012, doi: 10.1007/s11999-012-2269-8.

[16] J. J. Cherian, B. H. Kapadia, S. Banerjee, J. J. Jauregui, K. Issa, and M. A. Mont, "Mechanical, anatomical, and kinematic axis in TKA: Concepts and practical applications," *Current Reviews in Musculoskeletal Medicine,* vol. 7, pp. 89-95, 2014, doi: 10.1007/s12178-014-9218-y.

[17] D. L. Miranda, M. J. Rainbow, E. L. Leventhal, J. J. Crisco, and B. C. Fleming, "Automatic determination of anatomical coordinate systems for three-dimensional bone models of the isolated human knee," *Journal of Biomechanics,* vol. 43, pp. 1623-1626, 2010, doi: 10.1016/j.jbiomech.2010.01.036.

[18] T. O. Smith, L. Davies, A. P. Toms, C. B. Hing, and S. T. Donell, "The reliability and validity of radiological assessment for patellar instability. A systematic review

and meta-analysis," *Skeletal Radiology,* vol. 40, pp. 399-414, 2011, doi: 10.1007/s00256-010-0961-x.

[19] H. Dejour, G. Walch, L. Nove-Josserand, and C. Guier, "Factors of patellar instability: An anatomic radiographic study," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 2, pp. 19-26, 1994, doi: 10.1007/BF01552649.

Chapter 6 A robust and semi-automatic quantitative measurement of patellofemoral instability based on four dimensional computed tomography

Abstract:

Introduction: Patellofemoral instability is a motion related disease, featured as the patella dislocating from the trochlear groove. Four dimensional computed tomography (4DCT) enables full assessment of the patellofemoral movement. Nevertheless, the quantitative measurements of patellofemoral instability are still under research and currently of limited practical use. The aim of this study is to develop a robust and semi-automatic workflow to quantitatively describe the patellofemoral movement.

Methods: Eight patients suffering from patellofemoral instability were scanned in 3D/4D CT. For each patient, bone shape models were firstly constructed based on the segmented mask, after which a subject-specific anatomical coordinate frames and the bone landmarks for clinical measures were calculated based on the knee bone shape models. Coherent point drift registration was used to obtain the transformation matrices among the bone shape models to describe the knee joint deformation. The clinical measures of dynamic tibial tubercle – trochlear groove (TT-TG) and patellar centre – trochlear groove (PC-TG) distance were calculated for each patient based on the anatomical coordinate frames, bone landmarks and the transformation matrices.

Results: Using our method, we were able to obtain the dynamic TT-TG and PC-TG distance in automatic manner providing the bone shape models. The TT-TG distance based on proposed method successfully discriminate patients with patellofemoral instability. The obtained dynamic TT-TG distance shows a decreasing trend as the knee flexion angle increases, and the dynamic TT-TG distance generally follows a slightly different track either going into flexion or going into extension, indicating the clinical value compared with the conventional clinical used TT-TG distance.

Conclusions: The initial results showed agreement with manual observations of the tibial tubercle - trochlear groove (TT-TG) distance in routine practice, and the possibility to evaluate both TT-TG distance and patellar centre - trochlear groove (PC-TG) distance dynamically during active flexion-extension-flexion movement of the knee with an automatic workflow.

6.1 Introduction

Patellofemoral instability, presented as a dislocation of the patella from the trochlear groove, is related to knee pain and dysfunction. The majority of patients with patellofemoral pain syndrome can be treated non-operatively [1, 2], and the criterion of surgery on patellofemoral disorders has been subject to debate [3-9]. To assist in this debate there is a need for an accurate quantitative clinical description of patellofemoral instability.

Current clinical protocols of 3D knee imaging, e.g. CT, MRI, provide high quality images of the knee anatomy in terms of resolution, contrast and signal to noise ratio (SNR). However, an increasing number of studies have shown the benefit of having dynamic information of the knee over just one static image, resulting in improved diagnosis and more support in surgical decision making [10-13]. To provide motion related clinical parameters, a few research groups presented measurements based on several static scans obtained at different knee flexion angles on MRI image [10, 14] or the combination of fluoroscopic imaging and a 3D static scan [15, 16]. Nevertheless, these quasi-static approaches is of limited practical use, since it is hard to incorporate muscle factors which may affect patellofemoral functioning [12].

Recently, four dimensional computed tomography (4DCT, also known as dynamic CT) was introduced as a technological innovation in the field of CT. 4DCT is a three dimensional computed tomography volume imaged over a period of time [17], which provides the option of capturing continuous patellofemoral motion. With 4DCT, Tanaka et al. found a decrease in tibial tubercle – trochlear groove (TT-TG) distance (Figure 6.1) with increasing flexing angle of the knee joint, together with similar trends in bisect offset and patellar tilt [5], indicating the sensitivity of TT-TG to flexion angle and demonstrating that dynamic information may be beneficial to the diagnosis [11].

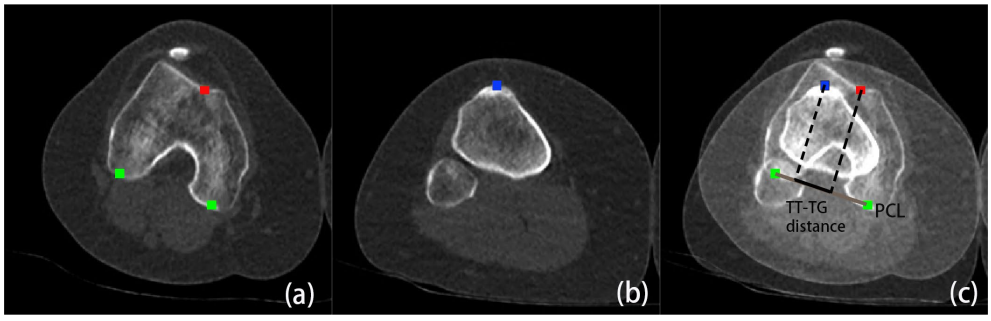


Figure 6.1: Manual measurement of TT-TG distance within the CT image ((a) red box indicates the TG, green boxes indicate the landmarks of PCRL; (b) blue box indicates the TT; (c) overlap of (a) and (b), and black line is the TT-TG distance)

In current clinical practice, healthy TT-TG distances have been reported to range from 9.4 to 13.6 mm [5]. This variation partially originates from the selection of slice direction, flexion angle of the knee and variation in landmark selection [5, 18]. To reduce the variation, the anatomical axial plane of the knee joint could be used to determine the landmarks (TG, TT, patellar centre (PC), posterior condylar reference line (PCRL) etc.) [13, 18, 19]. Moreover, the limited field of view (FOV) in 4DCT prevented any assessment over 30 degree flexion [5], in that the landmarks moved out of the FOV. Furthermore, the variance of the intra/inter-observation on landmark determination was a key factor in the variance of quantitative assessment of the knee movement [5, 12].

In this study, we aim to develop and evaluate a workflow for quantitative analysis (dynamic TT-TG distance and PC-TG distance) of patellofemoral dynamics during continuous active flexion and extension of the knee within a range between 0° and 90° on 4DCT images. The main contributions of this study are:

(1) automatic landmark determination of the trochlear groove, tibial tubercle, patella centre, posterior condyles reference line, and flexion angle between femur and tibia based on an anatomical coordinate frame instead of CT image coordinate frame;

(2) continuous measurement of TT-TG and PC-TG distance using 4DCT scanning;

(3) a robust workflow to perform the quantitative measurement within bone shape models;

(4) extending the quantitative knee joint movement measurement to 0° - 90°.

6.2 Materials and methods

The major abbreviations used in this study are summarized in Table 6.1.

Table 6.1: Main abbreviations

Abbreviation	Explanation
TT	tibial tubercle
TG	trochlear groove
PC	patellar centre
PCRL	posterior condylar reference line
4DCT	four dimension computed tomography

6.2.1 Data acquisition

Informed consent was obtained from eight patients, who would have otherwise gotten a conventional static CT, in accordance with the institutional review board–approved study protocol. These eight patients (4 men and 4 women, mean age of 23.5 (18-33)) underwent 3D and 4D CT scans with low dose on a Canon Aquilion One Genesis scanner. As Figure 6.2a illustrates, one 3D CT scan on both knees in full extension was acquired to provide anatomical information. The obtained volume had a fixed matrix size of $512 \times 512 \times 626$ voxels. The in-plane pixel spacing varied from 0.511×0.511 mm to 0.824×0.824 mm. Slice thickness was 1 mm with an increment of 0.8 mm. As Figure 6.2b and Figure 6.2c illustrate, the dynamic CT scan with eleven frames were acquired in both knees within 13 seconds during active flexion-extension-flexion movement i.e. temporal resolution of 0.85 Hz. This movement was achieved against gravitational resistance without any other limitations (see Figure 6.2). The obtained volumes had the fixed matrix size of $512 \times 512 \times 320$. The in-plane pixel spacing varied from 0.820×0.820 mm to 0.976×0.976 mm. Slice thickness was 0.5 mm with an increment of 0.5 mm. This CT protocol was associated with an ionising radiation dose of 0.32 millisievert (mSv), compared to normal preoperative CT scan of the knee of 0.02 mSv.

Figure 6.2: (a) The 3D scan with fully extension; (b) the 4D scan with approximately 90° knee flexion; (c) the 4D scan with fully extension

6.2.2 Manual measurement of static tibial tubercle - trochlear groove distance

Static TT-TG/PC-TG distance represents the distance calculated in 3D scan with full extension of the knee (0° flexion). As Figure 6.1 illustrates, to assess whether proposed method would quantify the static TT-TG distance in the same way as in clinical routine, the TT-TG distance in 3D scan for each patient was measured three times (one-week interval) by an experienced radiologist for intra-observer variability. The inter-observation measurement was performed by two experienced radiologists and one experienced orthopaedic surgeon. All manual measurements were measured using Lyon's protocol [20] (workflow to determine the TT-TG distance) and adjusted using the method from Yao et al. [18] (method to adjust the TT-TG distance based on the mis-alignment between axial plane

of the CT scanner and the anatomical axial plane of the knee). In this study, the static TT-TG distances were determined for both legs of 8 patients (16 legs).

6.2.3 Basic concept of quantitative description of patellofemoral motion

6.2.3.1 Flexion angle between femur and tibia

As Figure 6.3a illustrates, the flexion angle (α) between femur and tibia is calculated between the anatomical proximal-distal (PD) axes of femur and tibia. Full extension in 3D scan is defined as 0° flexion, negative flexion thus indicates hyperextension.

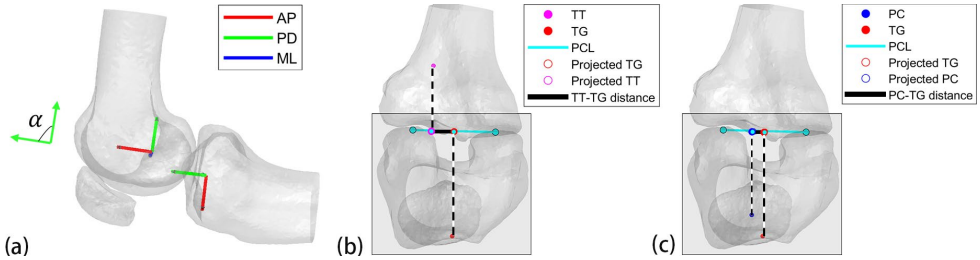


Figure 6.3: Illustration of the relevant coordinate systems and landmarks for the quantitative measurement of patellofemoral motion. (a) red, green and blue arrows indicate the corresponding anterior-posterior (AP), proximal-distal (PD) and medial-lateral (ML) directions, α indicates the flexion angle between the PD axes of femur and tibia; (b) red filled and magenta filled circles indicate the TG and TT, red and magenta circle indicate the projected TG and TT, cyan line indicates the PCRL and black thick line indicates the TT-TG distance, grey rectangle indicates the axial plane; (c) red filled and blue filled circles indicate the TG and PC, red and blue circle indicate the projected TG and PC, cyan line indicates the PCRL and black thick line indicates the PC-TG distance, grey rectangle indicates the axial plane.

6.2.3.2 Dynamic tibial tubercle - trochlear groove distance

As Figure 6.3b illustrates, like static TT-TG distance, the dynamic TT-TG distance was the distance between two projected points of trochlear groove and tibial tubercle, which was measured on a posterior condylar reference line (PCRL). The TG and PCRL were calculated in the same axial slice, with the most proximal part of the TG and slight subchondral condensation seen in the trochlea [20]. The TG point was defined as the deepest part of the trochlear groove [20]. The PCRL was defined as the tangent line to both posterior condyles [20]. The TT point was defined as the most anterior part of tibial tubercle [20]. In this study if the TT/PC point was on the lateral side of the TG point, we defined the distance as positive, otherwise negative if on the medial side. Dynamic TT-TG distance represents the TT-TG distance for each angle of flexion between femur and tibia, i.e. TT-TG distance as a function of flexion angle.

6.2.3.3 Dynamic patellar centre – trochlear groove distance

As Figure 6.3c illustrates, the dynamic PC-TG distance was the distance between the projected points of the trochlear groove and the centre of the patella which was measured on a PCRL. The definitions of TG and PCRL were the same as section 6.2.3.2. The centre of patella (PC) was identified as the centre of the patella shape model. Dynamic PC-TG distance indicates the distance as a function of flexion angle.

6.2.4 Proposed workflow for quantitative measurement of patellofemoral movement

Figure 6.4 illustrates a schematic representation of the pipeline for semi-automatic quantitative measurement of patellofemoral movement. The coordinate frames of femur and tibia and the landmarks for TT-TG/PC-TG distance are determined based on 3D scan data. The rotation matrices that describe the transformation information of the bone are determined based on 4D scan data. Manual segmentation is performed using Mimics software (V14, Materialise, Belgium). All the automatic methods are implemented using Matlab r2017a.

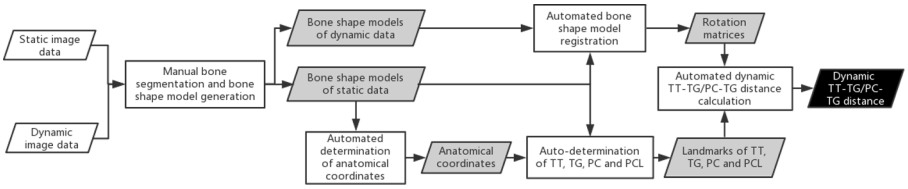


Figure 6.4: Pipeline of TT-TG distance calculation (white, grey and black parallelogram represents input data, intermediate results and output result, respectively; rectangle box represents the step of method)

6.2.4.1 Bone segmentation

Although many software tools have been published to automate the segmentation of knee bone in CT images [12], accuracy and robustness of the segmentation remains highly dependent on image quality. The dynamic CT images used in this study are of low contrast in femur condyles and contain motion artefacts, particularly in the tibia. To ensure a reliable registration of the initial bone shape obtained from the static scan, femur, tibia and patella were segmented using Mimics software, i.e. an experienced observer selected the seed points inside the structures where after a computer algorithm provided by Mimics determined the rest of the structures based on intensity similarity and filled in the holes to complete bone mask. These seed points were defined by the experienced observer by attempting to select the points of cortical bone based on intensity between 1800 to 2000 Hounsfield unit. Any leakage due to the low contrast or motion artefacts was removed manually, i.e. an

experienced observer went through the slices and removed the mis-segmented structure. The outputs from the software were the shape models of the femur, tibia and patella in STL format.

Table 6.2: The definition of the geometry threshold for articulate surface.

#	Rationale / main assumption	Definition
1 (Direction)	The articulating surface is cylindrical around the ML axis	The absolute of the cosine of the angle between the vertex normal and axis of the cylinder (i.e. the cylinder fitted to both condyles in Figure 6.5c) should be smaller than 0.6 i.e. in range of [52,127]
2 (Posterior/distal)	The articulating surfaces is located distally and posteriorly on the lower extremity of the femur	Any vertices closer than 0.35 times the furthest vertex from the separating plane (Figure 6.5a) was excluded.
3 (Lateral)	The articulating surfaces is located on medial and lateral side, but not in the central sagittal plane	Any vertices closer than 0.24 times the furthest vertex from the middle plane (Figure 6.5d) was excluded. The middle plane was defined as perpendicular to the cylinder axis, in the middle between the epicondyles.
4 (extrema)	The articulating surface is cylindrical around the ML axis	Any vertices closer than 0.725 times the furthest vertex from axis of the cylinder after applying the other measures was designated as non-articulating surface.

6.2.4.2 Automatic determination of anatomical coordinate frame for femur and tibia

a) Femur coordinate frame

The femur coordinate frame consists of three axes, medial-lateral (ML) axis, anterior-posterior (AP) axis, and proximal-distal (PD) axis. The ML axis aims to approximate the

flexion-extension axis based on the morphology articulating surfaces. First, the condyles were segmented based on the geometry of the femur shape model (Figure 6.5a). P1 was defined as the bottom part of the femur along the diaphyseal principle axis of inertia (inertial axes indicate the eigenvectors of the mass moment of inertia tensor; they were estimated using a flood filling algorithm that filled the shape model with the point masses (Figure 6.5b) [21]. P2 was defined as the outer rim along the anteroposterior inertial axis on the posterior side of the femur in slice direction perpendicular to the diaphyseal inertial axis, and at slice location where the variation in slice area was largest [21]. Then, a cylinder was fitted to both condyles based on the vertices of the isolated surface (Figure 6.5c). This cylinder was first order approximation of the condyles. In contrast to [21], we proposed to use two cylinders instead of one to represent the medial and lateral articulating surfaces separately. This allowed for asymmetry among the condyles [22] and restricted the fit to the mechanically relevant articulating surfaces [23]. To find these two cylinders, four geometrical criteria were defined based on the single cylinder fit from [21] to obtain the articular surfaces (black points in Figure 6.5d), as illustrated in Table 6.2. Next, as Figure 6.5e illustrates, the two cylinders were fitted to each articular surface. Then, the articular surface centres (centre of mass of the faces) were computed and projected on the respective cylinder axis. A line going through both points defined the ML axis. The origin of the femur anatomical coordinate frame was defined as the middle point of these two projected centres. The AP axis was determined as the cross product of ML axis and the long inertial axis of the femur shaft (Figure 6.5b 3rd IA). The PD axis of femur was determined as the cross product of ML axis and AP axis. This ensured an orthogonal reference frame, while keeping the diaphyseal inertial axis similar to the PD axis. Figure 6.5f illustrates all anatomical femur axes.

b) Tibia coordinate frame

The determination for tibial anatomical coordinate frame was identical to Miranda's method [21]. First, the centre of mass and inertial axes (IA) of the tibia were determined, as illustrated in Figure 6.6a. Along the 3rd IA (red line in Figure 6.6a), the cross-sectional area was calculated. At the point of maximum cross-sectional area, the tibial plateau was separated from the shaft (Figure 6.6b). As Figure 6.6c illustrates, the origin was determined as the centre of tibial plateau, and the AP, ML and PD axes, were determined by the 2nd, 3rd, and 1st inertial axis respectively.

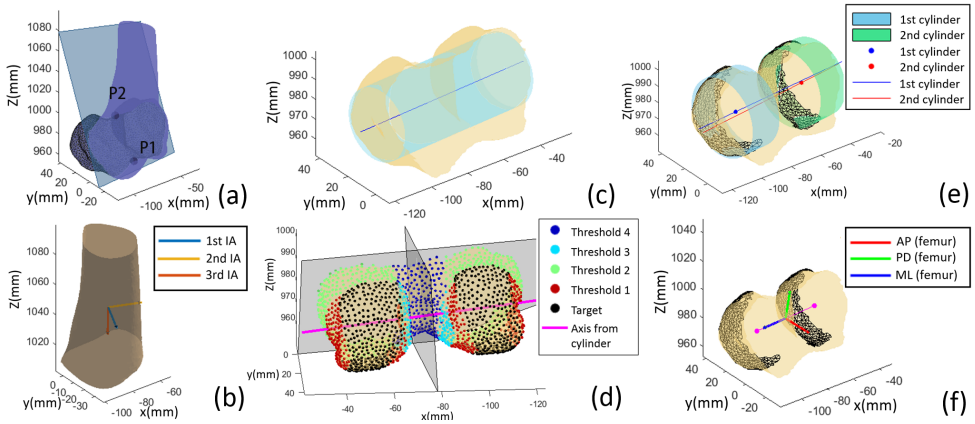


Figure 6.5: Determination of anatomical femoral coordinate. (a) separate the condyles with cut plane; (b) the inertial axes for femur shaft; (c) fit the condyles within the cylinder; (d) desired surface vertices to perform cylinder fitting; (e) cylinder fitting for each condyle; (f) the anatomical femur coordinate frame.

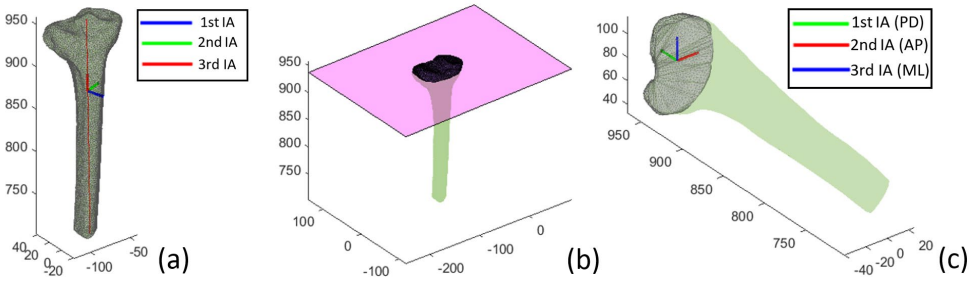


Figure 6.6: Determination of anatomical tibial coordinate. (a) Inertial axes of the tibia; (b) the tibial plateau is separated at the slice with maximum cross-sectional area; (c) the anatomical coordinate frame of tibia.

6.2.4.3 Automatic determination of the landmarks

a) Determination of the flexion angle of the knee joint

As Figure 6.3a illustrates, the knee flexion angle is determined as the angle between PD axes of femur and tibia, which are obtained from 6.2.4.2.

b) Determination of the plane with trochlear groove and posterior condyle reference line

The anatomical axial plane with TG and PCRL should be in the cut that traverses the bottom of TG in its most proximal part [20]. Firstly, the anatomical axial plane of tibia (as the mechanical PD axis and anatomical tibial PD axis is comparable [24]) was used to identify two cuts, one cut (green vertices in Figure 6.7a) was with largest external rectangle

area, and the other cut (yellow vertices in Figure 6.7a) was traversing the P1 in Figure 6.5a. As Figure 6.7a illustrates, the target slice with TG and PCRL was the magenta slice, which separated green slice and yellow slice with the distance ratio between D1 and D2 of 0.2:0.8 experimentally.

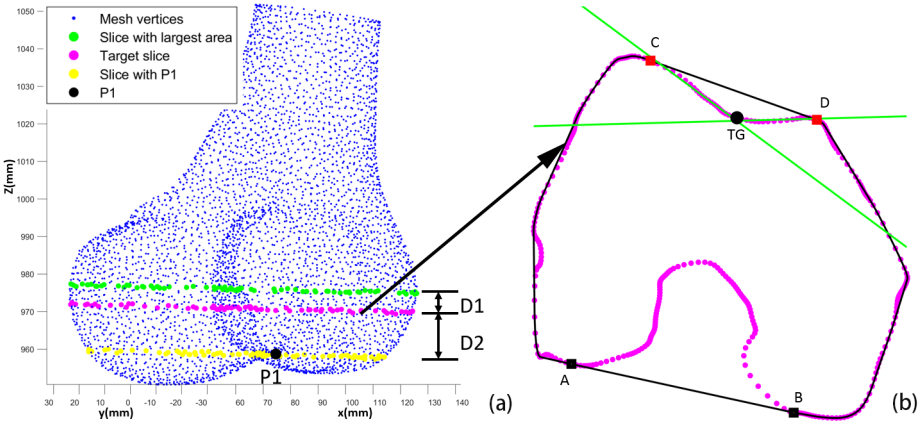


Figure 6.7: Illustration of the calculation of target TG and PCRL. (a) The bone shape model and the plane with largest area (green points), target slice (magenta points) and the slice with P1 (yellow points); (b) the method to calculate the target PCRL and TG (magenta points: vertices of target femur slice; black line indicates the convex hull of femur contour; black boxes represent the points of PCRL; points between red boxes are used to determine the point of TG; green lines indicate the fitting lines of vertices between red box; black circle indicates the point of TG).

c) Determination of the posterior condylar reference line

The PCRL was defined as the largest part of the convex hull of the magenta contour on the target plane (line AB in Figure 6.7b).

d) Determination of the trochlear groove

The trochlear groove was defined as one of the points between the top of the lateral and medial condyle (C and D in Figure 6.7.b) and following the downslope towards the groove. The lateral and medial condyle were defined by one part of the convex hull of the magenta contour, which is the most anterior one along the AP axis of femur. Two lines were simulated using least square distance method based on two groups of points separated by the candidate TGs. The point of TG was defined as the point (black point in Figure 6.7b) that had the minimum sum difference between the points in the selected magenta points (between C and D) and the corresponding points in the simulated lines (green lines in Figure 6.7b).

e) Determination of the tibial tubercle

The point of TT was defined as the most anterior part of the tibial tubercle. As Figure 6.8 illustrates, all the vertices in the tibia shape model were projected onto the anatomical tibial axial plane. The convex hull of the projected points was calculated as the red line (Figure 6.8). The centre points of all segments of the convex hull were calculated and projected on the AP axis of tibia. The two anterior segments with large length were selected (line AB and line CD in Figure 6.8). The point of TT was defined as the middle point (black point in Figure 6.8) of the convex combinations from A to C.

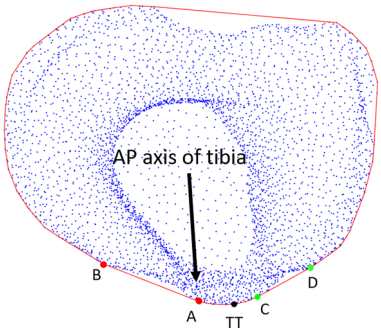


Figure 6.8: Illustration of the calculation of target TT (blue points indicate the projected vertices of tibial model on the tibial axial plane; red line indicates the convex hull; red and green points indicate the most anterior segments of the convex hull; black arrow indicates the AP axis of tibia; black point indicates the point of TT).

f) Determination of the patellar centre

As Figure 6.3c indicates, the centre of patella was determined straightforward by averaging the position of vertices of the patella shape model.

6.2.4.4 Bone shape model registration

To describe the rigid transformation relationship of the bone shape model in different time frames, a point cloud registration was performed to calculate the transformation matrix. In this study, bone shape model registration was done by using the coherent point drift (CPD) algorithm [25], which found the optimal transformation of a shape between consecutive time frames on point to point basis. Viewing the bone shape as a point cloud, CPD finds the transformation by maximizing a likelihood defined in a Gaussian mixture models of the complete set of points. The shape model from 3D scan was first registered to the first dynamic frame (Figure 6.9a), and then in step by step fashion to each dynamic frame [25]. Figure 6.9b gives an overview of a dynamic acquisition with the femur fixed and tibia overplayed for each timeframe. With the transformation matrix, the changing angle between femur and tibia and the changing position of the landmarks along the time frames were identified.

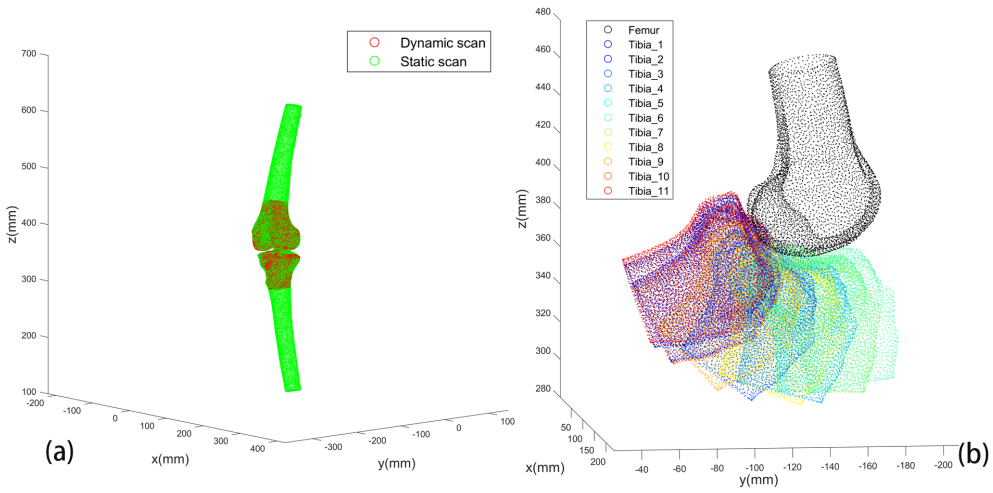


Figure 6.9: The coherent point drift registration among the shape models: (a) between static bone shape model and the dynamic bone shape model in the fifth frame; (b) between the dynamic shape models within a fixed femur

6.2.5 Descriptive Statistics

The manually obtained TT-TG measurements were from three observers where one observer scored one data set three times to attain a measure of intra-observer variability.

Intra- and inter- observer variability is summarized based on the mean TT-TG distance and its maximum offset (Figure 6.10). The results from proposed method is then compared to the manually obtained results.

For the registration error, the mean minimum distance between the vertices of the transformed shape model and the surfaces of the target shape model were calculated for each bone on each frame. Then, the means and the standard deviations based on the mean minimum distances were calculated for femoral condyles, top part of tibia (including the TT), and patella.

For the dynamic TT-TG curves, all the curves are calculated separately for each knee, with a summary of the curves illustrated on Figure 6.11 and Figure 6.12.

6.2.6 Outcome parameters

Firstly, a comparison of the static TT-TG distance between proposed method and manual observation is provided. Secondly, the registration errors of femur, tibia and patella are summarized. Thirdly, the dynamic TT-TG distance and PC-TG distance from the eight patients are provided. Fourthly, a graphical illustration of the dynamic TT-TG distance is shown.

6.3 Results

Figure 6.10 illustrates the range of manually observed TT-TG distance in 3D scan and the distance measured from proposed method for each knee, together with the general accepted pathological threshold and surgical threshold. In the figure, the intra-observer variation of static TT-TG distance in these data sets are associated with a mean difference of 1.2 mm, maximum difference of 4.3 mm, while the inter-observer variation with a mean difference of 2.4 mm, maximum difference of 7.3 mm.

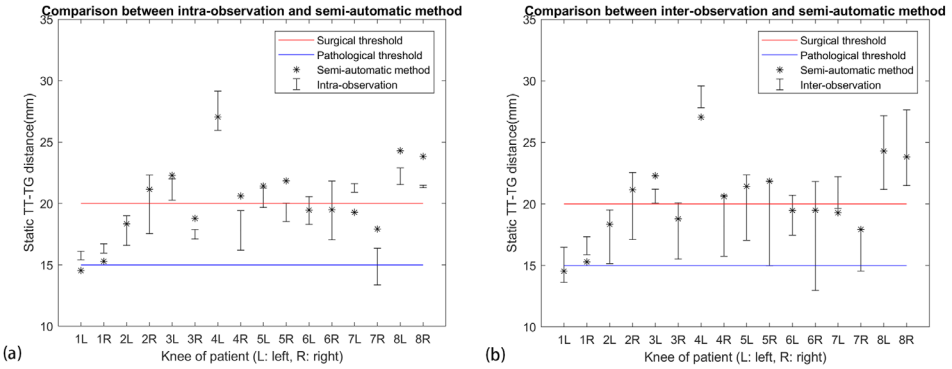


Figure 6.10: The comparison of the static TT-TG distance among the intra-observation (a) / inter-observation (b) and the semi-automatic method for eight patients (black star indicates the TT-TG distance from proposed method; black segment indicates the range of manual observations; blue and red line indicate the pathological threshold and surgical threshold).

For the registration accuracy, the average registration variability of femoral condyles, top part of tibia (including the TT), and patella among legs are associated with 0.56×0.14 mm, 0.57×0.13 mm and 0.55×0.20 mm respectively.

Figure 6.11 and 6.12 illustrate all the quantitative measurements (static/dynamic TT-TG distance, static/dynamic PC-TG distance) with varying knee flexion angle. The dynamic quantitative measurement reflects the knee joint movement starting with flexion then extension then flexion again. The 0° flexion angle was defined as the flexion angle in static scan. The flexion angle for dynamic scan was the difference between itself and the flexion angle in static scan. Besides, the left knee from subject 3 was out of field of view during the 4D CT scan. Therefore, the sample size of the knee for the measurement was 16 for 3D scan and 15 for 4D scan.

As Figure 6.11 illustrates, the TT-TG distance decreases as the flexion angle increases in most cases. Furthermore, the direction of flexion seems to influence the TT-TG distance

(different behaviour in flexion to extension vs extension to flexion). The trend of TT-TG distance variation for left and right are comparable in most cases.

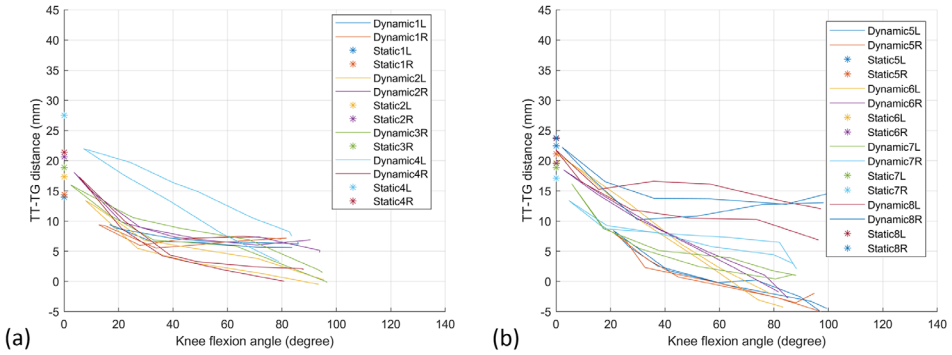


Figure 6.11: Dynamic TT-TG distance as a function of flexion angle for both knee for eight patients ((a): 1-4; (b): 5-8) (curves indicate the dynamic TT-TG distance; stars indicate the static TT-TG distance).

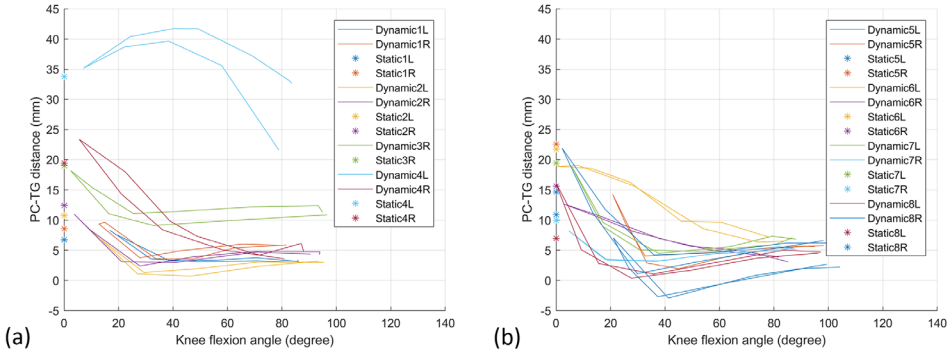


Figure 6.12: Dynamic PC-TG distance as a function of flexion angle for both knee for eight patients ((a): 1-4; (b): 5-8) (lines indicate the dynamic PC-TG distance; stars indicate the static PC-TG distance).

As Figure 6.12 illustrates, although the patellar movement correlates with tibial movement, the PC-TG distance is different from the TT-TG distance. The PC-TG distances decreases until around 30° of flexion angle and then possibly slightly increases or decreases or remains stable. One exception case of patient 4 (Figure 6.12a) illustrates an extreme abnormal curve, caused by patellar dislocation during complete flexion-extension. Static and dynamic PC-TG distance at equal flexion does not match but no clear offset can be seen. Direction of flexion seems to influence the PC-TG distance (not the same route back). The trend of PC-TG distance variation for left and right are comparable in most cases.

A 3D representation of the bones, anatomical frames and TT-TG distances at several selected time frames are illustrated in Figure 6.13.

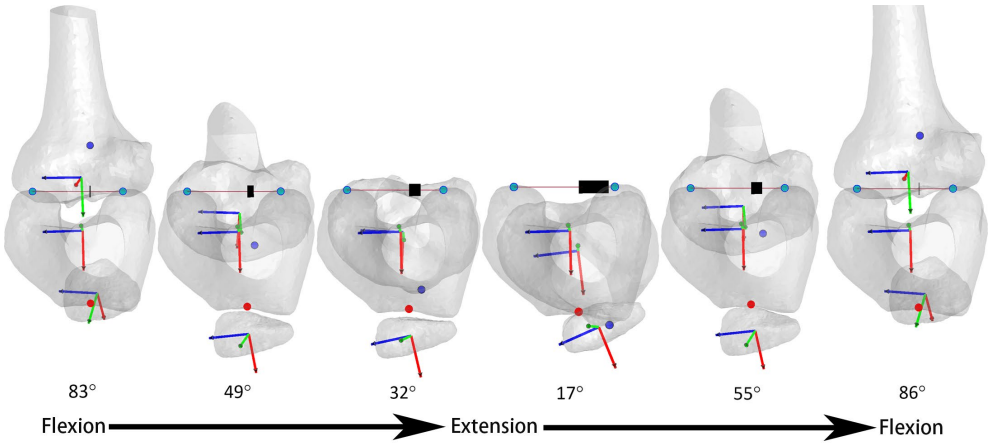


Figure 6.13: Dynamic TT-TG distance in a 3D axial view (the connected line with cyan circles represent the PCRL; red circle represents the point of TG; blue circle represents the point of TT; the black thick line indicates the dynamic TT-TG distance; red, green and blue vectors represent AP, PD and ML axes of femur, tibia and patella, respectively).

6.4 Discussion

Figure 6.10 illustrates the comparable TT-TG distances in the 3D scan between the manual measurements and the measurement from the semi-automatic method. The results of the variance of intra-/inter- observation on TT-TG distance assessed in this study are comparable to the earlier study [26]. Nevertheless, the inter observer variance is still relatively large and might affect the surgical decision in some cases. For that matter, the proposed method provides an option for clinicians to check validity of the measurement. The general accepted indication for a medializing osteotomy on TT-TG distance is 15 mm to 20 mm [5], and 20 mm as an accepted surgical threshold [20]. The method proposed in this paper successfully measured the TT-TG distance for all patients.

Tanaka et al. proposed a manual measurement of TT-TG distance for the dynamic CT image with the multi-planar reconstruction [5]. Nevertheless, in this study, we found it was not easy to reconstruct the femur and tibia with a suitable axial plane according to the definition of TT-TG distance [20]. To remove the large influence of any offset tilt of the axial plane to the TT-TG distance [18, 27] (almost 1 mm per degree), the anatomical coordinate frames of bones are used to calculate the landmarks. Furthermore, the key landmarks required for TT-TG measurement might be outside of the reconstructed FOV during the knee

movement. Hence, a registration of the full bone from static acquisition to the dynamics is applied to solve this limitation.

An important factor contributing to the robustness and reliability of this method is the variability in the registration from frame to frame in the dynamic's series. This variability originates from a range of factors such as the method of registration, resolution and frame rate of the data and accuracy of the segmentation of the bone. In this study we applied coherent point drift registration on the segmented mask, and the mean error was in the order of the one-pixel. The registration error mainly comes from the difference of the segmentation masks. This kind of difference caused from three aspects: 1) the intra-observation error; 2) anisotropic property of the CT images; 3) error from the generation of the bone shape model.

A further decrease in the variability of the outcome measures (i.e. TT-TG and PC-TG distance) might be obtained by optimizing the spatiotemporal resolution with the constraint of the used dosage (0.32 mSv) in this study per patient; the comparing of several registration algorithms [28, 29], i.e. Elastix toolbox, a software used to solve (medical) image registration problems [30], to the CPD algorithm used in this study; and an automated framework for accurate segmentation of the bone based on a statistical shape model [31] or perhaps a deep learning method, which has already been shown to be effective for pig skeleton [32], but to our knowledge not available for bone shape obtained from one 3D image data of the 4DCT dynamic series at this time. A necessary requirement for these kinds of improvements is an expansion of the database, which we hope to achieve soon. Nevertheless, the variabilities of the semi-automatic method as reported in this study are already relatively small compared to the intra- and inter- observation variabilities of the TT-TG, and thus confirming the feasibility of a semi-automated framework for quantitative assessment of patellofemoral disability.

From a clinical perspective, several findings can be reported according to Figure 6.11. Firstly, the TT-TG distance is decreasing as the knee flexion angle increases, which agrees with the findings of other authors [5, 10]. This is probably caused by an internal tibial rotation with higher flexion. Secondly, the TT-TG distance generally follows a slightly different track either going into flexion or going into extension (Figure 6.11), which may originate from active quadriceps contraction and relaxation in different movement direction and possibly from visco-elastic properties of the soft tissue structures around the knee or the cartilage. Registration of the full dynamic cycle (flexion-extension and extension-flexion) would thus ensure a better standardization of TT-TG assessment across subjects as the phase of this dynamic cycle can be synchronized whereas in static measurement this is simply unknown. Thirdly, the characteristics of TT-TG and PC-TG distance vs flexion angle appear to be consistent among patients, indicating it might be of clinical value. Furthermore, the PC-TG distance has the potential to show the clinical reverse J-sign symptom. For patient 4, a

reversed curve is found, because the patella of this patient is completely out of trochlear groove during the knee movement. Finally, the TT-TG and PC-TG distances are not the only option to assess the quantitative measurement of knee kinematics, e.g. patellar tilt [12], bisect offset measurement [5, 11]. To evaluate the best parameters to depict the knee joint kinematics, a larger dataset should be included, and we will explore more related parameters in the future.

Although the results from the proposed method show the promising dynamic description of movement between femur and patella, several limitations remain. Firstly, the method is evaluated with a patient group of eight patients. A larger dataset should be included to confirm the clinical findings. Secondly, the proposed method still requires substantial manual input both in segmentation of the bones in the static scan and the dynamic series. While segmentations of the lower extremity CT bone shapes are done manually, this manual segmentation still poses a large part of the work required for data processing i.e. a bottleneck in the goal of clinical implementation and exists the problems of reproductivity and accuracy. A deep learning based approach to automate robust segmentation of the bone shapes will be introduced in Chapter 7. Thirdly, a more comprehensive investigation of the intra and inter observer variation on a large dataset will be scheduled to determine the clinical value of the TT-TG / PC-TG distance. Fourthly, in the current workflow we rely on a 3D CT scan to determine the anatomical coordinate frame. This leads to an increase in the dosage, but still in the range of common clinical protocols (Pelvic CT 10 mSv; Head CT 2 mSv; mammography 0.4 mSv, for more information see for instance [33]). In the future, we will research on the coordinate system determination only based on the 4DCT data.

6.5 Conclusion

In this work, a semi-automatic workflow was developed to quantitatively assess the patellofemoral joint movement by robust determination of tibial tubercle – trochlear groove / patellar centre- trochlear groove distance during active flexion-extension-flexion movement of the knee. The workflow provides a robust and reproducible TT-TG/PC-TG distance, and the comparison with manually obtained data confirms its validity. With the registration process, the proposed workflow extends the quantitative joint movement measurement to 0°-90° in spite of limited field of view of dynamic CT. Initial quantitative results of the knee joint movement on a group of eight patients have shown that the complementary value of the dynamic tibial tubercle – trochlear groove / patellar centre- trochlear groove distance to that of the static measurement. Further research will be focused on the automatic segmentation of the knee joint and including more patients to confirm the clinical value of the dynamic quantitative description of the knee joint movement.

Conflict of interests

None

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP), n. 530-4CDP11 from Koninklijke Nederlandse Akademie Van Wetenschappen (KNAW).

Ethical Approval

Eight patients were included for data acquisition in this study. All Patients gave informed written consent and permission was granted by local Medical Ethical committee. The reference number from Medical Ethical committee is NL60392.091.17.

Acknowledgements

We greatly thank Léon Driessen for the bone shape models and Marieke Ploegmakers for the manual assessments of TT-TG distance.

References

[1] R. Nikku, Y. Nietosvaara, K. Aalto, and P. E. Kallio, "Operative treatment of primary patellar dislocation does not improve medium-term outcome," *Acta Orthopaedica,* vol. 76, pp. 699-704, 2005, doi: 10.1080/17453670510041790.

[2] T. O. Smith, F. Song, S. T. Donell, and C. B. Hing, "Operative versus non-operative management of patellar dislocation. A meta-analysis," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 19, pp. 988-998, 2011, doi: 10.1007/s00167-010-1355-2.

[3] A. T. Pennock, M. Alam, and T. Bastrom, "Variation in tibial tubercle-trochlear groove measurement as a function of age, sex, size, and patellar instability," *American Journal of Sports Medicine,* vol. 42, pp. 389-393, 2014, doi: 10.1177/0363546513509058.

[4] C. L. Camp *et al.*, "CT and MRI measurements of tibial tubercle-trochlear groove distances are not equivalent in patients with patellar instability," *American Journal of Sports Medicine,* vol. 41, pp. 1835-1840, 2013, doi: 10.1177/0363546513484895.

[5] M. J. Tanaka, J. J. Elias, A. A. Williams, J. A. Carrino, and A. J. Cosgarea, "Correlation Between Changes in Tibial Tuberosity-Trochlear Groove Distance and Patellar Position During Active Knee Extension on Dynamic Kinematic Computed Tomographic Imaging," *Arthroscopy : the journal of arthroscopic & related surgery,* vol. 31, pp. 1748-1755, 2015, doi: 10.1016/j.arthro.2015.03.015.

[6] N. Purohit, N. Hancock, and A. Saifuddin, "Surgical management of patellofemoral instability. I. Imaging considerations," *Skeletal Radiology,* 2018, doi: 10.1007/s00256-018-3123-1.

[7] P. Cao *et al.*, "Ratio of the tibial tuberosity–trochlear groove distance to the tibial maximal mediolateral axis: A more reliable and standardized way to measure the tibial tuberosity–trochlear groove distance," *Knee,* vol. 25, pp. 59-65, 2018, doi: 10.1016/j.knee.2017.10.001.

[8] K. E. Wilk, G. J. Davies, R. E. Mangine, and T. R. Malone, "Patellofemoral Disorders: A Classification System and Clinical Guidelines for Nonoperative Rehabilitation," *Journal of Orthopaedic & Sports Physical Therapy,* vol. 28, pp. 307-322, 1998, doi: 10.2519/jospt.1998.28.5.307.

[9] L. W.B., R. P.S., and M. M.A., "The appropriate use of patellofemoral arthroplasty: An analysis of reported indications, contraindications, and failures," *Clinical Orthopaedics and Related Research,* pp. 91-99, 2005.

[10] G. Seitlinger, G. Scheurecker, R. Högler, L. Labey, B. Innocenti, and S. Hofmann, "The position of the tibia tubercle in 0°–90° flexion: comparing patients with patella dislocation to healthy volunteers," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2396-2400, 2014, doi: 10.1007/s00167-014-3173-4.

[11] M. J. Tanaka, J. J. Elias, A. A. Williams, S. Demehri, and A. J. Cosgarea, "Characterization of patellar maltracking using dynamic kinematic CT imaging in patients with patellar instability," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 24, pp. 3634-3641, 2016, doi: 10.1007/s00167-016-4216-9.

[12] D. Forsberg, M. Lindblom, P. Quick, and H. k. Gauffin, "Quantitative analysis of the patellofemoral motion pattern using semi-automatic processing of 4D CT data," *International Journal of Computer Assisted Radiology and Surgery,* vol. 11, pp. 1731-1741, 2016, doi: 10.1007/s11548-016-1357-8.

[13] V. R. Carlson, F. T. Sheehan, A. Shen, L. Yao, J. N. Jackson, and B. P. Boden, "The Relationship of Static Tibial Tubercle–Trochlear Groove Measurement and Dynamic Patellar Tracking," *The American Journal of Sports Medicine,* vol. 45, p. 036354651770011, 2017, doi: 10.1177/0363546517700119.

[14] K. Izadpanah *et al.*, "Influence of knee flexion angle and weight bearing on the Tibial Tuberosity-Trochlear Groove (TTTG) distance for evaluation of patellofemoral alignment," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2655-2661, 2014, doi: 10.1007/s00167-013-2537-5.

[15] T. Ohnishi, M. Suzuki, A. Nawata, S. Naomoto, T. Iwasaki, and H. Haneishi, "Three-dimensional motion study of femur, tibia, and patella at the knee joint from bi-plane fluoroscopy and CT images," *Radiological Physics and Technology,* vol. 3, pp. 151-158, 2010, doi: 10.1007/s12194-010-0090-1.

[16] M. Ishimaru *et al.*, "Three-dimensional motion analysis of the patellar component in total knee arthroplasty by the image matching method using image correlations," *Journal of Orthopaedic Research,* vol. 32, pp. 619-626, 2014, doi: 10.1002/jor.22596.

[17] Y. Kwong, A. O. Mel, G. Wheeler, and J. M. Troupis, "Four-dimensional computed tomography (4DCT): A review of the current status and applications," *Journal of Medical Imaging and Radiation Oncology,* vol. 59, pp. 545-554, 2015, doi: 10.1111/1754-9485.12326.

[18] L. Yao, N. Gai, and R. D. Boutin, "Axial scan orientation and the tibial tubercle-trochlear groove distance: Error analysis and correction," *American Journal of Roentgenology,* vol. 202, pp. 1291-1296, 2014, doi: 10.2214/AJR.13.11488.

[19] V. R. Carlson, B. P. Boden, A. Shen, J. N. Jackson, L. Yao, and F. T. Sheehan, "The Tibial Tubercle-Trochlear Groove Distance Is Greater in Patients with Patellofemoral Pain: Implications for the Origin of Pain and Clinical Interventions," *American Journal of Sports Medicine,* vol. 45, pp. 1110-1116, 2017, doi: 10.1177/0363546516681002.

[20] W. N. Scott, "Surgery of the Knee," *Insall & Scott Surgery of the Knee,* pp. 1-1504, 2014, doi: 9781455727896.

[21] D. L. Miranda, M. J. Rainbow, E. L. Leventhal, J. J. Crisco, and B. C. Fleming, "Automatic determination of anatomical coordinate systems for three-dimensional bone models of the isolated human knee," *Journal of Biomechanics,* vol. 43, pp. 1623-1626, 2010, doi: 10.1016/j.jbiomech.2010.01.036.

[22] D. G. Eckhoff *et al.*, "Three-Dimensional Mechanics, Kinematics, and Morphology of the Knee Viewed in Virtual Reality," *The Journal of Bone and Joint Surgery (American),* vol. 87, p. 71, 2005, doi: 10.2106/JBJS.E.00440.

[23] S. D. Masouros, A. M. J. Bull, and A. A. Amis, "(i) Biomechanics of the knee joint," *Orthopaedics and Trauma,* vol. 24, pp. 84-91, 2010, doi: 10.1016/j.mporth.2010.03.005.

[24] J. J. Cherian, B. H. Kapadia, S. Banerjee, J. J. Jauregui, K. Issa, and M. A. Mont, "Mechanical, anatomical, and kinematic axis in TKA: Concepts and practical applications," *Current Reviews in Musculoskeletal Medicine,* vol. 7, pp. 89-95, 2014, doi: 10.1007/s12178-014-9218-y.

[25] A. Myronenko and X. Song, "Point Set Registration: Coherent Point Drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 2262-2275, 2010, doi: 10.1109/TPAMI.2010.46.

[26] J. M. Marzo, M. A. Kluczynski, A. Notino, and L. J. Bisson, "Measurement of tibial tuberosity-trochlear groove offset distance by weightbearing cone-beam computed tomography scan," *Orthopaedic Journal of Sports Medicine,* vol. 5, pp. 1-5, 2017, doi: 10.1177/2325967117734158.

[27] S. Harris and J. Cobb. Reliability of TT-TG measurements – standardizing the frame of reference.

[28] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: A review," in *Computer Methods in Biomechanics and Biomedical Engineering* vol. 17, ed: Taylor & Francis, 2014, pp. 73-93.

[29] G. K. L. Tam *et al.*, "Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid," *Visualization and Computer Graphics, IEEE Transactions on,* vol. 19, pp. 1199-1217, 2013, doi: 10.1109/TVCG.2012.310.

[30] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "elastix: A Toolbox for Intensity-Based Medical Image Registration," *IEEE Transactions on Medical Imaging,* vol. 29, pp. 196-205, 2010, doi: 10.1109/TMI.2009.2035616.

[31] T. Heimann and H. P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis,* vol. 13, pp. 543-563, 2009, doi: 10.1016/j.media.2009.05.004.

[32] J. Kvam, L. E. Gangsei, J. Kongsro, and A. H. S. Solberg, "The use of deep learning to automate the segmentation of the skeleton from CT volumes of pigs," *Translational Animal Science,* vol. 2, pp. 324-335, 2018, doi: 10.1093/tas/txy060.

[33] "Radiation Dose in X-Ray and CT Exams," March 27th 2017. [Online]. Available: https://www.radiologyinfo.org/en/pdf/safety-xray.pdf.

Chapter 7 Fully automatic quantitative measurement of knee joint movement based on 4DCT using deep neural network for automatic segmentation of bones

Abstract:

Introduction: The knee joint movement can be quantitatively assessed by four dimensional computed tomography (4DCT). However, the manual quantitative assessment is tedious and time consuming. In this study, the authors aim at developing an automatic workflow that can quantify the knee joint movement based on 4DCT image using the deep convolution neural network as the automatic bone segmentation method.

Methods: Automatic segmentation methods based on deep convolution neural networks (3D U-Net/3D V-Net) were proposed. An image standardization method was proposed to improve the robustness and performance of the neural networks. The results from the proposed segmentation method were then used as the input for generation of a clinically relevant quantitative measure (i.e. tibial tubercle – trochlear groove (TT-TG) distance) of knee joint movement and compared with manual segmentation input.

Results: In a test dataset of 4 patients (44 images from the 4D scan and 4 images from the 3D scan), the proposed segmentation method achieved a mean dice similarity coefficient (DSC) for femur, tibia, and patella of 0.98 ± 0.01, 0.97 ± 0.02, and 0.97 ± 0.01, respectively. The difference of TT-TG distance between the manual segmentation and automatic segmentation was relatively small (0.3 ± 0.3 mm), compared with the manual intra-observation errors of 1.2 ± 1.0 mm and the manual inter-observation of 2.4 ± 1.8 mm.

Conclusion: The deep neural network based segmentation method provides an accurate bone structure segmentation in 4D CT image and enables a robust and automated workflow for quantitative measurement of knee joint movement.

7.1 Introduction

Patellar instability, featured as dislocation of the patella from the trochlear groove, is one of the major causes of knee pain. Whereas lesions in cartilage and bone can be diagnosed by 3D medical imaging techniques (i.e. CT, MRI), the motion related pathologies are not easily confirmed by the 3D scanning. An increasing number of studies have shown the benefit of having dynamic information of the knee, resulting in improved diagnosis and more support in surgical decision making [1-3].

Recently, four dimensional computed tomography (4DCT) have been introduced because of possessing a substantially higher acquisition rate over conventional CT. Tanaka et al. demonstrated the possibility to quantify the mal-tracking pattern of patella movement in patients with patellofemoral instability based on 4DCT [3]. These authors found a decrease in tibial tubercle – trochlear groove (TT-TG) distance with increasing flexing angle of the knee joint, together with similar trends in bisect offset and patellar tilt [4], indicating the benefit of dynamic information on the diagnosis of patellar instability [3].

However, the quantitative measurements of patellar instability based on 4DCT is time consuming, in that 4DCT provides an order of magnitude more 3D images per subject. Hence, an automatic workflow to analyse the 4DCT data is beneficial to the knee joint related researches. The key steps to automate the quantitative analysis is to identify the bone landmarks and bone to bone transformation matrices between time frames. Both steps require the automatic segmentation of the bone.

The majority of commercial software available to perform CT bone segmentation was based on global thresholding and requires extensive manual post-processing [5]. The difficulties of the bone segmentation in CT images come from partial volume effects, CT value distortion caused by an implant and motion artefact (especially in 4DCT data) [6]. Recently, Statistical Shape Models (SSMs) [5] and atlas based segmentation algorithms [7, 8] have been suggested for fully automated segmentation of bone from CT images. Despite the encouraging results by model-based and atlas-based segmentation methods, the approaches might perform poorly when there is high subject variability and significant differences in local features [5, 9]. Therefore, a fully automatic and robust bone segmentation method for 4DCT is to our knowledge currently not yet available.

In recent years, tremendous progress has been made for automatic medical image analysis with the introduction of deep convolutional neural networks (CNN) [10]. CNN based method and its variants (SegNet [9, 11], U-Net [12-14], GAN [15, 16]) haven shown the significant advantages on image segmentation. Regarding knee joint segmentation, Liu et al. proposed a 2D semantic segmentation CNN model to perform cartilage and bone segmentation slice

by slice and refined the results using a 3D deformable model based on MRI [9]. Ambellan et al. proposed to firstly use a 2D CNN to obtain the bone mask and a 3D CNN to obtain the cartilage mask, and then used the statistical shape model to refine the outcomes of CNN based on MRI images [17]. Regarding CT bone segmentation, Minnema et al. proposed a 2D CNN model to segment the skull. To the best of our knowledge, the CNN based model has not been applied to knee joint bone segmentation based on 4DCT.

The purpose of this paper is to fully automate the quantitative analysis (dynamic TT-TG distance) of patellofemoral movement using the deep neural networks for automatic knee bone segmentation. The main contributions of this study are: (1) application of deep neural networks to knee bone segmentation for 4DCT data; (2) analysis of the possibility of a fully automatic workflow on quantitative measurement (TT-TG distance) of patellofemoral instability from 4DCT data.

7.2 Materials and method

7.2.1 Data acquisition

Informed consent was obtained in accordance with the institutional review board approved study protocol. Eight patients underwent 3D and 4D CT scans with low dose on a Canon Aquilion One Genesis scanner. As Figure 7.1a shows, one 3DCT scan on both knees in full extension was acquired to provide anatomical information. The obtained volume had a fixed matrix size of 512×512, 626 slices. The in-plane pixel spacing varied from 0.511×0.511 mm to 0.824×0.824 mm. Slice thickness was 1 mm with an increment of 0.8 mm. As Figure 7.1b and Figure 7.1c show, a 4DCT scan with eleven frames was acquired in both knees within 13 seconds during active flexion-extension-flexion movement i.e. a temporal resolution of 0.85 Hz. The obtained volumes had the fixed matrix size of 512×512, 320 slices. The in-plane pixel spacing varied from 0.820×0.820 mm to 0.976×0.976 mm. Slice thickness was 0.5 mm with an increment of 0.5 mm. This CT protocol was associated with an ionising radiation dose of 0.32 millisievert (mSv), compared to normal preoperative CT scan of the knee of 0.02 mSv.

Figure 7.1: (a) The 3D scan with fully extension; (b) the 4D scan with 90° knee flexion; (c) the 4D scan with fully extension

7.2.2 Manual segmentation of the bone and measurement of the TT-TG distance

The manual segmentation of bones were implemented using the Mimics software environment on 96 image datasets (8 for 3D scan and 88 for 4D scan). For the evaluation of the proposed workflow, the intra-observation of TT-TG distance was performed three times (one-week interval) by an experienced radiologist. The inter-observation was performed by two experienced radiologists and one experienced orthopaedic surgeon. All manual TT-TG distance were measured using Lyon's protocol [18] and adjusted using the method by Yao et al. [19]. The TT-TG distance was assessed on the 3D scan for both legs of 8 patients (16 legs).

7.2.3 Automatic workflow for quantitative measurement of knee joint movement

The quantitative measurement of the knee joint movement is based on the geometrical landmarks of the bone; therefore, an automatic, accurate and fast bone segmentation is important if we want to automate the quantitative measurement, especially for the clinical practises. In this study, each 4DCT data consisted of eleven 3D image volumes, which was a big workload and impossible to implement manual segmentation in clinical practice. Hence, a fully automatic workflow was proposed to quantify the knee joint movement in this study. As Figure 7.2a illustrates, the proposed automatic workflow for bone segmentation consisted of image normalization to pre-process the images, a deep neural network (3D U-Net/3D V-Net) to generate the bone classification probability map, a post-processing step to refine the bone masks and a bone shape model generation to convert the segmented masks to 3D models. The automatic workflow for the quantitative measurement of the knee joint movement is illustrated in Figure 7.2b. The anatomical coordinate frame of bones and the bone landmarks were determined based on bone shape models from segmented bone masks of 3D CT images, while motion related information was provided by the transformation matrices based on segmented bone masks of 4DCT images. Eventually, the dynamic TT-TG distance was

determined based on anatomical bone coordinate frames, bone landmarks and transformation matrices.

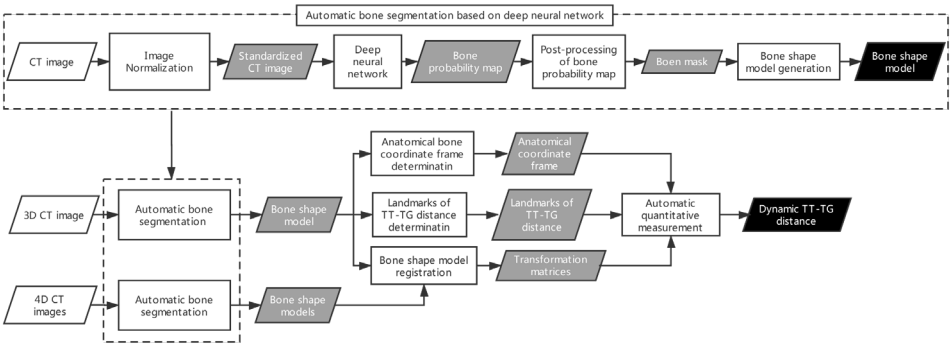


Figure 7.2: (a) Pipeline of the automatic segmentation of knee bone based on CT images; (b) pipeline of automatic quantitative measurement (TT-TG distance) of the knee joint movement. (white, grey and black parallelogram represent input data, intermediate results and output result, respectively; rectangle box represents a step of the workflow).

The proposed deep CNNs were implemented in Python 3.7 using Pytorch with a 3.7 GHz Intel(R) i7 Xeon (R) E5-1620 V2 CPU and a GTX 1080 Ti graphic card with 11GB GPU memory. The rest of automatic methods were implemented in Matlab r2017a.

7.2.4 Proposed methods for automatic segmentation of knee bones

7.2.4.1 Image normalization

The intensity range of each tissue in CT image is scale to a specific range based on Hounsfield units, which facilitates the segmentation task. However, in knee joint CT image, the metal implant might result in pixels with unexpected high and low intensity on the reconstructed CT images. Conventionally, a normalization step, min-max or z-score, would be applied to normalize the images into a similar scale before inputting to the neural network for training. However, they might fail to work if the pixels with extreme intensity exist.

To solve this issue, the metal implants were detected based on the number of pixels from metal intensity range. If the metal implant exists, all the pixels higher than the metal intensity threshold, 2100, would be set to 2100; and all the pixels lower than -2048 would be set to -2048. Moreover, a robust intensity cut-off was selected to prevent the long intensity tails effect [20]. In this study, the minimum and maximum cut-off was selected as the threshold with first and last 2% cumulative histogram. After above steps, a z-score normalization was applied to the image volumes.

7.2.4.2 Deep neural networks for bone segmentation

a) Basic architectures of 3D U-Net and 3D V-Net

The tested networks in this study were 3D U-Net and 3D V-Net, which had tremendous success in biomedical segmentation tasks in past years [12, 13, 17, 21, 22]. As Figure 7.3 shows, both architectures of 3D U-Net and 3D V-Net consisted of an encoding path to encode the valid features and a decoding path to perform a voxel based classification. For the proposed U-Net in this study, the encoding path contained the repeated layers of two convolutions, followed by a batch normalization, rectified linear unit and a max pooling operation with stride 2 for down-sampling. The up-sampling path also contained the repeated layers of convolution, but a skip connection was adopted by a concatenation of the correspondingly cropped features from contraction path and the output of the up-convolution result from last layer. At the final layer, a final $1 \times 1 \times 1$ convolution was used to map each 64 component feature vectors to the desired number of classes, and a soft-max calculation was followed at last to output a probability for each voxel of each class. The used V-Net in this study was depicted in Figure 7.3b, the main differences from U-Net were as follows, the convolution kernel was $5 \times 5 \times 5$; the max pooling was replaced by a $2 \times 2 \times 2$ convolution with a stride of 2; in each stage of convolutions, a residual function was learned [23]; the activation function was replaced by exponential linear units (ELU) [24].

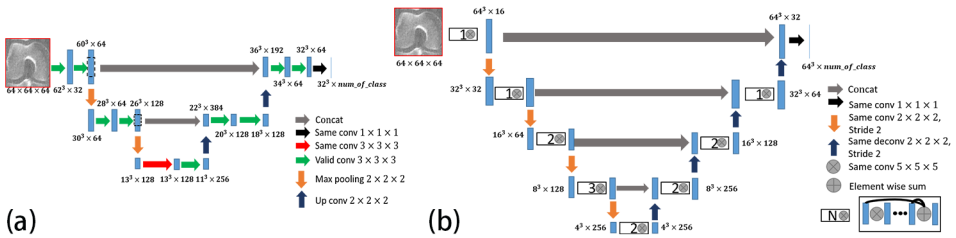


Figure 7.3: Architectures of the 3D U-Net and 3D V-Net (red square indicates one training patch)

The, loss function was computed by a voxel-wise soft-max over the final feature maps combined with the cross-entropy loss function as

$$p_{\mathrm{ij}}(x_{\mathrm{ij}}) = \mathrm{softmax}(x_{ij}) = \frac{e^{x_{ij}}}{\sum_{j=1}^{J} e^{x_{ij}}} \tag{7.4}$$

$$LOSS = -\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij} * \log p_{ij} \tag{7.5}$$

Where x_{ij} represents the output of the last convolution layer of class j for sample i, J represents the number of class, N represents the number of sample, y_{ij} means the ground truth label (0 or 1) and $\log p_{ij}$ means the predicted probability for sample i of class j, repectively.

b) Data augmentation:

In order to enlarge the variability of the training patch, image augmentations were applied before inputting the patch to the networks in this study, which consisted of random Gaussian noise, random rotation (-10°~10° rotation based on the x, y, z image axis).

c) Training

As data was available in two general levels of resolution: the static sets acquired in resolution $0.6 \times 0.6 \times 1\ \mathrm{mm}^3$ and the dynamic sets in resolution $0.9 \times 0.9 \times 0.5\ \mathrm{mm}^3$, the networks were trained for these two levels of resolution, using 1 testing set and 7 training sets for the "static" resolution, and 44 testing set and 44 training sets for the "dynamic" resolution. All sets were classified into four labels: background, femur, tibia and patella. The 3D volumes were split up into $64 \times 64 \times 64\ \mathrm{mm}^3$ patches with overlap by 32, after which 95% of the patches containing background only were discarded. In total, 13200 patches were used for training in one epoch for the dynamic data, and 10500 for the static data. The batch size was 4. The network was optimized using the Adam optimizer with an initial learning rate of 0.001. The stopping criterium was no improvement for the loss more than 5 epoch iterations or maximum of 200 iterations was achieved.

d) Testing

In the inference phase, the new input image volume was split into many sub-volume patches associated with the size of $64 \times 64 \times 64$ and the output of network was the probability maps of each class with the size of $\mathrm{J} \times 32 \times 32 \times 32$ for U-Net, $\mathrm{J} \times 64 \times 64 \times 64$ for V-Net, J is the number of class. Then, the class of each voxel was determined by the largest probability of the probability maps. At last, we needed to combine all the volume patches back to form a full volume.

7.2.4.3 Post-processing of bone probability map

Based on the probability map of each class from the deep neural network, voxels can be classified as femur, tibia, patella, and background. However, some voxels might still be mis-classified (Figure 7.4b). In this study, we used the image hole filling technique [25] to fill the mis-classified hole (Figure 7.4b, red circle). Regarding the mis-classified femur, tibia and patella (Figure 7.4b green circle), a flood-fill algorithm [25] was applied to label all non-

background voxels. Then, a component analysis was performed where the number of classifications of the connected labelled volumes were summarized and the connected volumes would be relabelled as the class with largest amount.

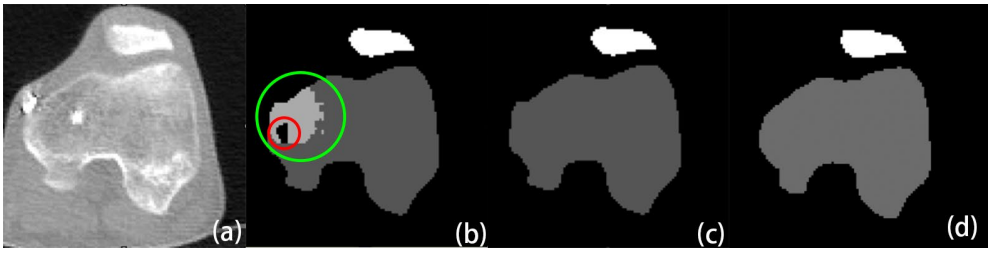


Figure 7.4: Example of mis-classification and its refinement: (a) original image; (b) classification from 3D U-Net (dark grey: femur; light grey: tibia; white: patella; area in red circle indicates the mis-classification of femur as background; area in green circle indicates the mis-classification of femur as tibia); (c) refined result after operations of fill hole and component analysis; (d) ground truth.

The final step in automatic segmentation was to generate the bone shape model from the binary volumetric image using iso2mesh toolbox [26] in order to perform the registration.

7.2.5 From bone surface to quantitative measurement of patella instability

7.2.5.1 Registration of bone

One essential step to quantitatively describe the patellar instability is to quantify the transformation relationship of bones in a time series. To achieve this, the bone shape model registration was implemented using the coherent point drift algorithm [27]. The bone shape model from the static 3D scan was first registered to the first frame of the 4D data set, and then in step by step fashion to each frame, as depicted in Figure 7.5.

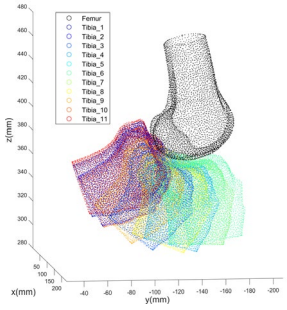


Figure 7.5: The coherent point drift registration between bone shapes (femur and tibia) from time frame to time frame

7.2.5.2 Quantitative measurement of patella instability

As Figure 7.6a shows, the flexion angle (α) between femur and tibia was calculated from the anatomical proximal-distal (PD) axes of femur and tibia (See Chapter 6 for the detail of automatic definition of anatomical coordinate frame of bones). Full extension was defined as zero flexion, negative flexion thus indicates hyperextension. As Figure 7.6b shows, the TT-TG distance was the distance between two projected points of trochlear groove and tibial tubercle and measured on a posterior condylar reference line (PCRL). The TG point was defined as the deepest part of the trochlear groove [18]. The PCRL was defined as the tangent line to both posterior condyles [18]. The TT-point was defined as the most proximal part of tibial tubercle [18]. In this study if the TT point was on the lateral side of the TG point, we defined the distance as positive, otherwise negative if on the medial side. We refer to Chapter 6 for details of the automatic detection of landmarks.

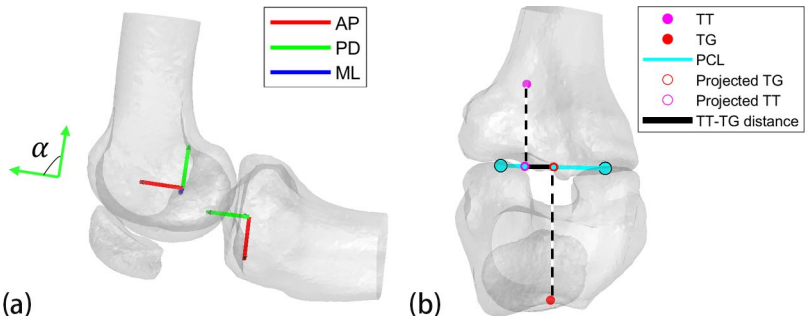


Figure 7.6: Illustration of the relevant landmarks for the quantitative measurement of patellofemoral motion. (a) red, green and blue arrows indicate the corresponding anterior-posterior (AP), proximal-distal (PD) and medial-lateral (ML) directions, α indicates the flexion angle between the PD axes of femur and tibia; (b) red filled and magenta filled circles indicate the TG and TT, red and magenta circles indicate the projected TG and TT, cyan line indicates the PCRL and black thick line indicates the TT-TG distance.

7.2.6 Outcome parameters

For evaluation the robustness of the proposed method, four outcomes were calculated to compare the manual and automatic segmentation. Firstly, manual and automatic segmented masks were compared directly by calculating the Dice Similarity Coefficient (DSC): $\frac{2T_P}{2T_P+F_P+F_N}$, Sensitivity: $\frac{T_P}{T_P+F_N}$, Specificity: $\frac{T_N}{T_N+F_P}$, and Positive Predictive Value (PPV): $\frac{T_P}{T_P+F_P}$ [28] for all 4D (dynamic) and 3D (static) scans, where T_P is the true positive, F_P is false positive and F_N is false negative. Secondly, a comparison of landmarks determination between using manual segmentation and using automatic segmentation was provided. Thirdly, the landmarks difference caused by the registration using manual segmentation and

using automatic segmentation was determined. Fourthly, the difference of TT-TG distance was calculated between using manual segmentation and automatic segmentation as input.

7.3 Results

Figure 7.7 depicts the segmentation results based on different normalization techniques, where first row represents the normal situation; second row represents the case with implant. Table 7.1 summarizes the performance of the femur segmentation based on 12 image dataset, 24 femurs (11 from 4D CT and 1 from 3D CT) with the implant.

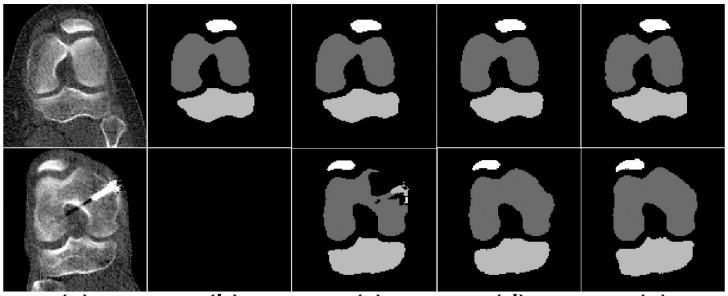


Figure 7.7: (a) Original image (display range is [-200,1000]); (b) normalized by min-max technique; (c) normalized by z-score; (d) normalized by proposed method; (e) ground truth.

Table 7.1: Comparison of femur segmentation between different normalization techniques based on image with implant

	DSC		**Sensitivity**		**Specificity**		**PPV**	
	U-Net	V-Net	U-Net	V-Net	U-Net	V-Net	U-Net	V-Net
Min-Max	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	NaN	NaN
z-score	0.92 ± 0.01	0.95 ± 0.01	0.85 ± 0.02	0.91 ± 0.02	1.00 ± 0.00	1.00 ± 0.00	0.98 ± 0.01	0.99 ± 0.01
Proposed method	0.98 ± 0.01	0.97 ± 0.00	0.97 ± 0.01	0.96 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.98 ± 0.00	0.99 ± 0.00

In Table 7.2, the DSC (mean ± standard deviation) of femur, tibia and patella based on proposed 3D V-Net were 0.99 ± 0.01, 0.96 ± 0.04, and 0.97 ± 0.01, respectively, and the DSC of femur, tibia and patella for 3D U-Net were 0.98 ± 0.01, 0.97 ± 0.02 and 0.97 ± 0.01, respectively.

Table 7.2: Performance comparison between different methods based on similar topics (Statistical shape model (SSM); Hybrid-level set (HLS), Intensity and histogram (IH); Pseudo 3D U-Net (PNet); Whole body (WB))

	3D U-Net			3D V-Net			SSM[7]	HLS[30]	IH[30]	PNet[14]
	Femur	Tibia	Patella	Femur	Tibia	Patella	Femur	Bone	Bone	WB
DSC	0.98 ± 0.01	0.97 ± 0.02	0.97 ± 0.01	0.98 ± 0.01	0.97 ± 0.03	0.97 ± 0.01	0.97 ± 0.01	0.73 ± 0.21	0.88 ± 0.14	0.92 ± 0.05
Sensitivity	0.98 ± 0.02	0.97 ± 0.03	0.98 ± 0.02	0.97 ± 0.01	0.96 ± 0.04	0.96 ± 0.02	-	0.85 ± 0.09	0.94 ± 0.15	0.91 ± 0.08
Specificity	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.04	-	0.99 ± 0.01	1.00 ± 0.00	1.00 ± 0.00
PPV	0.99 ± 0.01	0.97 ± 0.02	0.97 ± 0.02	0.99 ± 0.00	0.98 ± 0.02	0.98 ± 0.02	-	0.90 ± 0.15	0.90 ± 0.15	0.94 ± 0.04

Figure 7.8a shows the difference of automatic landmarks detection from 3D scan between using the masks from our proposed method and using the manual masks for four patients, where TG was 1.2 ± 0.2 mm, TT was 1.5 ± 0.5 mm and the landmarks of PCRL were 1.4 ± 1.1 mm. Figure 7.8b shows the landmark difference between the registration of the masks using the proposed method and using the manual masks. A total of 77 leg data sets contributed to the statistics (11 images of a left leg from patient 3 were out of the field of view). The distance difference between the registration of TG, TT, the landmarks of PCRL were 0.75 ± 0.15 mm, 0.91 ± 0.28 mm, and 0.81 ± 0.16 mm (mean $\pm$ standard deviation). In contrast, the intra-observed variation of TG, TT and the landmarks of PCRL were 2.2±1.5mm, 2.1±1.3mm, and 2.2±1.2mm, while the variation of the inter-observed variations were 5.5±3.6mm, 4.3±3.2mm, and 5.3±3.4mm (Figure 7.8c and Figure 7.8d).

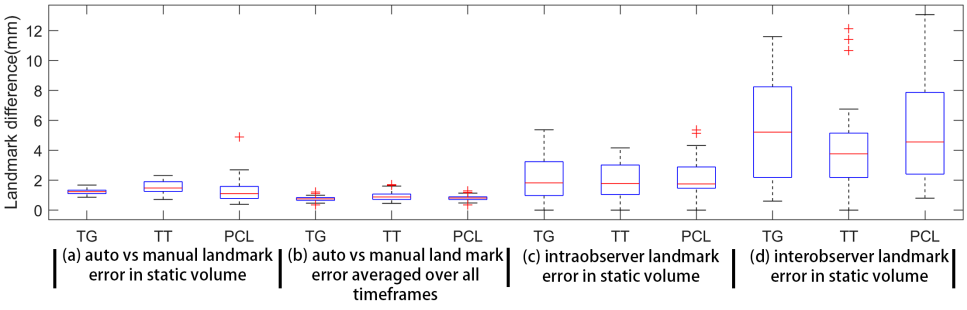


Figure 7.8: The position difference in detected landmarks: (a) point to point distance in landmarks determined from using automatic mask vs using manual mask in 3D scan; (b)

average difference in landmarks between registered by automatically and manually generated masks of all dynamic time frames; (c) point to point intra observer landmark error in 3D scan based on three separate measurements from one experienced observer; (d) point to point inter observer landmark error in 3D scan based on three separate measurements from three experienced observers

Figure 7.9a shows the TT-TG distance difference between manual masks and the masks from proposed method as a function of time frames, and each time frame has seven data points. The overall TT-TG distance difference was 0.3 ± 0.3 mm from 44 image frames. Figure 7.9b depicts the quantitative measurement of knee joint movement based on dynamic TT-TG distance as a function of flexion angle. It is clear that for most knees the TT-TG reduces with higher flexion angles, which is conformed with the knee-kinematic expectations.

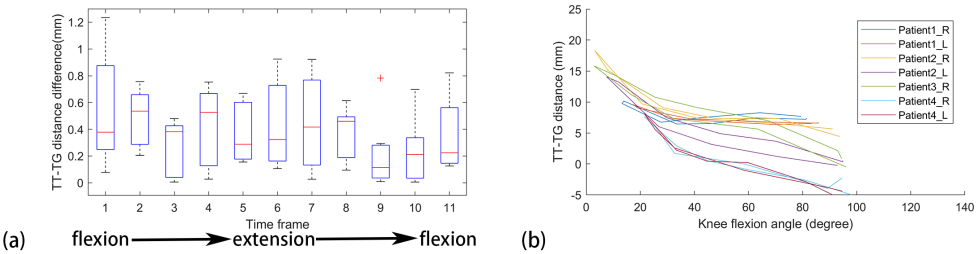


Figure 7.9: (a) The boxplot of TT-TG distance difference between the manual masks and masks from proposed method as a function of time frame for seven knees from four patients; (b) dynamic TT-TG distance as a function of flexion angle for 77 knee from four patients based on proposed method (lines indicate the dynamic TT-TG distance; there is only right knee information in patient 3 due to the fact that the left knee was out of the field of view (FOV))

7.4 Discussion

In this work, we developed a fully automatic workflow to quantify the knee joint movements using deep neural network for the bone segmentation. We proposed to have an implant detection and corresponding intensity correction to avoid the extreme intensity in order to ensure a robust and reliable segmentation performance. The clinical measurement (TT-TG distance) for patient with patella instability from 4DCT data using automatic segmentation was similar to that using manual segmentation.

The segmentation performance of unseen image with unexpected object is a challenge in the deep learning based segmentation. To solve this limitation, either adding the training samples to re-train the model or pre-process the input data is suggested. In knee joint CT image, the metal implant could be one of the unexpected objects when applying to the automatic knee joint segmentation. Our experiments showed that the extreme intensity value

caused by the implant could largely affect the segmentation performance no matter min-max or z-score normalization was applied. The poor performance by the min-max normalization (Figure 7.7b) is because the extreme intensity value largely altered the intensity distribution after the normalization. Although the z-score normalization could successfully segment the part of the bone, it failed to identify the region close to the implant (Figure 7.7c). This is because the neural network did not regard those area as the bone due to the pixels with unexpected intensity. As Figure 7.7d and Table 7.1 show, the proposed correction of the implant intensity could be an alternative to deal with the situation with implant.

Previous studies also aimed at automatic segmentation of bone in CT images in the purpose of diagnosis and 3D model construction. As Table 7.2 illustrates, both 3D U-Net and 3D V-Net based deep CNN methods achieved a high bone segmentation accuracy. Comparing the results to those reported in the literatures (SSM [7, 29], level set [28, 30], intensity and histogram [28] and Pseudo 3D U-Net [14]), the deep learning based segmentation method can achieve state-of-the-art accuracy based on the in-house dataset.

The robustness of the automatic landmark detection based on the automatic segmented bone masks was analysed in two ways. The first one was based on the comparison of the calculated landmarks between using automatic segmented and the manual segmented masks from 3D CT scan (Figure 7.8a). The error mainly comes from 1) the mismatch of these two segmented masks; 2) the bone shape model generation from the segmented mask. The second way to assess the robustness was the registration error of the TG, TT, the landmarks of PCRL (Figure 7.8b), which is mainly caused by 1) the difference of the bone shape models and 2) the robustness of the registration method. Although the result based on the automatic segmentation was not exactly the same as the one from manual segmentation, the resulting error was far smaller than the variance of inter-/intra- observation on landmarks, showing an adequate robustness of the propose method.

According to Chapter 6, the pure manual intra-observed variation of TT-TG distance was associated with a mean difference of 1.20 mm (SD = 1.04), while the inter-observed variation was 2.43 mm (SD=1.80 mm). As Figure 7.9a shows, the difference of TT-TG distance in the same time frame image between manual mask and the mask from proposed method were comparable. The TT-TG distance difference between these two masks over all dynamic time frames was 0.3±0.3 mm, which was much smaller than the difference of intra-/inter observation on TT-TG distance, and suggested the proposed automatic workflow was promising to replace manual labelling and therefore had the capacity to improve the effectiveness and efficiency of the clinical workflow.

From a clinical perspective, several findings can be noted from how the TT-TG distance depends on the flexion angle. Firstly, the TT-TG distance is decreasing as the knee flexion angle increases, which agrees with the findings of other authors [1, 4] and originates largely from an internal tibial rotation with higher flexion. Secondly, the TT-TG distance generally follows a slightly different track either going into flexion or going into extension (Figure 7.9b), which may originate from active quadriceps contraction and relaxation in different movement direction and possibly from visco-elastic properties of the soft tissue structures around the knee. Thirdly, the characteristics of TT-TG distance vs flexion angle appear to be consistent amongst patients, indicating it might be of clinical value.

Although promising results were shown in this study, several limitations remain. Firstly, we did not compare the conventional methods (SSM or atlas based method) to our proposed method with the same dataset. This is because the sample of the dataset in this study was limited to create a robust model that can perform a good segmentation. Instead, we compared the proposed method with above conventional methods on previous published literatures with similar topic. Secondly, the sample size in this study was obviously relatively small (4 patients (44 images) for training and 4 patients (44 images) for testing). However, this study was designed to test a potential workflow to automate the quantitative measurement of knee joint movement, so it is sufficient for this limited conceptual purpose. In the future, we will expand the training and testing groups to confirm the robustness of the proposed workflow. Thirdly, in the current workflow we relied on a 3D CT scan to determine the anatomical coordinate frame, which resulted in a higher dose for the patient. In a future refinement step, we plan to determine the coordinate system determination from the 4D scans.

7.5 Conclusions

In this work, an automatic bone segmentation method using deep neural network was proposed to apply on 4DCT images, and a fully automatic workflow to calculate the quantitative measurement of knee joint movement using the segmentation results was discussed. With this workflow, it is possible to efficiently quantify potential patella-femoral pathologies in a kinematic sense.

Funding

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 323091 awarded to N. Verdonschot. This work is also supported by China Exchange Programme (CEP) from Koninklijke Nederlandse Akademie Van

Wetenschappen (KNAW). Besides, this work is supported by Google Cloud Platform (GCP) research credits program.

Acknowledgements

We greatly thank Léon Driessen for the bone segmentation, Marieke Ploegmakers and Leo Kluijtmans for the manual measurements of TT-TG distance, meaningful discussion with Hans Dunning, and Yueyan Bian and Zhaoliang Lun for the helps on GPU.

References

[1] G. Seitlinger, G. Scheurecker, R. Högler, L. Labey, B. Innocenti, and S. Hofmann, "The position of the tibia tubercle in 0°–90° flexion: comparing patients with patella dislocation to healthy volunteers," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2396-2400, 2014, doi: 10.1007/s00167-014-3173-4.

[2] D. Forsberg, M. Lindblom, P. Quick, and H. k. Gauffin, "Quantitative analysis of the patellofemoral motion pattern using semi-automatic processing of 4D CT data," *International Journal of Computer Assisted Radiology and Surgery,* vol. 11, pp. 1731-1741, 2016, doi: 10.1007/s11548-016-1357-8.

[3] M. J. Tanaka, J. J. Elias, A. A. Williams, S. Demehri, and A. J. Cosgarea, "Characterization of patellar maltracking using dynamic kinematic CT imaging in patients with patellar instability," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 24, pp. 3634-3641, 2016, doi: 10.1007/s00167-016-4216-9.

[4] M. J. Tanaka, J. J. Elias, A. A. Williams, J. A. Carrino, and A. J. Cosgarea, "Correlation Between Changes in Tibial Tuberosity-Trochlear Groove Distance and Patellar Position During Active Knee Extension on Dynamic Kinematic Computed Tomographic Imaging," *Arthroscopy : the journal of arthroscopic & related surgery,* vol. 31, pp. 1748-1755, 2015, doi: 10.1016/j.arthro.2015.03.015.

[5] M. van Eijnatten, R. van Dijk, J. Dobbe, G. Streekstra, J. Koivisto, and J. Wolff, "CT image segmentation methods for bone used in medical additive manufacturing," *Medical Engineering and Physics,* vol. 51, pp. 6-16, 2018, doi: 10.1016/j.medengphy.2017.10.008.

[6] N. Sharma *et al.*, "Automated medical image segmentation techniques," *Journal of Medical Physics,* vol. 35, p. 3, 2010, doi: 10.4103/0971-6203.58777.

[7] C. Chu, C. Chen, L. Liu, and G. Zheng, "FACTS: Fully Automatic CT Segmentation of a Hip Joint," *Annals of Biomedical Engineering,* vol. 43, pp. 1247-1259, 2015, doi: 10.1007/s10439-014-1176-4.

[8] B. A. Besler, A. S. Michalski, N. D. Forkert, and S. K. Boyd, "Automatic Full Femur Segmentation from Computed Tomography Datasets Using an Atlas-Based Approach," in *Computational Methods and Clinical Applications in Musculoskeletal Imaging* vol. 3, ed, 2018, pp. 120-132.

[9] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, and R. Kijowski, "Deep convolutional neural network and 3D deformable approach for tissue segmentation

in musculoskeletal magnetic resonance imaging," *Magnetic Resonance in Medicine,* vol. 79, pp. 2379-2391, 2018, doi: 10.1002/mrm.26841.

[10] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis,* vol. 42, pp. 60-88, 2017, doi: 10.1016/j.media.2017.07.005.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, pp. 2481-2495, 2017, doi: 10.1109/TPAMI.2016.2644615.

[12] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation.

[13] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation.

[14] A. Klein, J. Warszawski, J. Hillengaß, and K. H. Maier-Hein, "Automatic bone segmentation in whole-body CT images," *International Journal of Computer Assisted Radiology and Surgery,* vol. 14, pp. 21-29, 2019, doi: 10.1007/s11548-018-1883-7.

[15] J.-y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *Computer Vision (ICCV), 2017 IEEE International Conference on,* pp. 2223-2232, 2017, doi: 10.1109/ICCV.2017.244.

[16] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017,* vol. 2017-Janua, pp. 5967-5976, 2017, doi: 10.1109/CVPR.2017.632.

[17] F. Ambellan, A. Tack, M. Ehlke, and S. Zachow, "Automated Segmentation of Knee Bone and Cartilage combining Statistical Shape Knowledge and Convolutional Neural Networks: Data from the Osteoarthritis Initiative," *Medical Image Analysis,* vol. 52, pp. 109-118, 2018, doi: 10.1016/j.media.2018.11.009.

[18] W. N. Scott, "Surgery of the Knee," *Insall & Scott Surgery of the Knee,* pp. 1-1504, 2014, doi: 9781455727896.

[19] L. Yao, N. Gai, and R. D. Boutin, "Axial scan orientation and the tibial tubercle-trochlear groove distance: Error analysis and correction," *American Journal of Roentgenology,* vol. 202, pp. 1291-1296, 2014, doi: 10.2214/AJR.13.11488.

[20] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping,* vol. 17, pp. 143-155, 2002, doi: 10.1002/hbm.10062.

[21] X. Xu *et al.*, "Simultaneous arteriole and venule segmentation with domain-specific loss function on a new public database," *Biomedical Optics Express,* vol. 9, p. 3153, 2018, doi: 10.1364/BOE.9.003153.

[22] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," pp. 1-11, 2016, doi: 10.1109/3DV.2016.79.

[23] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition.

[24] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs).

[25] R. C. Gonzalez and R. E. Woods, "Digital Image Processing (3rd Edition)," 2006.

[26] Q. Fang and D. A. Boas, "Tetrahedral mesh generation from volumetric binary and grayscale images," *Proceedings - 2009 IEEE International Symposium on*

Biomedical Imaging: From Nano to Macro, ISBI 2009, pp. 1142-1145, 2009, doi: 10.1109/ISBI.2009.5193259.

[27] A. Myronenko and X. Song, "Point Set Registration: Coherent Point Drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 2262-2275, 2010, doi: 10.1109/TPAMI.2010.46.

[28] J. A. Pérez-Carrasco, B. Acha, C. Suárez-Mejías, J. L. López-Guerra, and C. Serrano, "Joint segmentation of bones and muscles using an intensity and histogram-based energy minimization approach," *Computer Methods and Programs in Biomedicine,* vol. 156, pp. 85-95, 2018, doi: 10.1016/j.cmpb.2017.12.027.

[29] T. Albrecht, M. Lüthi, T. Gerig, and T. Vetter, "Posterior shape models," *Medical Image Analysis,* vol. 17, pp. 959-973, 2013, doi: 10.1016/j.media.2013.05.010.

[30] Y. Z. Y. Zhang, B. J. Matuszewski, L.-K. Shark, and C. J. Moore, "Medical Image Segmentation Using New Hybrid Level-Set Method," *2008 Fifth International Conference BioMedical Visualization: Information Visualization in Medical and Biomedical Informatics,* pp. 71-76, 2008, doi: 10.1109/MediVis.2008.12.

Chapter 8 General discussion and conclusion

As a part of the BioMechTools project, the overall goal of this thesis was to develop automatic and highly accurate segmentation of anatomical structures of the knee joint, and fully automated knee joint motion analysis providing clinical measures from 3D and 4D medical image data.

In medical research, there are various imaging modalities available to visualise the human body. This abundance of imaging tools are capable of providing a vast amount of information in itself already of great value to radiologists, but in order to be of use to the more and more advanced techniques in medicine, such as computer-assisted surgery, 3D pre-planning of surgeries and automatic detection of pathologies, a greater amount of processing and quantification of the imaging data are required.

The main task described in this thesis was the segmentation of knee joint anatomy in high detail and 3D/4D, allowing it as input for finite element models (FEM), analysis of orthopaedic kinetics, and the diagnosis of knee joint motion-related diseases. Among the available imaging modalities, MRI and CT were deemed most suitable, as they both give excellent contrast of the anatomical structures of interest and provide the best 3D resolution. MRI provides soft tissue contrast in a non-invasive manner; CT provides high spatiotemporal resolution and (in general) poses fewer motion restraints on the subject compared to MRI. As such, in this thesis, MRI data were used in **Chapter 2-5** to develop several methods of automated segmentation of trabecular bone, cortical bone and cartilage. 4DCT, i.e. dynamic 3DCT, was used to assess the development of current static diagnostic measures into dynamic ones. As Figure 8.1 illustrates, to automatically acquire the geometrical information of the knee joint structure, level set based, deep learning based and posterior shape model based methods were discussed in this thesis. In order to assess the knee joint dynamic information, a fully automatic workflow to calculate the clinical parameters was developed.

In this final chapter, the main contributions, limitations and recommendations for further steps towards clinical application and related fields of research will be discussed.

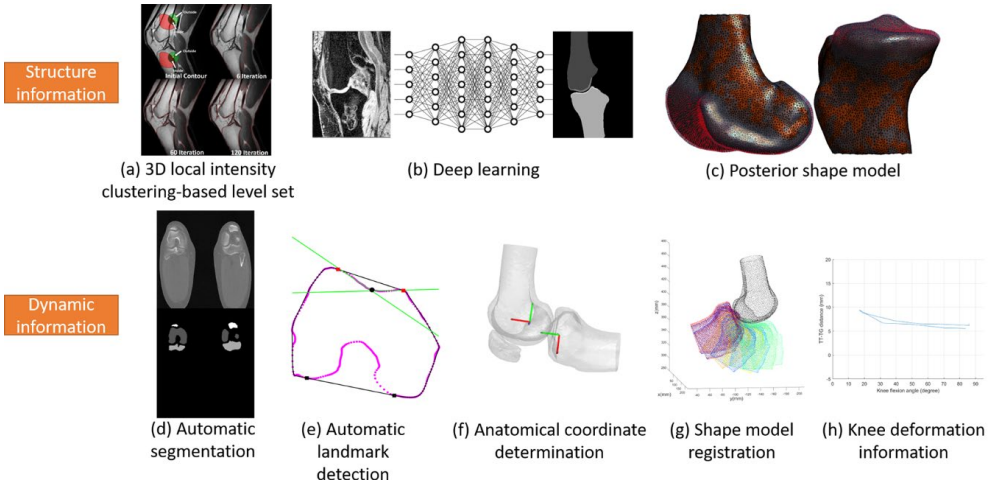


Figure 8.1: Key contents in this thesis

8.1 Automatic knee joint structure segmentation

8.1.1 Summary of the key results in automatic knee joint segmentation

MRI provides a vast number of different contrasts between the various soft tissues based on the manipulation of several atomic and molecular natural processes, i.e. T1 and T2 decay of excited hydrogen in both water and fat, diffusion of water through microstructures, the orientation of macromolecules to magnetic field etc. As such, registration/segmentation of MRI data poses different challenges and requirements depending on MRI contrast and the requirements/utilisations of the registration/segmentation results. In our experience, the time effort for an accurate segmentation of knee bones and cartilages by an experienced reader was around 2.5 hours. In this thesis, we developed two different methods of knee bone and cartilage segmentation for several frequently used types of MRI contrasts, proton density weighted, T1 and T2, of which T1 and T2 were acquired with fat suppression.

Chapter 2 describes a non-supervised 3D local intensity clustering based level set method to determine the trabecular bone areas of the knee despite any slow varying inhomogeneity, then the cortical bone was determined with sub pixel accuracy from intensity profiles along normal vectors of the trabecular bone shape. An average dice score of 0.96 was achieved in this study, together with segmentation time of 250 seconds.

Chapter 3 describes a deep learning approach for geometry extraction of the knee bone and cartilage from fat suppressed contrast data obtained from the publicly available grand challenge SK110 dataset [1]. The extensive comparisons of different networks showed that a simple U-Net architecture with fined tuned hyper-parameters could achieve state-of-the-art

results, but still required improvement in case of specific pathological data. We proposed a bone network including the adversarial loss to segment the resampled image data by the idea of enlarging the contextual information and ensuring the shape consistency for the bone segmentation, and then used a restoration network to further improve the segmentation accuracy of the result from bone network. The cartilage was segmented by a typical V-Net like network. It was the first time the total score of bone and cartilage segmentation was achieved greater than 76 using SKI10 validation dataset.

In **Chapter 4**, predictions of cartilage thickness and regularity based solely on the bone geometry was explored, with the goal of providing the diagnostic potential of cartilage status and construction of 3D knee joint model from CT images. Based on 99 fold data set (from SKI10 dataset [1], the accuracies of 0.64 mm and 0.59 mm for the femur and tibia cartilage thickness were achieved respectively in the tibiofemoral contact area. Although this accuracy may not always be high enough for diagnostic or modelling purposes, it does give more insight in how cartilage thickness is distributed and its local condition, instead of having to assume an average thickness as is typically using CT scans.

8.1.2 Deep learning in knee joint segmentation

Generally speaking, to date, deep learning based methods have surpassed other methods (active contour base methods, atlas based methods, statistical shape based methods etc.) in many medical image segmentation tasks [2, 3]. The reasons for this are not yet fully explained/understood. Image segmentation is the process of partitioning a digital image volume into multiple meaningful segments. In this sense, any segmentation tasks can be viewed as a direct dense classification task of each pixel/voxel in the input image volume. From a machine learning perspective, the classification task is simply to seek a suitable function where we feed in an image volume represented by a set of greyscale intensity pixels and get as output probability distributions over a set of classes that could separate the image volume into segments. The universal approximation theorem has proved that a shallow network structure with one hidden layer can approximate arbitrary (smooth) functions [4-6], and in this sense, already meets the requirements. Nonetheless, deep network structures have shown substantial benefits in finding a suitable function in a faster and more straightforward manner.

Deep neural networks provide exponentially growing complexity with the adding of more layers. Each layer adds more ability to define and isolate feature transformations, which can subsequently be trained with previous layers' output. This results in powerful networks capable of segmenting the most complex structures, which, however, also require more and more training data to avoid overfitting. Increasing amount of suggestions were proposed to

optimise and accelerate the training process (pre-trained model [7], normalization (batch normalization [8], instance normalization [9], group normalization [10]), suitable optimizers (Adam [11], Adagrad [12], Adadelta [13], RMSprop [14], SGD+momentum [15]), dropout [16], regularization technique (L1, L2), learning rate schedule [17], skip connections [18], suitable activations function (RELU [19], leaky RELU [20], ELU [21], etc.)). Moreover, the GPU is utilized to accelerate the computation. All these improvements in different areas help the development of the deep learning based method, which leads to its success not only in different segmentation challenges but also in speech recognition, object detection, etc.

In terms of discernible structures relevant for this thesis, the knee joint structure consists of bone, cartilage, menisci, ligament, and muscles. The first challenge is that the size of the image volume is too large to input to the network. In that case, the patch strategy might be an option to solve this issue. However, according to the experiments in **Chapter 3**, the range of the bone structure is larger than the maximum patch size, which failed to segment the regions of lower contrast or mis-segment the target-liked background areas. To solve this issue, we reduced the image volume into a reasonable size to allow the input patch to include more contextual information.

Similar to other deep learning based medical image applications [22], limited annotated data and the class imbalance of segmented objects are key reasons leading to the poor performance of knee joint segmentation, such as patient-specific osteophyte growth and unexpected artefacts [2]. To improve the performance, including a larger number of annotated samples is a straightforward way, but this task is tedious and expensive, in that experienced clinicians or people with professional training are required for this task. Data augmentation, e.g. flip, rotate, translate, mirror, etc., elastic deformation and addition of random noise, to increase the variability of the training dataset is the first option to solve this limitation. This approach has not yet lead to sufficient handling of unexpected artefacts, and methods such as active contours [23], statistical shape model [2], and conditional random fields [24] were designed successfully resolving this limitation in previous studies. In **Chapter 3**, we introduced the adversarial loss into the network training to solve this limitation within an end-to-end manner. As the bone is a compact structure, the adversarial loss is introduced as a shape regulator that ensures the shape consistency of the bone mask. Nevertheless, due to the small volume size of the cartilage, the cartilage segmentation cannot gain much benefits from the resampling and the adversarial loss, and thus, a typical V-Net was used in this study for cartilage segmentation in original resolution. Hence, although a main network with a single loss function would be preferable for simplistic and insight full determination of segmented structures, a more complex network employing a number of different loss functions is likely to deliver better results. On the other hand, increased

complexity, whether it be in terms of the number of layers or the variety in loss functions used will also increase the risk of the network being over-trained, thus requiring larger training data sets, and less likely to be of value to the network is trained for. Overall, when there are sufficient training data, a powerful GPU, a suitable network and a well-defined loss functions, deep learning should be the first option to implement fully automated segmentation of the knee joint.

8.1.3 Limitations and future challenges

Deep learning based medical image segmentation is a fast-developing field in recent years [22], and tremendous brilliant ideas have been proposed to improve the segmentation performance. Nevertheless, it is not easy to follow the work reported in the literatures and demonstrate the effectiveness of the modifications [3]. Isensee et al. proposed the nnU-net to set a baseline in six publicly available segmentation challenges. We attempted to set the baseline for the knee joint structure segmentation using nnU-Net as well, but the results were tested by a Nvidia 1080 Ti, which might lose some performance caused by a reduction of feature numbers due to lack of memory. A more complete comparison between the published networks and proposed network with a better GPU should be considered.

Also, the foundation of a good result using deep learning is a large labelled data set, which is a time and cost consuming task. A better tool to accelerate annotation task should be developed. Çiçek et al. proposed a semi-automatic workflow by annotating some slices in the volume to be segmented [25]. On the other hand, the non-supervised methods, level set, graph-cut, etc., are still good tools to assist or simplify the annotation procedure.

Moreover, the discussion of the knee joint segmentation was limited to the bone and cartilage in this thesis. To further investigate the segmentation of other knee joint parts, the annotation should be the first issues to be solved when deep learning based methods are utilized. Nonetheless, if the structures can hardly be discerned on the CT or MRI images (i.e. ligament), it is not a wise option to roughly annotate a bunch of data, then provide that as input to deep neural network and hope for surprising satisfied segmentation results. In that case, an improvement of the image quality should be considered rather than attempting all kinds of latest network that have been reported.

The recent development of the techniques, especially deep learning, indeed largely improves the performance of the medical image segmentation. As the methods become increasingly more sophisticated, evaluation of the quality and accuracy measures should also be considered. A high Dice coefficient score might also well conceal mis-segmentation of essential regions depending on the application of the segmentation. Therefore, a more clinical related evaluation measurement should be considered depending on the clinical application.

8.2 Knee joint motion analysis

8.2.1 Summary of the key results in the automatic knee joint motion analysis

Sensitivity analysis (**Chapter 5**) showed that the small disagreement between the mechanical axis of the lower extremity and craniocaudal axis of the CT scanner resulted in a significant impact on the tibial tubercle – trochlear groove (TT-TG) distance calculation. This simulated sensitivity analysis provided quantitative error information (1 mm per degree) on the effect of the disagreement to the TT-TG distance. Since most of the evaluation system to diagnose the knee joint status requires the alignment of mechanical axis and the craniocaudal axis of CT scanner, a necessity of pre-correction of the axial plane before any manual measurements of the knee joint relationship is required.

To avoid the variability introduced by manual pre-alignment of above disagreement [26, 27], an automatic determination of the anatomical coordinate frame of knee bones (**Chapter 6**) was designed to automate the correction of disagreement of the planes. Furthermore, according to the experiments in **Chapter 5** and **Chapter 6**, the differences in intra- and inter-observations of the key landmarks of TT-TG distance was substantial, i.e. differences were in the order of the scale ranging between different diagnostic outcomes. Automation of this procedure on landmarks detection led to more robust results, confirmed by TT-TG measurements in a patient group, where the TT-TG outcomes corresponded to the patient pathology. However, the workflow in **Chapter 6** still required a large manual workload for the knee joint segmentation.

To remove the need of manual input, a deep learning segmentation network for 3D and 4D CT data was proposed (**Chapter 7**). A mean DSC for femur, tibia, and patella of 0.98 ± 0.01, 0.97 ± 0.02, and 0.97 ± 0.01, respectively was achieved. The difference of TT-TG distance based on automatic segmentation and manual segmentation (0.3 ± 0.3 mm) was a factor of ten lower compared to TT-TG distances solely obtained from manual measurement 1.2 ± 1.0 mm for intra-observer error and 2.4 ± 1.8 mm for inter-observer error.

Overall, the developed fully automatic workflow has been proved to be capable of successful assessment of knee measurement using the TT-TG distance in a dynamic fashion within an in-house dataset. The total time required for this process for one dataset (both left and right knee, 24 knees) is around 0.5 h, which includes automatic segmentation of the 3D and 4D segmentation (< 20 min), generation of the 3D model based on the segmented masks (< 1 min), anatomical coordinate frame determination (< 2 min), landmarks detection (< 1 min), the registration among the 3D models (< 4 min), generation of the dynamic TT-TG distance (< 1 min).

8.2.2 Is the automatic workflow for knee joint motion measurement effective?

Although the workflow for knee joint motion analysis presented in this thesis was fully automatic, improvement of accuracy and reproducibility still needs to be confirmed in a clinical setting and on a larger scale. Quantification of motion and kinetics is a deeply wrought topic of research in the biomechanical field. State-of-the-art methods to acquire the knee joint dynamic information range from skin-mounted marker tracking, quasi dynamic assessment in CT or MRI [28, 29] and highly accurate measurements from (stereo) fluoroscopic systems [30-32]. Where each of these methods surpasses the others in terms of resolution and accuracy, they lack in terms of cost, invasiveness and laboriousness. As a part of the BioMechTools project, an A-mode probe ultrasound marker system was designed, attempting to combine the accuracy of CT, MRI, fluoroscopy, with the low invasiveness, lower cost and broader range of motion that skin marker methods offer. This system showed very promising results, but like all new techniques, requires more study involving validation on a larger scale [33-35] and was not available for the experiments of this thesis. Nevertheless, the none-invasive, accurate real-time tracking of joint motion is still in an unmet need. To make a baseline of knee joint movement tracking within enough real-time accuracy, the 4DCT was used in this study, in that the bone can be imaged clearly within a CT device [27].

The proposed workflow in this thesis effectively extended the knee joint movement measurement ranging from 0° to 90° and successfully calculated the corresponding required landmarks in non-fully extended knee joint status using the registration techniques. Previously, even though the 4DCT was capable of capturing the knee joint movement, the tracking capability was limited to 5° to 30°, because the required landmarks were not included in the limited field of view (FOV) [13]. Moreover, since most of the evaluation criterions of knee joint status were based on a fully extended knee joint, it was not easy to identify the correct position of the landmarks in a flexed knee joint according to the clinicians' experience. Hence, the proposed workflow not only shows a definite improvement of detected range but also ensures the robustness of the landmark detection.

Furthermore, as analysed in **Chapter 5**, the disagreement between the anatomical axial plane and the CT axial plane was one of the primary sources of error in knee joint status measurement. The proposed automatic determination of the anatomical coordinate frame provided an anatomical axial plane to automatically determine the suitable axial plane to calculate the required landmarks for the measurements, which saves the workloads from adjusting the disagreement.

When comparing the traditional manual measurement, the inter- and intra- observation of the TT-TG distance was up to 1.2±1.0 mm and 2.4±1.8 mm, while the proposed workflow

could provide repeatable and robust results (0.3 ± 0.3 mm compared to the results based on manual segmentation) and successfully predict the knee joint status based on the in-house dataset.

To test the clinical value of proposed workflow, **Chapter 6** describes the quantification of dynamic TT-TG distance, revealing hysteresis associated with the direction of flexion, i.e. flexion to extension vs extension to flexion. This hysteresis in TT-TG distance indirectly demonstrated the added value of dynamic assessment, since the static imaging is unable to take this into account. Whether this new parameter has any clinical meaning (e.g. pointing towards a more lax or unstable knee joint) remains to be demonstrated in future clinical studies.

8.2.3 Limitations and future challenges

Although the dynamic TT-TG distance and PC-TG distance showed the promising results with inclusion of the dynamic information in this thesis and [27, 36, 37], a larger patient group should be included to summarise the clinical finding as a clinical standard. Also, it is essential to note that the diagnosis of the knee joint is not only based on the TT-TG distance and PC-TG distance, and that more related parameters, such as patellar tilt, bisect offset, etc. [38] should be included in the automatic workflow in order to have a deeper understanding of knee joint status. Another improvement point lies on the image sample frame rate. As a pilot study, the frame rate to capture the 3D data was set to 11 frames in 13 seconds in order to reduce the motion artefact. In future studies, a higher sample frame rate should be applied to test the robustness of the workflow.

Moreover, a full CT scan was required in the current workflow in order to capture the accurate landmarks, which increased the total amount of the dosage. The introduction of the automatic segmentation based on MRI images (**Chapter 2** and **Chapter 3**) suggests a promising solution to have a full MRI scan instead. Nevertheless, one should be aware of the disagreement of TT-TG distance between CT and MRI measurements reported in the literature [39, 40]. Furthermore, due to the extensibility of the proposed workflow, it can further apply to non-invasive image techniques (4D MRI [41], A-mode based ultrasound [33, 35]). However, regarding 4DMRI, several issues should be researched, such as the limited range of knee joint motion, pain caused by the repeated movement, the insufficient image quality for a successful registration. Regarding the application of A-mode ultrasound based tracking devices to measure bone kinematics, the accuracy for the bone position detection and a successful registration should be considered in future studies.

8.3 Recommendations for next steps towards clinical application

Even though significant improvements have been achieved to automatically segment the knee bone and cartilage, and automatically quantify the knee joint movement in an effective way, several steps are still required before reaching the ambitious goal of routinely using the proposed workflow.

The applicability of proposed deep learning based knee joint segmentation methods to larger image sources remains to be investigated. Due to different focuses on tissue structures, different MRI sequences might be adopted, leading to the different intensity distribution of the same knee joint structure. The straightforward approach is to create the corresponding segmented masks to train the network again, but it will cost a lot of time to annotate the masks. Another possible approach is to perform an image-to-image translation on the new source image, in other words, to train a network that translates an image from source domain A to a target domain B in an absence of paired examples. Recently, Zhu et al. proposed a cycle-consistent adversarial network to solve this kind of problem [42]. Some attempts have been applied to medical field (from CT to MRI [43, 44], MRI to CT [43, 45, 46], among MRI sequences [47], CT to PET [48, 49]). Further investigation on the generalizability of the proposed segmentation workflow is required.

The automatic workflow for knee joint motion analysis presented in this thesis is a part of the BioMechTools project, where a successful example of fully automatic quantification of knee joint movement was proven. However, to reproduce the automatic workflow in clinical practices, more dataset is required to confirm the robustness of the proposed workflow and the clinical findings.

Moreover, although all the different steps of automatic segmentation and motion analysis of the knee joint were successfully linked in **Chapter 7** and it proved the possibility to automatically generate the clinical related parameters to describe the status of the knee joint dynamics, further efforts are still required to put them to the final user of the tools, i.e. the clinicians or surgeons without a technical background. In order to have a clear path to guide the required steps towards the clinical application, we summarised the status of the developed software based on technology readiness level (TRL) [50] in Table 8.1. Based on Table 8.1, it is clear that future works should focus on testing the findings of the thesis on a large scale and making the tools and graphical interfaces simpler and more user-friendly in order to guarantee that clinical personnel can use the software in an efficient way.

Table 8.1: Developed software and their corresponding technology readiness level

Chapter	Software	Language	TRL	Steps towards to clinical use
2	3d local intensity clustering based level set and its post-processing	Matlab	3	1. A simple GUI should be developed for the user without a technical background.
3	Deep learning based knee bone and cartilage segmentation	Python	5	1. The medical image and the corresponding annotation are prepared; 2. A simple GUI should be developed for the user without a technical background.
4	Posterior shape model for knee cartilage estimation	Matlab	2	1. A better dataset should be collected (isotropic scan and label with the KL score); 2. Proof of concept of the estimation of the cartilage from the CT images; 3. Evaluation of the error tolerance for the application, as there must be some errors from this method.
6	Detection of key landmarks of bone	Matlab	3	1. A simple GUI should be developed for the user without a technical background
	Determination of anatomical coordinate frame	Matlab	4	
	Registration techniques	Matlab	4	
	Calculation of TT-TG / TT-PC distance	Matlab	4	

8.4 Conclusions

To conclude, we presented several technical developments in the field of automatic segmentation tools on knee joint images, and a complete workflow to automatically quantify the knee joint motion (from images to quantitative description of the knee joint movement). Although further technical developments will be needed in the future, the segmentation tools presented in this thesis could facilitate researchers, clinicians and surgeons in the

biomechanical-orthopaedics field to build a musculoskeletal model, perform a clinical diagnosis or execute a pre-operative planning. Furthermore, the described workflow of automatic knee joint motion quantification in dynamic images are likely to present added clinical value, as it introduces new knee joint motion information. With these tools, we hope to contribute to an improvement of care of patients suffering from pathologies to the knee joint.

References

[1] T. Heimann and B. Morrison, "Segmentation of knee images: A grand challenge," 2010.

[2] F. Ambellan, A. Tack, M. Ehlke, and S. Zachow, "Automated Segmentation of Knee Bone and Cartilage combining Statistical Shape Knowledge and Convolutional Neural Networks: Data from the Osteoarthritis Initiative," *Medical Image Analysis,* vol. 52, pp. 109-118, 2018, doi: 10.1016/j.media.2018.11.009.

[3] F. Isensee, J. Petersen, S. A. A. Kohl, P. F. Jäger, and K. H. Maier-Hein. nnU-Net: Breaking the Spell on Successful Medical Image Segmentation.

[4] G. Cybenkot, "Approximation by Superpositions of a Sigmoidal Function," *Math. Control Signals Systems,* vol. 2, pp. 303-314, 1989.

[5] B. Csáji, "Approximation with artificial neural networks," *MSc. thesis,* p. 45, 2001, doi: 10.1.1.101.2647.

[6] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks,* vol. 4, pp. 251-257, 1991, doi: 10.1016/0893-6080(91)90009-T.

[7] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A Survey on Deep Transfer Learning," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2018arXiv180801974T

[8] S. Ioffe. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.

[9] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2016arXiv160708022U

[10] Y. Wu and K. He, "Group Normalization," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2018arXiv180308494W

[11] D. P. Kingma and J. L. Ba. Adam: A Method for Stochastic Optimization.

[12] J. Duchi, E. Hazan, and Y. Singer, "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," *Journal of Machine Learning Research,* vol. 12, pp. 2121-2159, 2011.

[13] M. D. Zeiler, "ADADELTA: An Adaptive Learning Rate Method," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2012arXiv1212.5701Z

[14] A. Graves, "Generating Sequences With Recurrent Neural Networks," *arXiv e-prints*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2013arXiv1308.0850G

[15] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," presented at the Proceedings of the 30th

International Conference on International Conference on Machine Learning - Volume 28, Atlanta, GA, USA, 2013.

[16] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research,* vol. 15, no. 1, pp. 1929-1958, 2014.

[17] F. Isensee *et al.* nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation.

[18] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation.

[19] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," presented at the Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 2010.

[20] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolutional Network," 2015.

[21] D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs).

[22] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges," *Journal of Digital Imaging,* vol. 3, pp. 582-596, 2019, doi: 10.1007/s10278-019-00227-x.

[23] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, and R. Kijowski, "Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging," *Magnetic Resonance in Medicine,* vol. 79, pp. 2379-2391, 2018, doi: 10.1002/mrm.26841.

[24] Z. Zhou, G. Zhao, R. Kijowski, and F. Liu, "Deep convolutional neural network for segmentation of knee joint anatomy," *Magnetic Resonance in Medicine,* vol. 80, pp. 2759-2770, 2018, doi: 10.1002/mrm.27229.

[25] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3D U-net: Learning dense volumetric segmentation from sparse annotation.

[26] L. Yao, N. Gai, and R. D. Boutin, "Axial scan orientation and the tibial tubercle-trochlear groove distance: Error analysis and correction," *American Journal of Roentgenology,* vol. 202, pp. 1291-1296, 2014, doi: 10.2214/AJR.13.11488.

[27] M. J. Tanaka, J. J. Elias, A. A. Williams, J. A. Carrino, and A. J. Cosgarea, "Correlation Between Changes in Tibial Tuberosity-Trochlear Groove Distance and Patellar Position During Active Knee Extension on Dynamic Kinematic Computed Tomographic Imaging," *Arthroscopy : the journal of arthroscopic & related surgery,* vol. 31, pp. 1748-1755, 2015, doi: 10.1016/j.arthro.2015.03.015.

[28] G. Seitlinger, G. Scheurecker, R. Högler, L. Labey, B. Innocenti, and S. Hofmann, "The position of the tibia tubercle in 0°–90° flexion: comparing patients with patella dislocation to healthy volunteers," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2396-2400, 2014, doi: 10.1007/s00167-014-3173-4.

[29] K. Izadpanah *et al.*, "Influence of knee flexion angle and weight bearing on the Tibial Tuberosity-Trochlear Groove (TTTG) distance for evaluation of patellofemoral alignment," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 22, pp. 2655-2661, 2014, doi: 10.1007/s00167-013-2537-5.

[30] D. L. Benoit, D. K. Ramsey, M. Lamontagne, L. Xu, P. Wretenberg, and P. Renström, "Effect of skin movement artifact on knee kinematics during gait and cutting motions measured in vivo," *Gait and Posture,* vol. 24, pp. 152-164, 2006, doi: 10.1016/j.gaitpost.2005.04.012.

[31] S. Guan, H. Gray, F. Keynejad, and M. Pandy, "Mobile Biplane X-ray Imaging System for Measuring 3D Dynamic Joint Motion during Overground Gait," *IEEE Transactions on Medical Imaging,* vol. 0062, pp. 1-1, 2015, doi: 10.1109/TMI.2015.2473168.

[32] G. Li, S. K. Van de Velde, and J. T. Bingham, "Validation of a non-invasive fluoroscopic imaging technique for the measurement of dynamic knee joint motion," *Journal of Biomechanics,* vol. 41, pp. 1616-1622, 2008, doi: 10.1016/j.jbiomech.2008.01.034.

[33] K. Niu *et al.*, "In situ comparison of A-mode ultrasound tracking system and skin-mounted markers for measuring kinematics of the lower extremity," *Journal of Biomechanics,* vol. 72, pp. 134-143, 2018, doi: 10.1016/j.jbiomech.2018.03.007.

[34] K. Niu, J. Homminga, V. Sluiter, A. Sprengers, and N. Verdonschot, "Measuring relative positions and orientations of the tibia with respect to the femur using one-channel 3D-tracked A-mode ultrasound tracking system: A cadaveric study," *Medical Engineering and Physics,* vol. 57, pp. 61-68, 2018, doi: 10.1016/j.medengphy.2018.04.015.

[35] K. Niu, V. Sluiter, J. Homminga, A. Sprengers, and N. Verdonschot, "A Novel Ultrasound-Based Lower Extremity Motion Tracking System," ed, 2018, pp. 131-142.

[36] M. J. Tanaka, J. J. Elias, A. A. Williams, S. Demehri, and A. J. Cosgarea, "Characterization of patellar maltracking using dynamic kinematic CT imaging in patients with patellar instability," *Knee Surgery, Sports Traumatology, Arthroscopy,* vol. 24, pp. 3634-3641, 2016, doi: 10.1007/s00167-016-4216-9.

[37] D. Forsberg, M. Lindblom, P. Quick, and H. k. Gauffin, "Quantitative analysis of the patellofemoral motion pattern using semi-automatic processing of 4D CT data," *International Journal of Computer Assisted Radiology and Surgery,* vol. 11, pp. 1731-1741, 2016, doi: 10.1007/s11548-016-1357-8.

[38] W. N. Scott, "Surgery of the Knee," *Insall & Scott Surgery of the Knee,* pp. 1-1504, 2014, doi: 9781455727896.

[39] B. B. Hinckel *et al.*, "Are the osseous and tendinous-cartilaginous tibial tuberosity-trochlear groove distances the same on CT and MRI?," *Skeletal Radiology,* vol. 44, pp. 1085-1093, 2015, doi: 10.1007/s00256-015-2118-4.

[40] C. L. Camp *et al.*, "CT and MRI measurements of tibial tubercle-trochlear groove distances are not equivalent in patients with patellar instability," *American Journal of Sports Medicine,* vol. 41, pp. 1835-1840, 2013, doi: 10.1177/0363546513484895.

[41] V. Mazzoli *et al.*, "Accelerated 4D self-gated MRI of tibiofemoral kinematics," *NMR in Biomedicine,* vol. 30, pp. 1-11, 2017, doi: 10.1002/nbm.3791.

[42] J.-y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *Computer Vision (ICCV), 2017 IEEE International Conference on,* pp. 2223-2232, 2017, doi: 10.1109/ICCV.2017.244.

[43] Z. Zhang, L. Yang, and Y. Zheng, "Translating and Segmenting Multimodal Medical Volumes with Cycle- and Shape-Consistency Generative Adversarial Network," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 9242-9251, 2018, doi: 10.1109/CVPR.2018.00963.

[44] C.-B. Jin *et al.*, "Deep CT to MR Synthesis Using Paired and Unpaired Data," *Sensors,* vol. 19, p. 2361, 2019, doi: 10.3390/s19102361.

[45] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. Išgum, "Deep MR to CT synthesis using unpaired data," *MICCAI 2017 Workshop on Simulation and Synthesis in Medical Imaging,* vol. 10557 LNCS, pp. 14-23, 2017.

[46] Y. Hiasa *et al.*, "Cross-modality image synthesis from unpaired data using cyclegan: Effects of gradient consistency loss and training data size," *MICCAI 2018 Workshop on Simulation and Synthesis in Medical Imaging,* vol. 11037 LNCS, pp. 31-41, 2018, doi: 10.1007/978-3-030-00536-8_4.

[47] S. Ul *et al.*, "Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks," *IEEE Transactions on Medical Imaging,* vol. 90, pp. 1-1, 2019.

[48] A. Ben-Cohen, E. Klang, S. P. Raskin, M. M. Amitai, and H. Greenspan, "Virtual PET Images from CT Data Using Deep Convolutional Networks: Initial Results," in *MICCAI 2017 Workshop on Simulation and Synthesis in Medical Imaging* vol. 10557 LNCS, ed, 2017, pp. 49-57.

[49] A. Ben-Cohen *et al.*, "Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection," *Engineering Applications of Artificial Intelligence,* vol. 78, pp. 186-194, 2019, doi: 10.1016/j.engappai.2018.11.013.

[50] NASA, "Technology Readiness Level Definitions," p. 1, 2019.

List of Publications

Journal contributions

Hao Chen, André Sprengers, Yan Kang, Nico Verdonschot. Automated segmentation of trabecular and cortical bone from proton density weighted MRI of the knee. *Medical & Biological Engineering & Computing*. 2019;57(5):1015–1027. doi:10.1007/s11517-018-1936-7.

Hao Chen, Tao Tan, Yan Kang, Nico Verdonschot, André Sprengers. Knee bone and cartilage segmentation based on 3D deep neural network using adversarial loss for prior shape constraint. *Plan to Submit to Computerized Medical Imaging and Graphics.*

Hao Chen, André Sprengers, Tao Tan, Nico Verdonschot. Cartilage Estimation based on Bone Geometry using Posterior Shape Model. *Submitted to Journal of Biomechanics.*

Hao Chen, André Sprengers, Leo Kluijtmans, Sebastiaan van de Groes, Nico Verdonschot. Axial orientation in CT scanner is of large influence on tibial tubercle-trochlear groove distance. *Submitted to Knee Surgery, Sports Traumatology, Arthroscopy.*

Hao Chen, Leo Kluijtmans, Hans Dunning, Sebastiaan van de Groes, André Sprengers, Nico Verdonschot. A robust and semi-automatic quantitative measurement of patellofemoral instability based on 4D computed tomography. *Conditionally accepted by Medical Engineering & Physics.*

Hao Chen, Tao Tan, Sebastiaan van de Groes, Yan Kang, André Sprengers, Nico Verdonschot. Fully automatic quantitative measurement of knee joint movement based on 4DCT using deep neural network for automatic segmentation of bones. *Submitted to Journal of Orthopaedics.*

Shengyu Fan, Yueyan Bian, **Hao Chen**, Yan Kang, Tao Tan, Qi Yang, Danny JJ Wang Unsupervised Cerebrovascular Segmentation of TOF-MRA Images based on Deep Neural Network and Hidden Markov Random Field Model. *Frontiers in Neuroinformatics*. 2020; 13: 77. https://doi.org/10.3389/fninf.2019.00077.

Conference contributions

Hao Chen, André Sprengers, Yan Kang, Hong Li, Nico Verdonschot. A localized level set method for automatic segmentation of bone MRI data without use of prior knowledge.

ESB 2016, Lyon, France. *22nd Congress of the European Society of Biomechanics*. European society of biomechanics.

Bruna Vaughan, Franklin Kuraoka Oda, **Hao Chen**, Nikita Kruis, André Sprengers, Nico Verdonschot. Inter and intra observer variability in segmentation of long bones related to the mean curvature. ESB 2016, Lyon, France. *22nd Congress of the European Society of Biomechanics*. European society of biomechanics.

Hao Chen, André Sprengers, Max Bakker, Sebastiaan van de Groes, Leo Kluijtmans, Yan Kang, Nico Verdonschot. Automatic quantitative assessment for patellofemoral joint within dynamic TT-TG distance of 4D CT data. WCB 2018, Dublin, Ireland. *8th World Congress of Biomechanics*.

Hao Chen, André Sprengers, Yan Kang, Nico Verdonschot. Automated quantitative assessment of cartilage thickness using scale space analysis. WCB 2018, Dublin, Ireland. *8th World Congress of Biomechanics*.

Hao Chen, André Sprengers, Max Bakker, Sebastiaan van de Groes, Leo Kluijtmans, Yan Kang, Nico Verdonschot. Automatic quantitative assessment for patellofemoral joint within dynamic TT-TG distance of 4D CT data. *ORS 2019 Annual Meeting*, Austin, USA. Orthopaedic Research Society.

Acknowledgement

It all started one year after I completed a master's degree of biomedical engineering from Northeastern University, China. I was an algorithm developer at that time in Shanghai United Imaging Healthcare Co., Ltd. One day when I took the bus to work, I was thinking was I decide to give up the pursuit of PhD study? I truly remembered the answer was a hard no, and meanwhile, I truly realized how difficult it would be for a person who had started to work for a company more than years in China to seek an Ph.D. position abroad. Luckily, prof. Nico Verdonschot and prof. Yan Kang jointly applied a joint research project based on scientific cooperation between China and the Netherlands in 2015. Prof. Kang and Han J. W. van Triest recommended me to apply the position and this began my Ph.D. story.

When I look back the Ph.D. journey I have experienced since 2015, I would like to express my sincere gratitude to all the people who helped me in the past four and a half years.

First of all, I would like to express my deepest gratitude to my supervisor Nico Verdonschot and Yan Kang, and co-supervisors André Sprengers. Without your kindness help during the past years, I would not have been able to complete my thesis.

Dear Nico, thanks for giving me the chance to join the BioMechTools group. Thanks for your guidance for my research and effort in correcting my paper and thesis. You are a tireless researcher who is always full of all kinds of meeting and I am really appreciated that you can always find some space for me to discuss the progress especially after you became the scientific director of the technical medical centre and when I was in China. I still remembered in 2016 when you travelled to China for a conference, you spent all your spare time to discuss my first paper, rather than travelled around the city. You taught me that scientific researches are not only fantastic good results. Any rational results are worth to report and always remind me to move forward. I am very grateful for having you as my supervisor.

Dear prof. Kang, I have been your student since 2007, from undergraduate to PhD. Your passion for work and enthusiasm for computer aided diagnosis (CAD) is encouraging. It is you that let me find my interest is also on the CAD. Until now, most of my work focus are still related to CAD. Also, thanks for recommending me to apply the PhD position, which provide me a chance to realize my PhD dream. Thanks for giving me the chance to join the clinical business unit, Neusoft Medical System, which let me learn a lot of experience of how to translate research results to actual medical cases. I wish all the best to your new career in Shenzhen Technology University, and hope we can have a further collaboration in the AI based medical image application.

Dear André, words cannot express my gratitude and appreciation to you. Thanks for always having time for a chat, no matter it is a life related discussion or a scientific discussion. I am really appreciated your encouragement in the darkest moment of my Ph.D. It was you that trusted me and gave me the energy to move forward. I still remembered the first day we met, you said I can always contact you if I need helps, and I did, no matter the "fake fine issue" or all kinds of weird scientific questions. You are always so kind to help me, and you solved the problems I had every time. Every time! I am so lucky that having you as my co-supervisor. "Hao, it's gonna be fine!" will always encourage and support me when I face to the unknown challenge in the future.

Secondly, I want to express my gratitude to all the committee members prof. dr.ir. C. H. Slump, prof. dr. ir. G. J. Verkerke, prof. dr. ir. B.M. ter Haar Romeny, prof. dr. ir. A. J. Nederveen and dr. T. van Tienen for spending time to read my thesis, for allowing me to give defence, and for taking part in the defence ceremony.

Dear Tao, thanks for your helps in all the research discussions, especially for the deep learning researches. It is you that actually bring me into the deep learning world. I still remember in 2017 I thought deep learning was too far away for me. With your helps, I started from U-Net 2D, then U-Net 3D, to now all kinds of different networks. I wish we can keep discussing the scientific research in the future and wish all the best to your new position in GE.

I want to express special thanks to my (former) colleagues in the department of biomechanics. Firstly, I want to thank Lianne and Jeanine. You make me feel like at home. Although it was my first time to work and study abroad, with your helps, I didn't have any troubles to live and work in Enschede. I felt lucky that we have you as our secretaries in BW group. I would also like to thank Bart, the big boss of our group. Thanks for your helps on my work permit in the department. Nikolai, thanks for all your IT supports, from software installation to hardware replacement. Vincenzo, Kenan, Iris, Riza, Hans and Corine, I am so fortunate to have you as my office mates during the past years. We have so much fun in W213 and W107. I hope we can do the dinner party once more! I also want to thank the lovely colleagues in BW group. Here I express my gratitude to the (former) BW members: Jasper, Victor, Kostas, Verros, Wouter, Haarman, Halfwerk, Dorien, Massimo, Jiena, Allan, Simone, René, Ravi, Hielke, Alexander. Thank you all once again for giving me such a nice time in the group.

I would also like to express my gratitude to all BioMechTools members and colleagues in Nijmegen. Léon, thanks for your help on the manual segmentation! To be honest, sometimes I am afraid that you are scare to hear my name, cause every time I contact you, you need to

use Mimics to do the tedious works again. I am really appreciated what you did for the projects, and it saved me a lot of time. Dear Max, Leo, Hans and Sebastian, our collaboration resulted in Chapter 5 and Chapter 6. I learnt a lot of biomechanics knowledge from you. Valentina, thanks for scanning MRI image for all volunteers for my research. Hamid, thanks for your helps and explanation of the biomechanical modelling. I am really enjoying the discussion of the statistical shape model with you and I hope we can collaborate further for the publication. Macro, thanks for your detailed explanation of the anatomical coordinate and sharing of state-of-the-art techniques, and also the organized codes on gitlab, since then I started to keep my codes on gitlab too. I also want to thank Kaj, Dennis, Richard, Bruna, Nikita, and Franklin. Your kindness helps made me feel like at home when I was in Nijmegen.

I am grateful to have wonderful Chinese friends in Enschede. Hao, I spent unforgettable time with you in the past years, basketball, wangzherongyao, shopping in Gronau, watching TV drama series, “in the name of the people”, etc. Best wishes to your future career both in Europe and China. Kenan, I am so lucky have you as my colleague, classmate and good friend. You helped me a lot no matter in life or research. I am really enjoying the time with you, watching Douyu, bilibili, also discussing the research. Best wishes to your newborn. XiaoCai and FanGe, we share a lot of unforgettable memories in the Netherlands, PokémonGo, PS4, travelling around Eindhoven and Leiden, etc. Best wishes to your new career. Many thanks to Xiaohua, Yidan, Jingya, Zhantao, Li Qian, Gao Lan, Jinhui, Mengdi, Youwen, Xiaolin, Wenlong, Hongxu and all others. I am also grateful to have so many BMIE friends in Europe, Han, Fenghua, Jingqi, Siyu, Xiaofu, and Dr. Qi. It was my fortunate to have you together for all the parties, trips, laughs and joys

I will use the Chinese for the final thanks to my family:

首先，我非常感谢父母对我的支持，没有你们默默的支持，我是无法完成博士论文的；其次，我非常感谢我的妻子李佳，过去的四年半你付出了太多，没有你作为坚实的后盾和对嘟嘟的悉心照顾，我是无法专心进行博士研究的；同时也非常感谢我的岳父岳母，谢谢你们对我工作的理解和支持；最后，我感谢我的女儿嘟嘟，是你让我有足够的动力成为更好的榜样。

December 2019

Hao

About the author

Hao Chen was born on 26th October 1988, in Zhaoqing, Guandong Province, China. He has obtained a Master degree (2013) and a Bachelor degree (2011) on biomedical engineering from Biomedical and Information Engineering School, Northeastern University, Shenyang, China. He worked on the topic of 3D CT image reconstruction and its acceleration for both his master and bachelor thesis, under the supervision of prof. Yan Kang and dr. Dan Xia.

Afterwards, he joined Shanghai United Imaging Healthcare Co., Ltd to research the image reconstruction based on incomplete projections of digital breast tomosynthesis (DBT) and image enhancement of digital radiography (DR).

In April 2015, he started his Ph.D. on a joint research project with the support of KNAW grant and ERC advanced grant of "BioMechTools" under the supervision of Prof. Nico Verdonschot at Department of Biomechanical Engineering of the University of Twente, Enschede, The Netherlands and prof. Yan Kang at College of Medicine and Biological Information Engineering of Northeastern University, China and dr. André Sprengers at Department of Biomedical Engineering & Physics of Amsterdam University Medical Center, The Netherlands. The results of Hao's PhD are described in this thesis.

Since April 2019, he has started as an mathware engineer in Sioux to work on the topics of application of artificial Intelligence / machine Learning.

ISBN: 978-90-365-4947-9