

**AN AUTOMATED, DEEP LEARNING APPROACH TO
SYSTEMATICALLY & SEQUENTIALLY DERIVE THREE-DIMENSIONAL
KNEE KINEMATICS DIRECTLY FROM TWO-DIMENSIONAL
FLUOROSCOPIC VIDEO**

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Viet Dung Nguyen

August 2023

Copyright © 2023 by Viet Dung Nguyen

All rights reserved.

Dedication

To my parents, Lang Nguyen & Oanh Nguyen

To my grandfathers, in memories

To my grandmothers

To my sister

Acknowledgements

First, I would like to thank Dr. Richard Komistek for his exceptional guidance, encouragement, and mentorship throughout my time at the Center for Musculoskeletal Research (CMR).

I am deeply grateful for the immense support and appreciation I received from Dr. Michael LaCour. Our time working together at CMR was filled with countless memories, lessons, and stories that I will cherish forever.

I want to acknowledge the CMR lab for a vast dataset source and state-of-the-art equipment in enabling me to complete my research successfully.

I extend my gratitude to my colleagues at CMR for their exceptional support during my PhD journey, particularly Dr. Brad Meccia, Garrett Dessinger, Jarrod Nachtrab, Lauren Smith, Seth Coomer, and the undergraduate assistant, Sarah Miller. Additionally, I want to express my appreciation to Dr. Thang Nguyen, who not only served as a great colleague but also became like a brother to me.

I want to express my gratitude to my seniors, Dr. Manh Ta and Dr. Bo Phan. Dr. Manh Ta played a crucial role in connecting me to the CMR lab. Dr. Bo Phan's expertise in machine learning proved invaluable to me, and he generously shared many valuable techniques and insights with me.

I would also like to thank Dr. Taejoon Kim who was my academic advisor for my Master program at the University of Kansas.

I would also like to thank Dr. Hong Vu and Dr. Son Luu Nguyen who are like my family members and accompany me with a lot of helpful advice.

Finally, I feel an overwhelming sense of gratitude toward my family, including my parents, my grandparents, and my little sister, whose love and support has been helping me complete this journey. Their belief in me has instilled a deep sense of motivation, enabling me to overcome countless challenges and obtain this degree.

Abstract

Total knee arthroplasty (TKA), also known as total knee replacement, is a surgical procedure to replace damaged parts of the knee joint with artificial components. It aims to relieve pain and improve knee function. TKA can improve knee kinematics and reduce pain, but it may also cause altered joint mechanics and complications. Proper patient selection, implant design, and surgical technique are important for successful outcomes. Kinematics analysis plays a vital role in TKA by evaluating knee joint movement and mechanics. It helps assess surgery success, guides implant and technique selection, informs implant design improvements, detects problems early, and improves patient outcomes. However, evaluating the kinematics of patients using conventional approaches presents significant challenges. The reliance on 3D CAD models limits applicability, as not all patients have access to such models. Moreover, the manual and time-consuming nature of the process makes it impractical for timely evaluations. Furthermore, the evaluation is confined to laboratory settings, limiting its feasibility in various locations.

This study aims to address these limitations by introducing a new methodology for analyzing in vivo 3D kinematics using an automated deep learning approach. The proposed methodology involves several steps, starting with image segmentation of the femur and tibia using a robust deep learning approach. Subsequently, 3D reconstruction of the implants is performed, followed by automated registration. Finally, efficient knee kinematics modeling is conducted. The final kinematics results showed potential for reducing workload and increasing efficiency. The algorithms demonstrated high speed and

accuracy, which could enable real-time TKA kinematics analysis in the operating room or clinical settings. Unlike previous studies that relied on sponsorships and limited patient samples, this algorithm allows the analysis of any patient, anywhere, and at any time, accommodating larger subject populations and complete fluoroscopic sequences. Although further improvements can be made, the study showcases the potential of machine learning to expand access to TKA analysis tools and advance biomedical engineering applications.

Table of Contents

Chapter I Introduction.....	1
1.1 Motivation.....	1
1.2 Method Overview	4
1.3 Image Segmentation.....	8
1.4 3D Reconstruction	10
1.5 Deep Learning.....	12
Chapter II Literature Review	15
2.1 Image Segmentation.....	15
2.1.1 Manual Segmentation	15
2.1.2 Thresholding	16
2.1.3 Edge-based Segmentation.....	16
2.1.4 Region-based Segmentation.....	18
2.1.5 Deep Learning Segmentation.....	18
2.2 3D Reconstruction	20
2.2.1 CAD/Laser Scan	21
2.2.2 Medical Imaging	22
2.2.3 Non-rigid/Deep Learning Registration	23
2.3 3D-2D Registration.....	25
2.3.1 Manual Registration.....	26
2.3.2 Optimization Methods	28

2.4 Machine Learning	29
Chapter III Objectives and Contributions	32
3.1 Dissertation Objectives	32
3.2 Contributions.....	33
Chapter IV Methods Overview	35
4.1 Problem Definition.....	35
4.2 Current Processing Methodology: Distortion Correction	39
4.3 Image Segmentation Using Deep Learning	44
4.4 3D Reconstruction Using Non-Rigid Registration.	45
4.5 3D-2D Registration Using Deep Learning	47
4.6 Kinematics Calculation	49
Chapter V Image Segmentation	51
5.1 Introduction.....	51
5.2 Method	56
5.2.1 Data Preparation: Convert Registration Data to Segmentation Data...	57
5.2.2 Data Preparation: Convert to COCO's format.....	57
5.2.3 Deep Learning: Training.....	59
5.2.4 Deep Learning: Testing and Evaluation	59
5.3 Results.....	61
5.3.1 Evaluation Criteria	63
5.3.2 Result of Several Representative Cases	65

5.3.3 Comparison with the Ground Truth.....	68
5.3.4 Comparison with Other Methods.....	70
5.3.5 Challenging Cases: Blurry Images.....	70
5.3.6 Challenging Cases: Edge Cases Images	78
5.3.7 Challenging Cases: Overlap Cases Images.....	78
5.3.8 Challenging Cases: Two Implanted Knees	79
5.3.9 Challenging Cases: Combinations	79
5.3.10 Application in Segmenting a Fluoroscopic Video	80
5.4 Discussion	80
Chapter VI 3D Reconstruction.....	84
6.1 Introduction.....	84
6.2 Methods.....	87
6.2.1 Model Processing: Boundary Vertices Statistics	92
6.2.2 Image Processing: Edge Detection	92
6.2.3 Morphing Algorithm.....	95
6.2.4 Model Refinement	96
6.2.5 Posts Reconstruction.....	98
6.2.6 Model Comparison Evaluation	98
6.3 Computational Results	98
6.4 Discussion	104
Chapter VII Automated 3D-to-2D Registration	110

7.1 Introduction.....	110
7.2 Methods.....	115
7.2.1 Camera Calibration	115
7.2.2 Deep Learning Model: Training with Fluoroscopic Images Data	122
7.2.3 Deep Learning Model: Training without Fluoroscopic Images Data	125
7.2.4 Tuning Optimization: Edge-based Registration.....	127
7.2.5 Tuning Optimization: Intensity-based Registration.....	129
7.3 Results.....	130
7.3.1 Comparison with the Ground Truths	130
7.3.2 Comparison with other Methods.....	134
7.4 Discussion	137
Chapter VIII Implanted Knee Kinematics Analysis	140
8.1 Introduction.....	140
8.2 Kinematics Model	141
8.3 Results of Knee Kinematics Analysis.....	144
8.4 Discussion	147
Chapter IX Conclusions and Future Work.....	149
9.1 Conclusions.....	149
9.2 Assumptions.....	154
9.3 Limitations and Future Work.....	155
List of References	157

Appendices.....	175
I. Extended Applications to THA and Patella	175
II. Software User Interfaces.....	175
Vita.....	179

List of Tables

Table 1. mAP comparison between the ground truth and the predicted registered knee image.....	69
Table 2. The evaluation of 18 reconstruction models.....	103
Table 3. Cruciate retaining and posterior stabilized TKA reconstructions evaluations.	105
Table 4. The average errors of the registration comparisons with the ground truths.	135
Table 5. The average errors of the benchmarks for the registration.	136

List of Figures

Figure 1. Total knee replacement surgery (templehealth.org).....	2
Figure 2. Conventional manual kinematics calculation procedure for recent years.	5
Figure 3. Proposed kinematics evaluation framework using deep learning.	7
Figure 4. Image segmentation example (telusinternational.com).	9
Figure 5. 3D reconstruction example (medium.com).	11
Figure 6. Artificial intelligence, machine learning, and deep learning (metaload.com). .	13
Figure 7. Manual image segmentation using Slicer software.	17
Figure 8. Example of image segmentation using thresholding (towardsdatascience.com).	17
Figure 9. Example of edge-based segmentation in image processing.	19
Figure 10. Example of region-based segmentation in image processing (researchgate.net).	19
Figure 11. Manual 3D-2D registration process.....	27
Figure 12. Comparison between deep learning and machine learning (semiengineering.com).....	31
Figure 13. (Top) Fluoroscopic image sequence for DKB from full extension to maximum flexion at 30° increments. (Bottom) Same image sequence with the respective CAD models overlayed (after completion of the 3D to 2D image registration process). ..	36
Figure 14. Fluoroscopic units with (a) not distorted images and (b) distorted images.....	38

Figure 15. Main methods framework for the proposed machine learning approach.	40
Figure 16. Example of image distortion correction.....	40
Figure 17. Beadboard examples.....	43
Figure 18. The image before (Left) and after (Right) correct the distortion.....	43
Figure 19. Example of the input and outputs of the image segmentation.....	46
Figure 20. The input and output of 3D reconstruction step.	46
Figure 21. Input and output of 3D-2D registration step.....	48
Figure 22. Example of the implanted knee fluoroscopic image (left) and its segmentation with femur (red) and tibia (blue).....	53
Figure 23. Examples of several representative challenging cases of implanted knee image segmentation that might be impossible by manual human vision.	55
Figure 24. Data preparation of one subject for training the deep neural network of image segmentation.	58
Figure 25. Example of converting registration data into segmentationd data.	58
Figure 26. The TKA dataset is verified by the open-source COCO viewer.	60
Figure 27. The framework of using the trained segmentation deep neural network.	62
Figure 28. Illustration of IoU (pyimagesearch.com).	64
Figure 29. A sample of an original fluoroscopy image of the PS implant patient and the corresponding of the segmented images.	66
Figure 30. A sample of an original fluoroscopy image of the CR implant patient and the corresponding of the segmented images.	67

Figure 31. The comparison of the deep learning method with other traditional methods.	71
Figure 32. The comparison of the deep learning method with other traditional methods (continued from the previous page).	72
Figure 33. A challenging case of the blurry image (left) and the predicted segmentation result (right).	73
Figure 34. A challenging case of the blurry image (left) and the predicted segmentation result (right).	73
Figure 35. A challenging case of the out-of-captured range image (left) and the predicted segmentation result (right).	74
Figure 36. A challenging case of the out-of-captured range image (left) and the predicted segmentation result (right).	74
Figure 37. A challenging case of the overlap case image (left) and the predicted segmentation result (right).	75
Figure 38. A challenging case of the overlap case image (left) and the predicted segmentation result (right).	75
Figure 39. A challenging case of the multiple implants image (left) and the predicted segmentation result (right).	76
Figure 40. A challenging case of the multiple implants image (left) and the predicted segmentation result (right).	76
Figure 41. A challenging case of the image (left) with multiple challenges (overlap & blurry) and the predicted segmentation result (right).	77

Figure 42. A challenging case of the image (left) with multiple challenges (edge & overlap & blurry) and the predicted segmentation result (right).	77
Figure 43. Application of deep learning in segmenting a fluoroscopic video.	81
Figure 44. The proposed problem, wherein 2D-3D registration is needed for analysis, but no CAD models are available to generate results.	86
Figure 45. 3D TKA reconstruction framework.....	88
Figure 46. Evaluation of the number of vertices to determine the ideal positions.	90
Figure 47. 3D TKA reconstruction framework with an appropriate number of fluoroscopy images to maximize data output and minimize patient risk.....	90
Figure 48. Registration of the template model with errors.	91
Figure 49. Schematic diagram showcasing the defined directions used for analysis of the kinematic variables.	91
Figure 50. Boundary vertices and edge detection in the morphing algorithm: Model Processing Input.....	93
Figure 51. Boundary vertices and edge detection in the morphing algorithm: Model Processing Output.	93
Figure 52. Boundary vertices and edge detection in the morphing algorithm: Image Processing Input.....	94
Figure 53. Boundary vertices and edge detection in the morphing algorithm: Image Processing Output.	94

Figure 54. An example of a worse morphed CAD model needed to be smoothed: Multiple sections in one model.....	97
Figure 55. An example of a worse morphed CAD model needed to be smoothed: Before smoothing.....	97
Figure 56. An example of worse morphed CAD model needed to be smoothed: After smoothing.....	97
Figure 57 The wrong position of the posts after morphing (left) and the correction by using rigid registration (right).....	99
Figure 58. The results of overlaying the unmorphed model (left) and morphed model (right) onto the fluroscopic image.....	101
Figure 59. Morphed model overlays.	101
Figure 60. Heat maps show comparisons between the ground truth and the morphed model: Top view with largest error.....	102
Figure 61. Heat maps show comparisons between the ground truth and the morphed model: Bottom view with smaller error articulating surfaces.....	102
Figure 62. Several morphed models result.	103
Figure 63. Implant morphing algorithm applications.	105
Figure 64. A sequence of fluoroscopic images for a DKB from full extension to maximum flexion at 30° increments is show in the top row, and the same sequence of images with the respective CAD models overlayed (after completion of the 3D to 2D image registration process) is shown on the bottow row.....	111

Figure 65. Filmstrip showcasing the implant during DKB activity from full extension to maximum flexion.	111
Figure 66. Example of clear image (left) and blurry image (right) that is a challenge of manual registration.....	113
Figure 67. 3D-2D registration methodology using deep learning.	116
Figure 68. Virtual camera setup (left) to simulate the fluoroscopy (right).	116
Figure 69. Camera setup in the data collection system (Karbs).	119
Figure 70. Translated camera setup in the data preparation for training of the deep learning model.....	120
Figure 71. Training step for deep learning neural network structures with fluoroscopic data.	124
Figure 72. Training step for deep learning neural network structures without the fluoroscopic data.	126
Figure 73. Tuning optimization framework.....	128
Figure 74. Registration results represented by the 3D box for a clear image (left) and blurry image (right).....	131
Figure 75. Registration results of a fluoroscopic sequence using machine learning.	131
Figure 76 Registration results with deep learning (left) and then applied optimization (right).	132

Figure 77. Registration results with initial position (pink), machine learning result (blue), optimization result (red), and the ground truth (green) of the PS implant case (top) and CR implant case (bottom).	133
Figure 78. Parameters of interest: Femorotibial contact (LAP, MAP, LSI, MSI) and axial rotation (AXR).	142
Figure 79. Converting the models from the global space to the tibia space.	142
Figure 80. Femorotibial contact (top) and axial rotation (bottom).	145
Figure 81. Kinematics outputs of different approaches.	146
Figure 82. The summarization of the entire framework of 5 main steps of the dissertation: (i) data collection, (ii) image segmentation, (iii) 3D reconstruction, (iv) 3D-to-2D registration, and (v) kinematics modelling.	151
Figure 83. Segmentation results for the cup (left) and stem (right) of THA.	176
Figure 84. Registration results of the implants for THA.	176
Figure 85. Segmentation and registration of the patella at the non-implanted knee.....	177
Figure 86. User interface of Karbs.....	177
Figure 87. User interface for the registration in Karbs.	178
Figure 88. User interface for the 3D reconstruction tool.	178

List of Abbreviations and Symbols

TKA	Total Knee Arthroplasty
THA	Total Hip Arthroplasty
CAD	Computer Aided Diagnosis
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
ICP	Iterative Closest Point
CPD	Coherent Point Drift
SSM	Statistical Shape Model
LAP	Lateral Anterior/Posterior Position
MAP	Medial Anterior/Posterior Position
LSI	Lateral Superior/Interior Position
MSI	Medial Superior/Interior Position
AXR	Axial Rotation
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
DoF	Degrees-of-Freedom

Chapter I

Introduction

1.1 Motivation

Total knee arthroplasty (TKA) is a surgical procedure that is used to treat severe knee joint pain and disability due to various conditions such as osteoarthritis, rheumatoid arthritis, and post-traumatic arthritis, among others [1]. The procedure involves resurfacing a damaged or worn-out knee joint with an artificial joint, which is made of metal, plastic, or ceramic as seen in Figure 1. The primary goal of TKA is to relieve pain and restore joint function, mobility, and stability. TKA can also improve the quality of life for people who have been suffering from chronic knee pain, stiffness, and limited movement, and who have not found relief through non-surgical treatments such as medications, physical therapy, and lifestyle modifications. In general, TKA is recommended when non-surgical treatments are ineffective, and the knee joint damage is severe enough to impair the patient's ability to perform daily activities and significantly affect their quality of life.

TKA is a complex surgical procedure that requires precise surgical measurements and techniques combined with high-quality design of component parts as well as surgical jigs [1]. Thus, engineering plays a critical role in the development and advancement of total knee arthroplasty (TKA) technology. Engineers in the field of orthopaedics use their knowledge of materials science, biomechanics, and other engineering disciplines to design and develop the various components of the artificial knee joint. The engineers work closely with orthopedic surgeons to understand the unique challenges and requirements of TKA

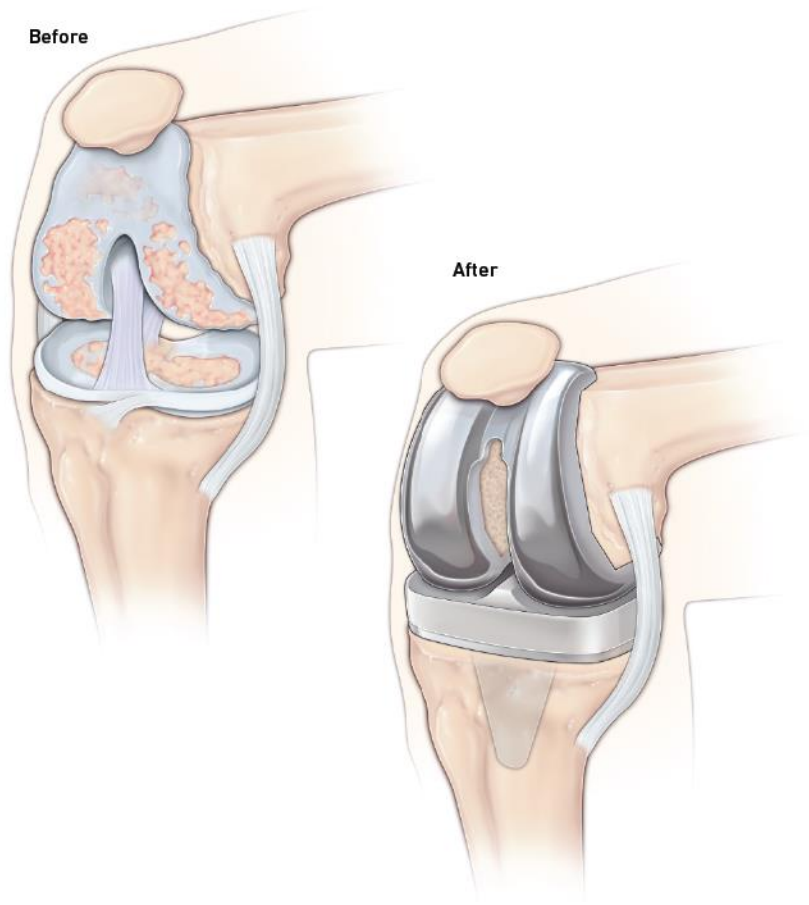


Figure 1. Total knee replacement surgery (templehealth.org).

and create solutions that meet the specific needs of both surgeons and patients. Collaborations between engineers and surgeons have helped to improve the precision, accuracy, and durability of TKA components and procedures, ultimately leading to better outcomes for patients. For example, recent engineering advancements has led to the development of computer-aided design and manufacturing (CAD/CAM) technologies that help create custom-made knee prostheses for patients [2]. These technologies specifically use imaging and modeling software to create a 3D model of the patient's knee joint, which is then used to create a customized knee prosthesis with a patient-specific fit to yield optimal functionality.

One specific contribution of engineering technology relates to analyzing the in vivo mechanics of the knee joint after total knee arthroplasty (TKA) [3, 4]. By analyzing knee kinematics, both engineers and surgeons can make more informed decisions, based on these outcomes, about implant positioning and soft tissue balancing, leading to better function and longevity of the implant. Poor alignment and balance can lead to complications such as instability, wear, and loosening of the implant, so such evaluations have helped improve optimal clinical outcomes after TKA by finding the proper alignment and balance of the knee implant. Additionally, by analyzing knee kinematics preoperatively, surgeons can identify potential problems early and take corrective measures to reduce the risk of these complications. Every patient has a unique anatomy, and analyzing knee kinematics can help surgeons to customize implant positioning and soft tissue balancing to fit the patient's specific needs. This can lead to better function and a

more natural feel of the knee after surgery. Finally, analyzing knee kinematics can also help to advance the technology of knee implants. By understanding the complex movements of the knee joint, engineers can design more advanced and sophisticated implants that mimic the natural movement of the knee more closely.

Thus, efficiently analyzing knee kinematics in TKA is crucial for optimizing clinical outcomes, reducing complications, customizing implant positioning, and advancing the technology of knee implants. While analyzing knee kinematics is critically important in TKA, the current techniques used to analyze kinematics pertains to relying on manual 3D-to-2D image registration software algorithms that are quite labor and time consuming [5, 6]. This dissertation proposes a novel approach to analyze the knee kinematics from fluoroscopy with high accuracy and automatic by utilizing a deep learning technique and a non-convex optimization method. The automated approach by machine learning can improve the accuracy of the kinematics output by evaluating the lower quality of the data that cannot be completed by manual human vision. Also, it is very important to note that analyzing knee kinematics is nearly impossible without the 3D implanted CAD model [3, 7]. The proposed approach resolves the 3D implant CAD model creation problem by using a new robust non-rigid registration algorithm.

1.2 Method Overview

The conventional manual kinematics calculation existing for many recent years is shown in Figure 2. Fluoroscopy is used as the input for analyzing patient kinematics, because fluoroscopy is an effective medical imaging technique for getting the real-time



Figure 2. Conventional manual kinematics calculation procedure for recent years.

images inside the body. Initially, subjects are instructed to perform specific movements, such as walking or bending the knee, while the fluoroscopy unit captures real-time video of the knee joint. The subject's fluoroscopy video allows easy visualization of both the bone and the components during the activity.

In the current manual methodology, a limited number of “good” images are extracted manually by human vision, usually at 0, 30, 60, 90, and 120 degrees of the flexion, from the fluoroscopy video collected at the hospital. At the same time, the 3D model of the corresponding bone in the fluoroscopy video might be obtained using CT scans or generated by the implant design company for analyzing TKA. Then, this 3D implant model is registered manually onto the chosen fluoroscopy images. Finally, the interested kinematics parameters are computed in 3D space from the above 3D-2D registration step such as LAP, MAP, LSI, MSI, AXR so that surgeons and experts can judge the outcomes of the patient after TKA. The method allows for accurate and detailed analysis of joint movement, which can help improve patient outcomes and inform future research on knee replacement surgery.

While most of the steps of the conventional approach were completed manually, the overall proposed automated framework of standard kinematic evaluations of the knee is proposed as in Figure 3. There are five main steps of getting the kinematics from fluoroscopy: i) data collection, ii) implanted knee image segmentation, iii) 3D CAD model creation, iv) automated 3D-2D registration using deep learning, and v) implanted knee kinematics calculation. In our novel methodology, the femoral and tibial component

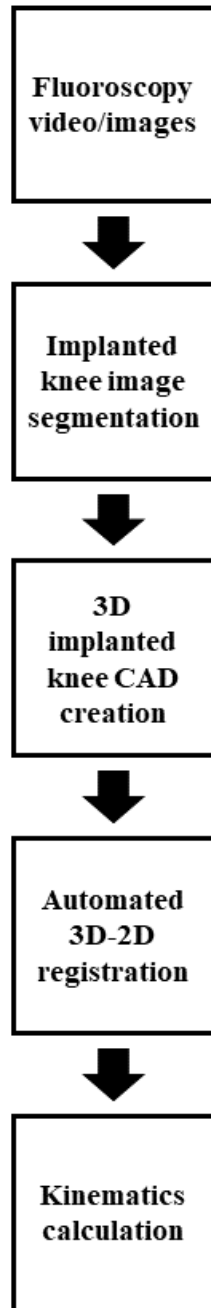


Figure 3. Proposed kinematics evaluation framework using deep learning.

silhouettes are extracted from the fluoroscopy using automated image segmentation. The output of the image segmentation, called the mask image, is a formatted image of either femur or tibia in isolation, such that the image can be used for the next steps of 3D reconstruction and 3D-to-2D image registration.

As mentioned above, these kinematics analyses are nearly impossible without CAD models, so the 3D implanted knee CAD creation is an optional step (if necessary) to reconstruct the implant CAD model from the segmented images. Once the CAD models are obtained or derived, the automated 3D-2D registration algorithms automatically overlay the implant CAD models, including the femur and tibia, onto the fluoroscopy images so that the movement and motion of the joint components are evaluated via the kinematics calculation. With the similar steps of the proposed approach in Figure 3 compared to the conventional manual approach in Figure 2, the automated approach is highly advanced for the determination of kinematics using all frames in the fluoroscopic video stream, even at low quality. By improving the current kinematics evaluation approach with advanced techniques, the proposed approach in Figure 3 has the potential to be used in practice and is discussed in more detail in the next chapters.

1.3 Image Segmentation

As the second step in Figure 3, image segmentation is the process of dividing an image into multiple regions or segments based on the visual characteristics of the pixels in the image, as seen in Figure 4. The goal of image segmentation is to simplify or change the



G - Original Image



H - Semantic segmentation



I - Instance segmentation



J - Panoptic segmentation

Figure 4. Image segmentation example (telusinternational.com).

representation of an image into something that is more meaningful and easier to analyze. Image segmentation is an important technique in computer vision and image processing, as it is a fundamental step in many applications such as object detection, recognition, tracking, and medical image analysis. It is also used in robotics, autonomous vehicles, and many other areas where visual perception is important. There are many different approaches to image segmentation, including thresholding [8], edge detection [9], clustering [10], and machine learning-based methods [11, 12]. Each method has its advantages and disadvantages, and the choice of method depends on the specific application and the characteristics of the images being analyzed. Overall, image segmentation is a critical step in many image analysis tasks, as it provides a way to extract meaningful information from images and to enable further analysis and processing.

1.4 3D Reconstruction

The third step in Figure 3 pertains to three-dimensional (3D) reconstruction is a fundamental task in computer vision, which involves the creation of a 3D model of an object or scene from two-dimensional images, as seen in Figure 5. It has broad applications in areas such as robotics, virtual reality, and augmented reality. In recent years, significant progress has been made in 3D reconstruction, thanks to the advancement of machine learning algorithms, improved camera technology, and the availability of large-scale datasets. Traditional 3D reconstruction methods are based on geometric and photometric approaches that rely on handcrafted features. However, in recent years, deep learning-based methods have shown superior performance in 3D reconstruction tasks [13],

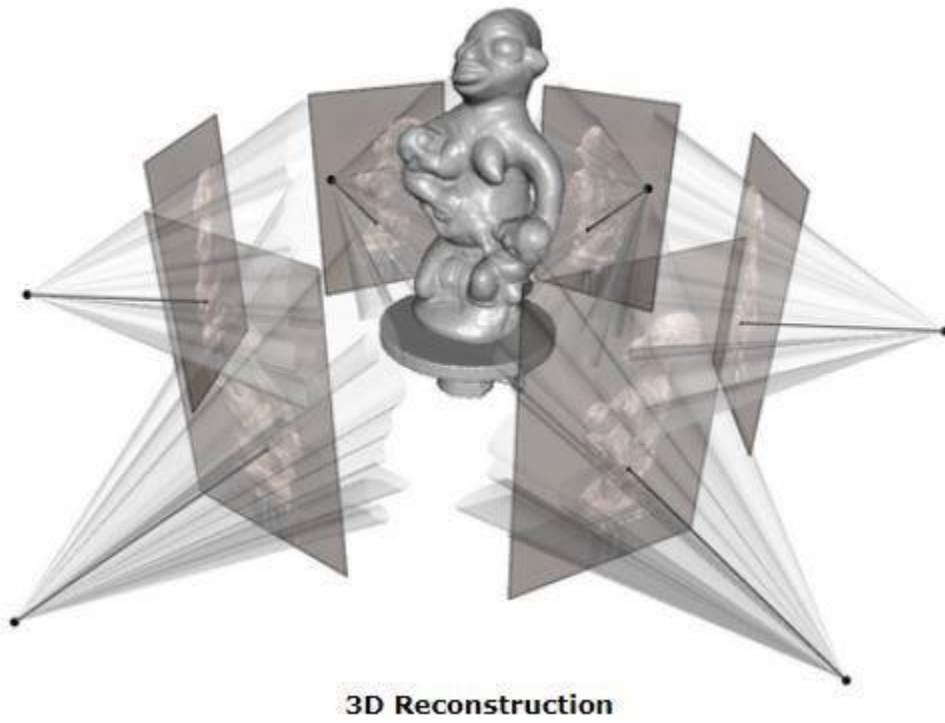


Figure 5. 3D reconstruction example (medium.com).

leveraging their ability to learn complex representations from large amounts of data. In this review, we will discuss the state-of-the-art techniques for 3D reconstruction, including both traditional and deep learning-based methods, their strengths and limitations, and future research directions.

1.5 Deep Learning

As seen in Figure 6, deep learning is a subfield of artificial intelligence (AI) that uses artificial neural networks to analyze and learn from large datasets [14]. These neural networks are modeled after the structure and function of the human brain, allowing them to learn from and make predictions based on complex data patterns. Deep learning has become increasingly popular in recent years due to advances in computing power, big data, and algorithmic developments. Deep learning algorithms can be used in a wide range of applications, including image and speech recognition, natural language processing, robotics, and autonomous driving. Utilizing advanced deep learning techniques for handling large amounts of data eliminates manual human work, deep learning is then used for the image segmentation and automated 3D-2D registration.

In deep learning, data is processed through multiple layers of artificial neural networks, with each layer learning a different aspect of the data. The neural network learns from the data through a process of trial and error, adjusting the weights of connections between neurons to minimize the error between the predicted output and the actual output. This process is called backpropagation. One of the key advantages of deep learning is that

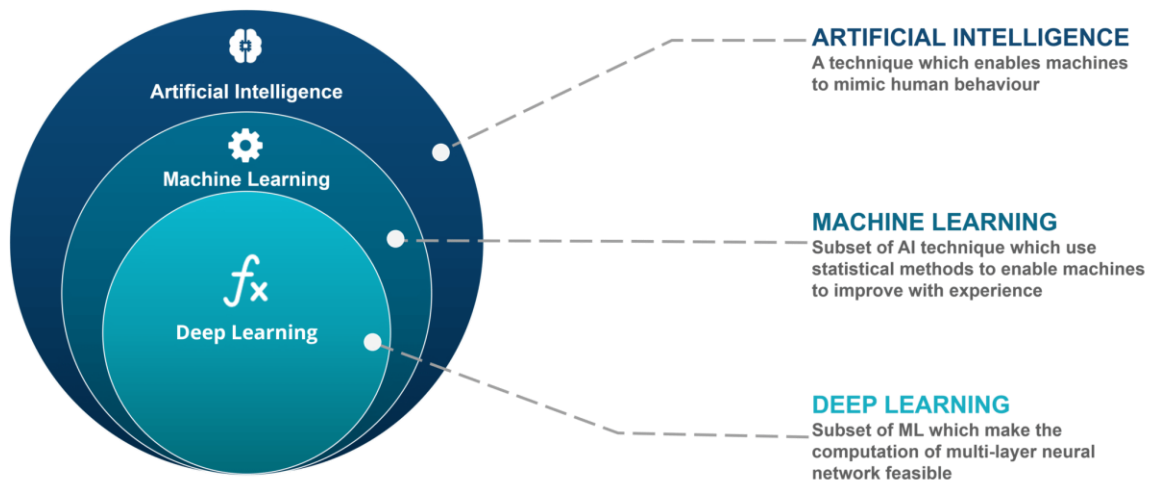


Figure 6. Artificial intelligence, machine learning, and deep learning (metaload.com).

it can automatically extract and learn features from raw data, without the need for manual feature engineering. This makes it well-suited to handle complex and high-dimensional data, such as images, video, and speech. Deep learning algorithms can be supervised, unsupervised, or semi-supervised. In supervised learning, the algorithm learns from labeled data, while in unsupervised learning, the algorithm learns from unlabeled data. In semi-supervised learning, a combination of labeled and unlabeled data is used. Some of the most used deep learning algorithms include convolutional neural networks (CNNs) for image and video analysis [15], recurrent neural networks (RNNs) for sequential data, and generative adversarial networks (GANs) for generating new data.

Chapter II

Literature Review

The algorithms developed herein demonstrate an automated methodology for analyzing and conducting the overlay step with regard to a fluoroscopic image to ultimately reduce the time required for the current manual overlay framework. To achieve this goal, we need to investigate three specific fields: image segmentation, 3D reconstruction, and automated 3D-2D registration. This chapter reviews the literature of the current techniques used for solving those problems and their applications in total knee arthroplasty.

2.1 Image Segmentation

Image segmentation is a fundamental task in computer vision that involves partitioning an image into multiple regions, each of which represents a separate object or background. The goal of image segmentation is to identify the boundaries of these regions accurately and efficiently, allowing for further analysis and processing of the image data. Over the past several decades, numerous methods have been proposed for image segmentation, utilizing a wide range of techniques and algorithms. In this literature review, we will explore some of the key approaches and advances in the field of image segmentation.

2.1.1 Manual Segmentation

Manual image segmentation is the process of manually separating the foreground object or region of interest from the background in an image. This technique is often used

in computer vision and image processing applications where accurate object detection is essential, as an example seen in Figure 7. Manual segmentation is typically performed by human annotators who use specialized software or tools to mark the boundaries of the object or region of interest. The process can be time-consuming and labor-intensive, but it allows for precise control over the segmentation, making it a valuable technique for tasks such as medical imaging, object recognition, and video processing. Despite advances in automated segmentation techniques, manual segmentation remains a crucial method in many areas of image analysis where accuracy is critical.

2.1.2 Thresholding

Thresholding is a simple and widely used technique for image segmentation. It involves selecting a threshold value, which is used to separate the image pixels into two groups based on their intensity values, as an example in Figure 8. Pixels with intensities above the threshold are classified as one group, while those with intensities below the threshold are classified as another. While thresholding is effective for separating objects from the background in simple images, it may not be sufficient for complex images with multiple objects or varying lighting conditions.

2.1.3 Edge-based Segmentation

Edge-based segmentation is another popular approach to image segmentation, which relies on identifying and extracting the edges or boundaries between different regions in an image. This can be accomplished through techniques such as the Canny edge

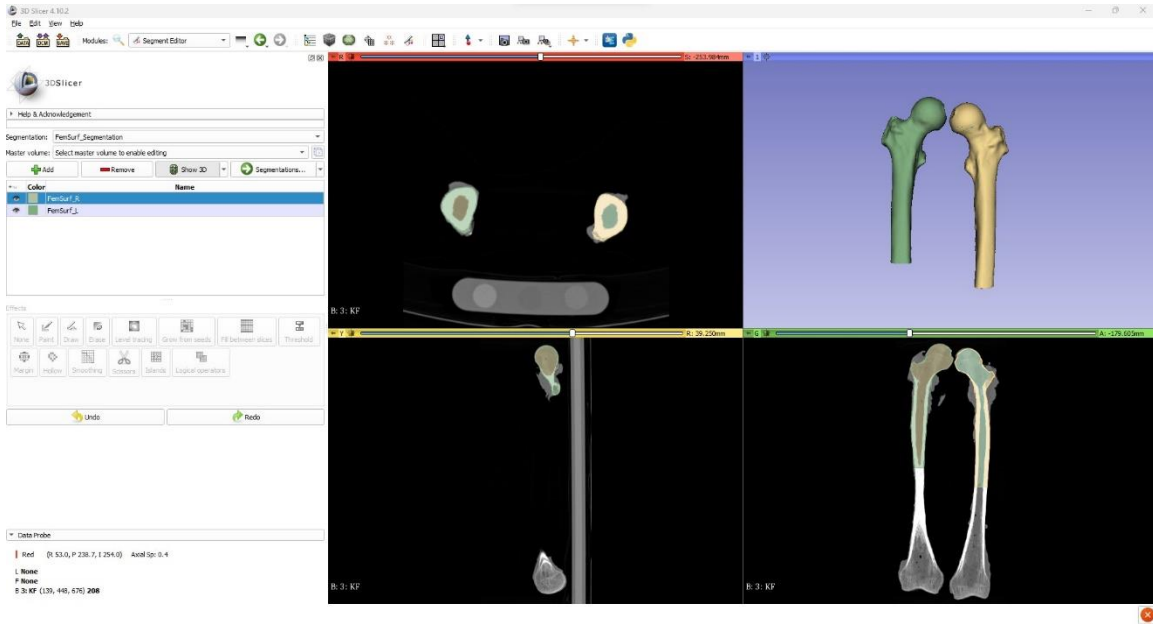


Figure 7. Manual image segmentation using Slicer software.

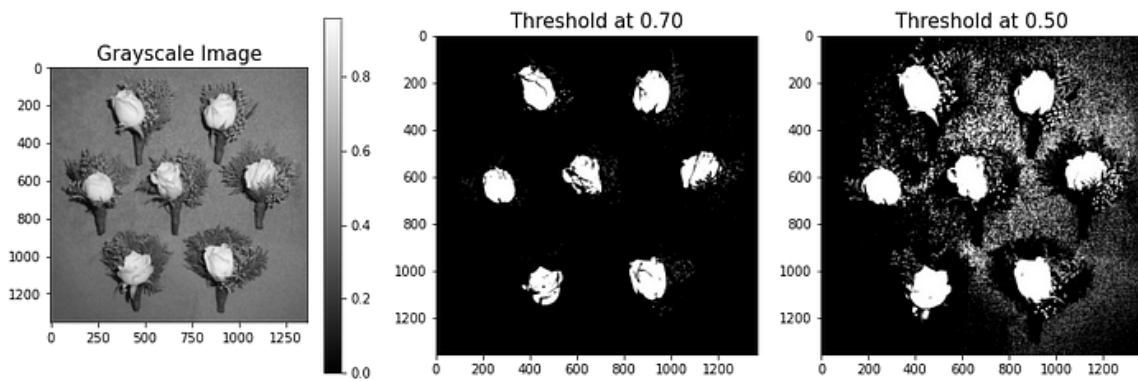


Figure 8. Example of image segmentation using thresholding (towardsdatascience.com).

detector, which applies a series of filters to highlight significant changes in intensity and direction in the image, as seen in Figure 9. Edge-based segmentation is particularly useful for images with well-defined boundaries but may be less effective for images with more diffuse or indistinct boundaries.

2.1.4 Region-based Segmentation

Region-based segmentation is a more complex approach to image segmentation that involves dividing an image into multiple regions based on similarities in color, texture, or other features, as seen in Figure 10. This can be accomplished through techniques such as clustering, where pixels are grouped together based on their similarity to other pixels in the image. Region-based segmentation can be more effective than thresholding or edge-based segmentation for complex images with multiple objects, but it can also be computationally intensive and may require careful tuning of parameters.

2.1.5 Deep Learning Segmentation

Image segmentation is the process of dividing an image into multiple parts or regions to extract useful information. Machine learning techniques, particularly deep learning, have been widely used for image segmentation in recent years due to their ability to learn complex patterns and features from data [12]. The most common approach for image segmentation using machine learning is to train a deep neural network (DNN) model on a dataset of labeled images, where each pixel is assigned to a particular class. The DNN

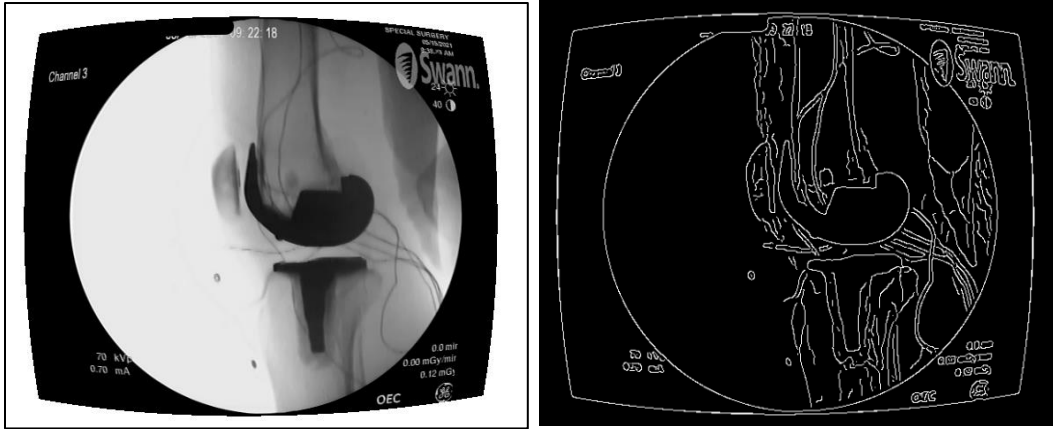


Figure 9. Example of edge-based segmentation in image processing.

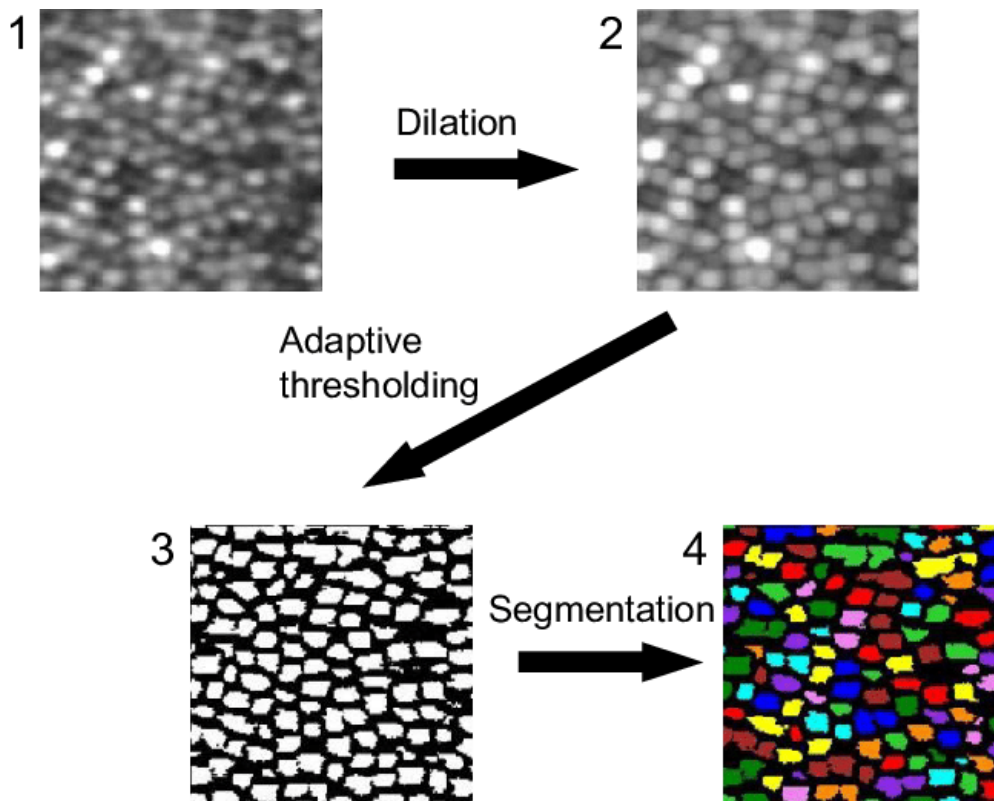


Figure 10. Example of region-based segmentation in image processing (researchgate.net).

model learns to identify patterns and features in the input images and maps each pixel to its corresponding class label. The output of the DNN model is a segmented image where each region or part is labeled with a different color or shade. There are several popular DNN architectures for image segmentation, such as U-Net, Mask R-CNN, and Fully Convolutional Networks (FCN). These models use different techniques to capture spatial information and context in the input image, such as convolutional layers, pooling layers, and skip connections. Image segmentation using machine learning has many applications, including medical image analysis, object detection and recognition, and autonomous driving. It is an active area of research, and new models and techniques are constantly being developed to improve the accuracy and efficiency of image segmentation.

In conclusion, image segmentation is a fundamental task in computer vision, with numerous techniques and approaches available for partitioning an image into multiple regions. While thresholding and edge-based segmentation are simple and widely used techniques, more complex approaches such as region-based segmentation and deep learning-based segmentation may be more effective for complex images with multiple objects. Additionally, multi-modal segmentation can help improve the accuracy and robustness of image segmentation by combining information from multiple imaging modalities.

2.2 3D Reconstruction

Three-dimensional (3D) reconstruction is a technique that aims to reconstruct the 3D shape of an object or scene from 2D images or sensor data. It has a wide range of

applications in fields such as computer graphics, robotics, and medical imaging. In this literature review, we will explore some of the key approaches and advances in the field of 3D reconstruction.

2.2.1 CAD/Laser Scan

Computer-aided design (CAD) models are used extensively in various industries to create 2D or 3D digital representations of products or structures. A literature review of CAD models reveals that they have significantly improved the design process by reducing the time and cost associated with physical prototyping. CAD models have also enabled designers to create more complex geometries and test various design iterations easily. The review also highlights the importance of accurate CAD models in simulation and analysis, particularly in the fields of engineering and architecture. Additionally, CAD models have been integrated with various technologies such as virtual reality and 3D printing, opening up new avenues for visualization and rapid prototyping. However, challenges such as data interoperability, model accuracy, and software compatibility still exist and require ongoing research to overcome.

Laser scanning is a popular technology for capturing and modeling 3D information of physical objects or environments. The technology uses a laser scanner to emit a laser beam that sweeps across the target area, measuring the distance to the objects in its path. The collected data is then processed to generate a point cloud, which represents the surface of the scanned object or environment. Laser scanning has been widely used in various fields, including architecture, engineering, archaeology, and geology, due to its high

accuracy and ability to capture detailed geometric and spatial information. It has also been utilized in various applications such as reverse engineering, quality control, heritage documentation, and virtual reality. However, challenges such as data processing and management, accuracy and resolution limitations, and the need for specialized skills and equipment, have been identified in the literature. Nonetheless, laser scanning remains a valuable tool in 3D modeling and has the potential for future developments and applications. Unfortunately, laser scanning cannot be utilized to create a 3D bone of a live subject and for cadaver specimens, all soft-tissues will have to be removed and the bone thoroughly cleaned before the laser scan process can be conducted. Therefore, laser scanning does have limitations in the field of orthopaedics.

2.2.2 Medical Imaging

A CT scan of a TKA knee is a medical imaging technique used to visualize the structure and position of an artificial knee joint. TKA, also known as total knee replacement, is a surgical procedure that involves replacing the damaged or diseased knee joint with an artificial joint made of metal and plastic components. CT scans use X-rays to create detailed images of the knee joint from different angles, which can help doctors evaluate the position of the implant, assess any bone or tissue damage, and detect any complications such as infection or loosening of the implant. CT scans are typically performed post-surgery as a part of follow-up care or if the patient experiences any issues or complications with the TKA knee. The procedure is non-invasive but contains radiation.

An MRI scan of an implanted knee is a medical imaging technique used to visualize the internal structure and soft tissue surrounding an artificial knee joint. MRI stands for Magnetic Resonance Imaging and uses a magnetic field and radio waves to produce detailed images of the knee joint. This type of imaging is particularly useful for evaluating the integrity of the soft tissues surrounding the knee implant, such as ligaments, tendons, and muscles, and can help detect any damage or abnormalities. MRI scans are also useful for identifying potential complications with the implant, such as infection, loosening, or dislocation. However, patients with certain types of implants, such as those made of metal, may not be able to undergo an MRI scan due to safety concerns. The procedure is non-invasive but is expensive. Unfortunately, it is very difficult to use either a CT or an MRI scan of a TKA because of the metal scatter that occurs. A metal artifact removal algorithm must be used, but again, this is not the most accurate method for creating a 3D image of an implanted knee.

2.2.3 Non-rigid/Deep Learning Registration

While the rigid registration only provides the information of transformation and scale of the 3D object, the non-rigid registration can design the shape of the 3D object. 3D reconstruction of the knee using non-rigid registration is a medical imaging technique used to create a 3D model of the knee joint and its surrounding structures by aligning and merging multiple medical images obtained from different imaging modalities [16]. Unlike rigid registration, non-rigid registration involves aligning images that have a variable relationship with each other, such as those obtained from MRI scans with different patient

positions. By merging these images using complex algorithms, a more accurate and detailed 3D model of the knee joint can be created.

3D reconstruction using non-rigid registration is a non-invasive and safe imaging modality that can be used to visualize and diagnose knee conditions such as ligament tears, meniscus injuries, and cartilage damage. The technique can also be used to monitor the progression of conditions such as osteoarthritis and to plan surgical interventions. The 3D reconstruction of the knee using non-rigid registration can provide highly detailed information about the knee joint, which can help doctors make more accurate diagnoses and treatment plans. However, it is important to note that this technique can be more time-consuming and computationally intensive than rigid registration and requires specialized software and expertise. Additionally, the accuracy and reliability of the 3D model may be influenced by factors such as patient motion during imaging and the quality of the input images.

3D reconstruction using deep learning is a medical imaging technique that uses artificial intelligence algorithms to create a 3D model of the knee joint and its surrounding structures. Deep learning is a subset of machine learning that involves training neural networks to identify patterns and relationships in complex data sets, such as medical images. 3D reconstruction using deep learning is a non-invasive and safe imaging modality that can be used to visualize and diagnose knee conditions such as ligament tears, meniscus injuries, and cartilage damage. The technique can also be used to monitor the progression of conditions such as osteoarthritis and to plan surgical interventions. The 3D

reconstruction of the knee using deep learning can provide highly accurate and detailed information about the knee joint, which can help doctors make more accurate diagnoses and treatment plans.

Additionally, deep learning algorithms can improve the efficiency and accuracy of the 3D reconstruction process and can be used to automatically segment and classify structures within the knee joint. However, it is important to note that the accuracy and reliability of the 3D model may be influenced by factors such as the quality and quantity of input data used to train the deep learning algorithm. Additionally, the interpretation of the 3D model may require specialized expertise and training and may not be suitable for all patients depending on their individual needs and medical conditions. In conclusion, 3D reconstruction is a fundamental task in computer vision, with numerous techniques and approaches available for reconstructing the 3D shape of an object or scene. While traditional approaches such as CT scanning and MRI scan are widely used, new approaches such as non-rigid registration and deep learning-based reconstruction are promising and have shown remarkable performance in various applications.

2.3 3D-2D Registration

Three-dimensional (3D) to two-dimensional (2D) registration is the process of aligning a 3D image or model with a 2D image. It has a wide range of applications in fields such as medical imaging, robotics, and augmented reality. In this literature review, we will explore some of the key approaches and advances in the field of 3D-2D registration.

2.3.1 Manual Registration

Manual 3D-2D registration refers to the process of aligning a three-dimensional (3D) medical image with a two-dimensional (2D) medical image. This process is typically done manually by a radiologist or a medical imaging professional using specialized software. The purpose of 3D-2D registration is to improve the accuracy of medical imaging by aligning the 3D image with the 2D image to provide a more precise location of the area of interest. This is particularly useful in interventional procedures such as surgery, where the physician needs to accurately place a surgical instrument or implant within the patient's body. The manual 3D-2D registration process involves several steps as shown in Figure 11. First, the 2D image is loaded into the software, and the radiologist identifies a feature or landmark in the image. This process could be conducted on a bone or tissue structure that is easily recognizable. Then, the corresponding feature or landmark is identified in the 3D image.

The radiologist then adjusts the position and orientation of the 3D image until the identified feature in the 3D image matches the same feature in the 2D image. This process may involve multiple iterations and adjustments until the alignment is satisfactory. Once the 3D-2D registration is complete, the aligned images can be used for diagnostic or interventional procedures, such as guiding a surgical instrument or placing an implant with greater accuracy. While manual 3D-2D registration can be time-consuming and labor-intensive, it remains a valuable tool in medical imaging and intervention, particularly in

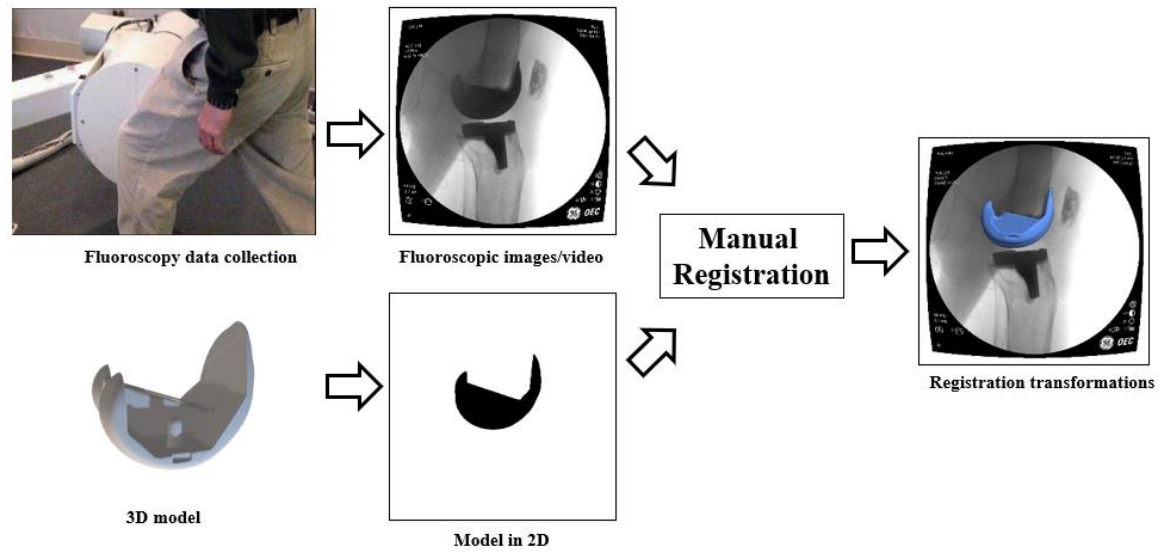


Figure 11. Manual 3D-2D registration process.

cases where high accuracy is required, usually considered as the “gold standard” for the other registration methods.

2.3.2 Optimization Methods

Intensity-based registration is a technique that aligns 3D and 2D images based on their pixel or voxel intensities [17]. It works by finding the optimal transformation that minimizes the difference between the two images. Intensity-based registration is widely used in medical imaging, where it can be used to align 3D medical images with 2D X-ray or fluoroscopy images. Edge-based registration is a computer vision technique that aligns images or point clouds by identifying and matching corresponding edges [18]. The edges are more stable and reliable features than other types of features, such as keypoints or textures, which can be affected by lighting conditions or other variations. The registration process involves estimating the transformation that maps one set of edges onto the other and applying it to align the images or point clouds.

Feature-based registration is a technique that aligns 3D and 2D images based on the correspondence of features or points [19]. It works by detecting corresponding features or points in both images and finding the optimal transformation that minimizes the distance between them. Feature-based registration is commonly used in robotics, where it can be used to align a 3D map of the environment with 2D camera images. Model-based registration is a technique that aligns a 3D model with 2D images. It works by extracting features or points from the 3D model and projecting them into the 2D images. The optimal transformation is then found by minimizing the distance between the projected points and

the detected features or points in the 2D images. Model-based registration is commonly used in augmented reality, where it can be used to align a virtual 3D model with real-world images.

Hybrid registration is a technique that combines multiple approaches, such as intensity-based and feature-based registration, to achieve higher accuracy and robustness [18]. Hybrid registration is commonly used in medical imaging, where it can be used to align 3D medical images with 2D X-ray or fluoroscopy images, as well as in robotics and augmented reality. Deep learning-based registration has recently emerged as a promising approach, where deep neural networks are trained to predict the optimal transformation between 3D and 2D images [20]. This approach has led to remarkable performances in various applications, including medical imaging and robotics.

In conclusion, 3D-2D registration is a fundamental task in various fields, with numerous techniques and approaches available for aligning 3D images or models with 2D images. While traditional approaches such as intensity-based, feature-based, and model-based registration are widely used, new approaches such as hybrid and deep learning-based registration are promising and have shown remarkable performance in various applications.

2.4 Machine Learning

Machine learning has revolutionized image segmentation by providing efficient and accurate methods for segmenting objects and regions within images. Image segmentation is the process of dividing an image into multiple segments or regions to

simplify and analyze the image's content. Machine learning algorithms can automatically learn the features of an image and use that knowledge to segment it. Convolutional Neural Networks (CNNs) are a popular machine learning technique used for image segmentation, as seen in Figure 12. CNNs learn to identify and extract important features of an image, which are then used to segment the image into different regions. Additionally, unsupervised learning techniques like clustering and anomaly detection can also be used for image segmentation tasks. Machine learning has made image segmentation faster, more accurate, and less prone to human error, making it a crucial tool for various applications, such as medical imaging, object detection, and self-driving cars.

Machine learning has demonstrated promising results in the challenging task of 3D-2D registration, which involves aligning a 3D image or model with a 2D image. This process is crucial in various applications, such as image-guided surgery, augmented reality, and robotics. Machine learning algorithms can learn the complex relationships between 3D and 2D images, which can be challenging to model using traditional methods. Convolutional Neural Networks (CNNs) have been used for 3D-2D registration, where the network learns to predict the 3D pose of an object based on its 2D projection. Other machine learning techniques, such as Random Forests, Support Vector Machines, and Deep Belief Networks, have also been used for 3D-2D registration. Machine learning can significantly improve the accuracy and speed of 3D-2D registration, making it an important tool for various medical and industrial applications.

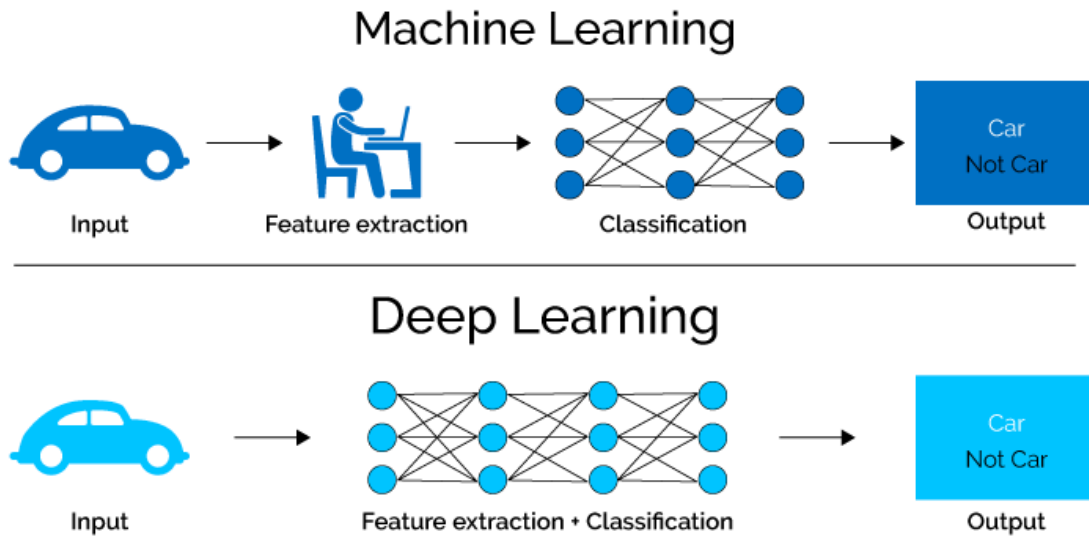


Figure 12. Comparison between deep learning and machine learning (semiengineering.com).

Chapter III

Objectives and Contributions

3.1 Dissertation Objectives

This dissertation aims to tackle the complex issues pertained to problems associated with evaluating joint kinematics by addressing several important sub-problems. To achieve this goal, the following objectives are outlined below:

1. First, the development of a fast and reliable image segmentation tool is necessary to extract the mask images of the implanted knee. This will be critical in enabling the subsequent steps of the evaluation process.
2. Second, a robust 3D reconstruction tool must be developed to accurately reconstruct the 3D implanted knee. This will require careful consideration of the complexities involved in imaging the knee joint.
3. Third, an automated and highly accurate 3D-2D registration tool is needed to overlay the 3D implanted knee model onto its 2D fluoroscopy images. This will enable the visualization of the knee joint during motion, providing valuable insights into its kinematics.
4. Finally, a fast and accurate kinematics model of the implanted knee must be developed to output the relevant kinematics parameters of interest. This will require the integration of data from the previous steps in order to produce a comprehensive model that accurately represents the knee joint during motion.

The successful completion of these objectives will represent a significant advance in the field of kinematics evaluation and will have important implications for the diagnosis and treatment of knee joint disorders.

3.2 Contributions

This dissertation represents a groundbreaking contribution to the fields of computer vision and medical imaging, making significant advances in four key areas.

1. First of all, a new algorithm for fast and efficient image segmentation has been introduced, which utilizes deep learning and can be applied to a wide range of medical images. This novel approach represents a significant improvement over existing segmentation techniques and has important implications for medical diagnosis and treatment.
2. Second, a novel method for 3D reconstruction has been presented, which accurately captures the structure and geometry of implanted models. This technique has the potential to significantly improve the accuracy and precision of surgical planning and implant design.
3. Third, an automated and highly accurate 3D-2D registration technique has been proposed, which enables the precise alignment of 3D models with 2D images obtained after surgery. This will facilitate the accurate visualization of surgical outcomes and enable more effective post-operative analysis.
4. Finally, an efficient implanted knee model for determining kinematics has been introduced, which can be used to simulate knee movements. This represents a

significant advance in the field of kinematics evaluation and has important implications for the diagnosis and treatment of knee joint disorders.

In summary, these contributions have significant implications for medical imaging, computer vision, and surgical planning. They represent important steps towards improving the accuracy and effectiveness of medical diagnosis and treatment and pave the way for further research in these fields.

Chapter IV

Methods Overview

Previous sections of this dissertation have provided a comprehensive understanding of the general framework and current techniques for addressing manual implanted knee kinematics calculations and their respective limitations. In contrast, this chapter offers a brief overview of the innovative methods proposed to tackle each step of this problem. The subsequent chapters will delve into the specifics of each method.

4.1 Problem Definition

This study presents a highly efficient and precise methodology for automatically computing the kinematics of a fluoroscopic video capturing an implanted knee, as demonstrated by the samples depicted in Figure 13. The proposed approach leverages the sequence of fluoroscopic frames, resembling the images showcased in the top row of Figure 13, as input. The resulting kinematic outputs encompass the movements of both the femur (green) and tibia (red) in three-dimensional space for the implanted knee, as illustrated in the bottom row of the same figure. Traditionally, these kinematic analyses necessitated manual execution, which proved to be time-consuming and labor-intensive. However, by harnessing previous datasets and implementing machine learning techniques, we introduce a novel approach that enables automated analysis of implanted knee kinematics. The subsequent step involves data preparation for the problem before delving into the specifics of the method.

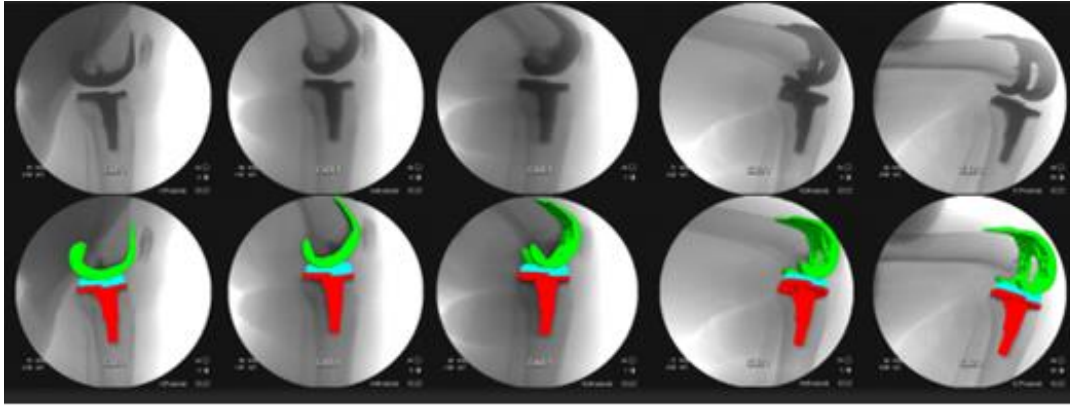


Figure 13. (Top) Fluoroscopic image sequence for DKB from full extension to maximum flexion at 30° increments. (Bottom) Same image sequence with the respective CAD models overlaid (after completion of the 3D to 2D image registration process).

More specifically, the Center for Musculoskeletal Research (CMR) research labs, and prior the Rocky Mountain Musculoskeletal Research Laboratory (later to become CMR) have been diligently accumulating patient data for over 30 years. Specifically, the team extracted high-quality fluoroscopy images from the video footage, at 0, 30, 60, 90, and 120-degree angles, for the purpose of kinematics analysis. These images were then corrected for distortion and employed in a virtual simulation, using the camera and fluoroscopy plane setup described in subsequent sections. This vast dataset serves as the foundation for training and testing the proposed method for knee kinematics analysis presented in this dissertation.

However, because the data was gathered from multiple sites with varying fluoroscopy unit types, it was necessary to standardize the data to enable effective utilization and validation. This process is known as data cleaning and includes distortion correction and camera calibration as the primary steps. To ensure the highest quality of data, this dataset was sourced from two distinct types of fluoroscopic units - one producing clear images and the other exhibiting distortion, as illustrated in Figure 14. Consequently, it is necessary to correct the distorted images for optimal accuracy. The distortion correction step is previously published [5] and has not been modified for this research. Nonetheless, an understanding of their general functionality is necessary before the machine learning algorithms can be correctly applied.



(a) Fluoroscopic image is not distorted



(b) Fluoroscopic image is distorted

Figure 14. Fluoroscopic units with (a) not distorted images and (b) distorted images.

After the data is corrected, calibrated, and standardized, our novel machine learning framework, as illustrated in Figure 15, can be employed. The first step in this framework involves image segmentation to generate masks of the femur and tibia in the fluoroscopy images. A comprehensive discussion of this step can be found in Chapter V Image Segmentation. If necessary, the predicted mask image can be used to reconstruct a 3D CAD model of the implanted knee, as detailed in Chapter VI 3D Reconstruction. The predicted segmented images are also passed through a 3D-2D registration deep neural network to obtain the initial pose, followed by refinement via an optimization tool to achieve final registration. This process is explored in-depth Chapter VII Automated 3D-to-2D Registration. Ultimately, after successfully registering 3D-2D, the movement of the femur and tibia in 3D space will provide the kinematic information through the mathematical knee model presented in Chapter VIII Implanted Knee Kinematics Analysis.

4.2 Current Processing Methodology: Distortion Correction

Distortion correction of X-ray images is an essential technique used in medical imaging to improve the accuracy and reliability of diagnostic images, as an example in Figure 16. X-ray images can often suffer from image distortion caused by image processing [21] because the process of developing and digitizing X-ray films can also introduce distortion in the image. For example, if the film is not developed correctly or the digitization process is not calibrated properly, it can cause the image to appear distorted. The distortion correction technique is aimed at correcting these distortions to improve the accuracy of the image and, in turn, enhance the accuracy of the diagnosis. The process of

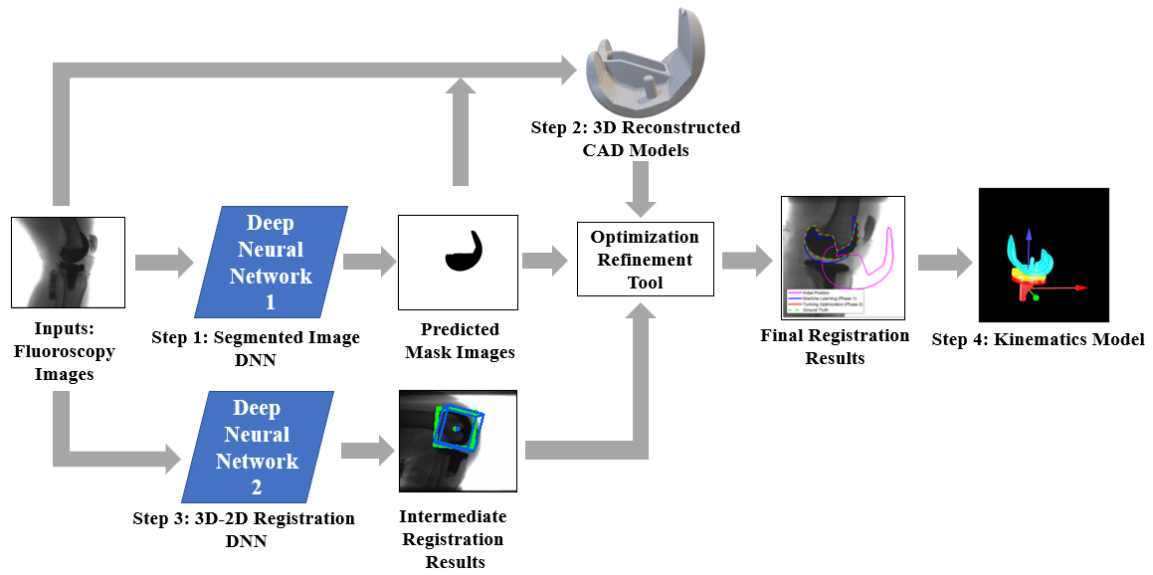


Figure 15. Main methods framework for the proposed machine learning approach.

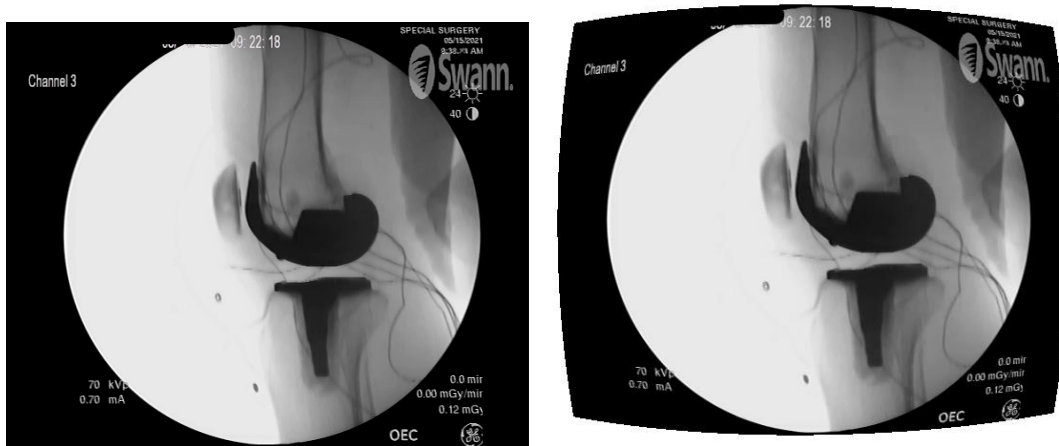


Figure 16. Example of image distortion correction.

distortion correction involves the use of software algorithms that analyze the image and identify any distortions present. Once identified, the algorithm applies a corrective transformation to the image, adjusting for the distortions and improving the accuracy of the image. This process can be automated, which allows for high throughput and faster diagnosis times.

One of the main advantages of distortion correction of X-ray images is its ability to improve the accuracy of the diagnosis. Correcting distortions that are present in an image leads to clearer and easier to interpret images, which in turn contributed to a more accurate diagnosis. This can be particularly beneficial for patients with complex conditions that require careful examination and analysis. Another advantage of distortion correction is that it can be applied to a wide range of X-ray imaging modalities, including CT scans, MRI scans, and ultrasound. This means that the technique can be used to improve the accuracy of a wide range of diagnostic tests, further improving patient outcomes and reducing the need for further testing. Distortion correction of X-ray images is a valuable technique that has the potential to improve the accuracy of diagnostic imaging and, in turn, enhance patient outcomes. Its ability to be applied to a wide range of imaging modalities and to be automated for high throughput makes it an essential tool in the field of medical imaging.

Prior to distortion correction, the distortion must be quantified and analyzed using a control image. Specifically, a highly engineered “beadboard” is to determine the nature of the image distortion. The beadboard is manufactured with tightly controlled tantalum beads placed exactly 1 cm apart, then it is positioned in front of the fluoroscopic unit and

captured accordingly (Figure 17). Standard image processing techniques are typically used to “clean” the image so the contrast between the beads and the image background is more significant. Using the information from these tantalum beads, a distortion correction algorithm is used to “unwarp” the virtual fluoroscopic image, so all beads are re-aligned as they are in the real world [21]. Sample images from before (Left) and after (Right) unwarping are seen in Figure 18. This distortion correction algorithm can then be applied to all fluoroscopic images captured from the same fluoroscopy unit on the day of data collection. It is wise to conduct the distortion process each day the fluoroscopic unit is used to take into account changes that may occur.

One of the main advantages of distortion correction of X-ray images is its ability to improve the accuracy of the diagnosis. Correcting any distortion that is present leads to images that become clearer and easier to interpret, which leads to more accurate diagnoses. This can be particularly beneficial for patients with complex conditions that require careful examination and analysis. Another advantage of distortion correction is that it can be applied to a wide range of X-ray imaging modalities, including CT scans, MRI scans, and ultrasound. This means that the technique can be used to improve the accuracy of a wide range of diagnostic tests, further improving patient outcomes and reducing the need for further testing. Distortion correction of X-ray images is a valuable technique that has the potential to improve the accuracy of diagnostic imaging and, in turn, enhance patient outcomes. Its ability to be applied to a wide range of imaging modalities and to be automated for high throughput makes it an essential tool in the field of medical imaging.

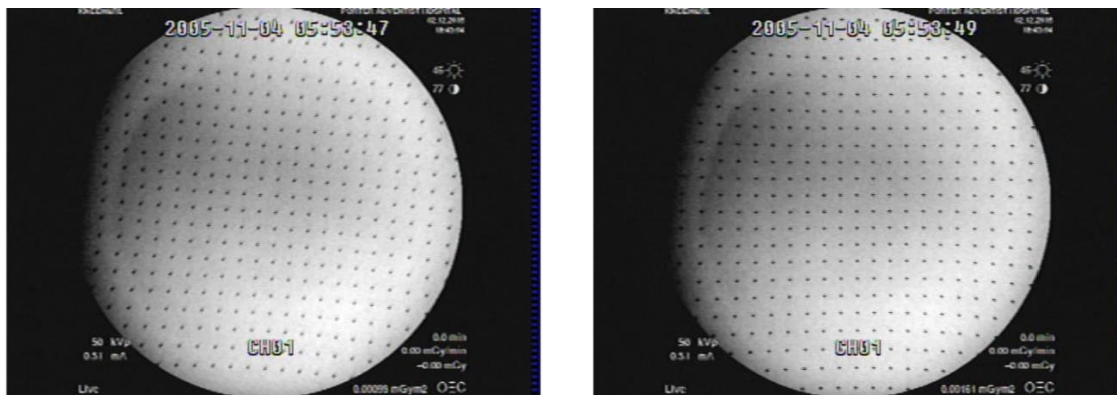


Figure 17. Beadboard examples.

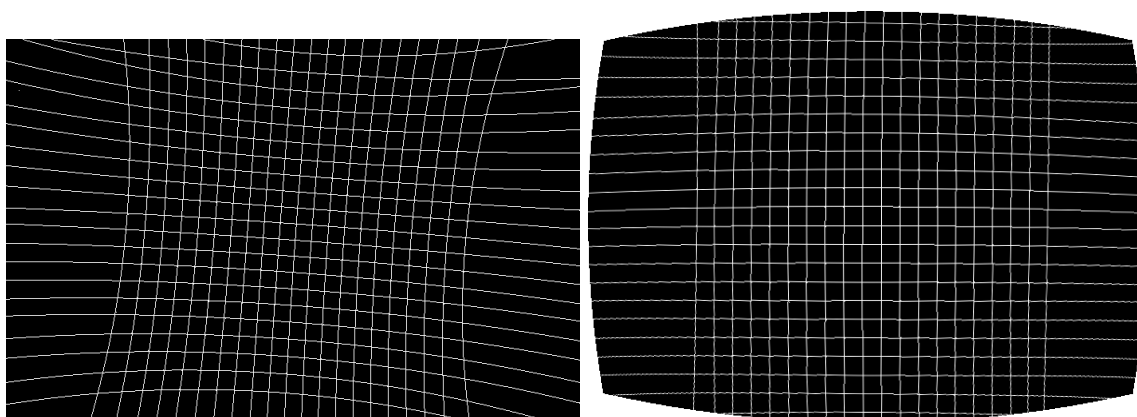


Figure 18. The image before (Left) and after (Right) correct the distortion.

In the context of medical imaging, it is commonly accepted that slightly distorted images may suffice for simple diagnosis. However, when it comes to conducting precise quantitative measurements on such images, a degree of distortion can lead to significant errors. It is therefore imperative for accurate assessment that any image warped image be accurately corrected before proceeding with further analysis. Failure to undertake such measures may result in overlay misalignment and the inability to accurately discern the shape of components within the video. In light of this, the unwarping of images is deemed an essential step in the pursuit of reliable and accurate results.

4.3 Image Segmentation Using Deep Learning

The first step of the machine learning framework as seen in Figure 15 is the image segmentation to take out the femur and tibia information from fluoroscopic image by removing the other around noises. Utilizing the above cleaned and labeled database, the segmentation of implanted knee images was achieved using a deep neural network model that shares similarities with Yolo [12]. This model was chosen for several reasons, primarily due to its ability to accurately predict multiple objects within the image such as the femur and tibia. In addition, the Yolo-based model is fast and can be applied in real-time scenarios, which is particularly important in medical settings where time is often of the essence.

The training data includes 5,000 human supervised subjects of both segmented femur and tibia images. After training the deep neural network with this dataset, the deep learning model can take any input of a fluorosocpy image and will output the segmented

images of both femur and tibia, called mask images as in Figure 19. This segmentation provides high accuracy, which is an important image processing step for the next problems of the 3D reconstruction and 3D-2D registration. The method details will be discussed in Chapter V Image Segmentation.

4.4 3D Reconstruction Using Non-Rigid Registration.

The step for 3D knee reconstruction is optional and required only when the 3D implanted knee models are not available. 3D reconstruction is the process of creating a 3D model of an object or a scene from a set of 2D images or videos. As in the above literature review, there are many registration techniques such as CT scanning, non-rigid registration, and machine learning. Since CT scanning contains radiation and is expensive, the proposed approach aims to replace the need for this method by reconstructing the 3D implant knee model from a few fluoroscopic images. Machine learning techniques, particularly deep learning, have shown promising results for 3D reconstruction in recent years. However, machine learning usually relies on a huge training dataset of 3D models that is not always available. Thus, utilizing the segmented images from the above section, a non-rigid registration method was proposed and used for reconstructing the 3D implanted knee models, including both the femur and tibia, by morphing a template model to match the silhouette of the actual model onto the fluoroscopy image, as an example in Figure 20. The method details will be discussed in Chapter VI 3D Reconstruction.

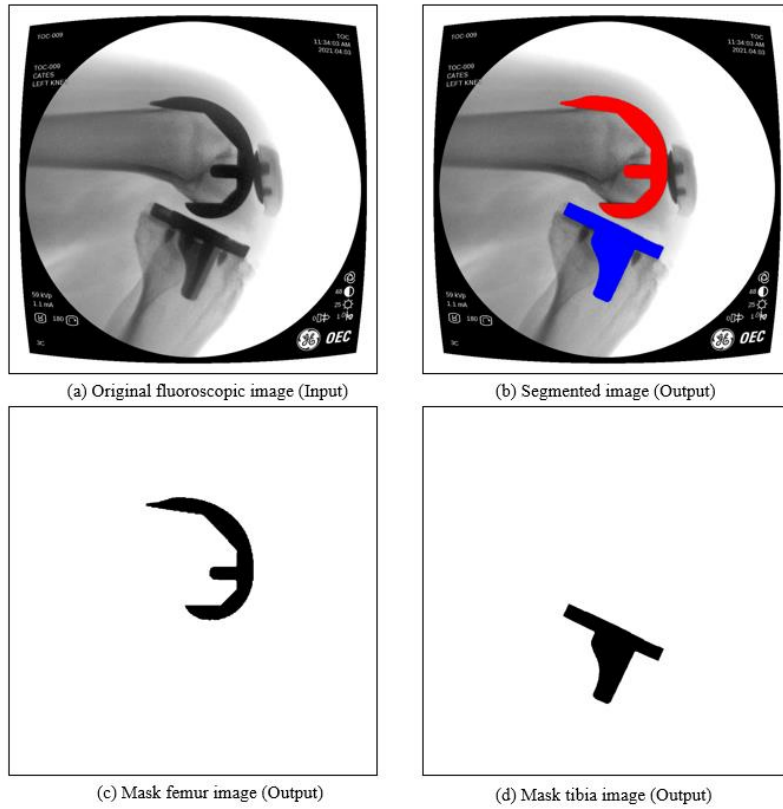


Figure 19. Example of the input and outputs of the image segmentation.

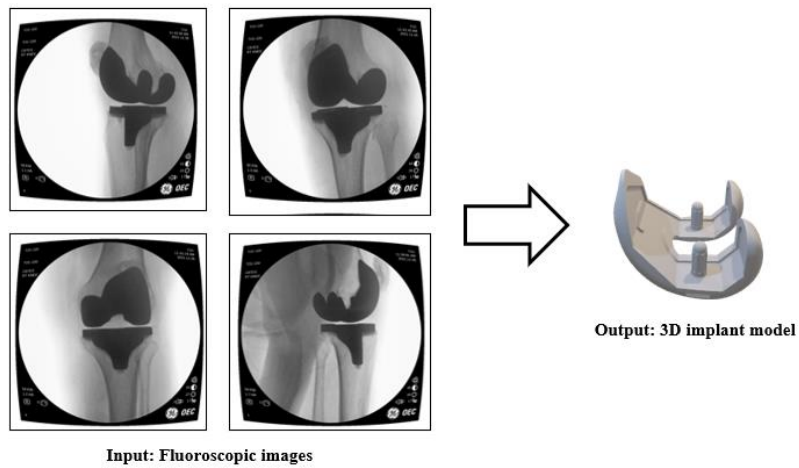


Figure 20. The input and output of 3D reconstruction step.

4.5 3D-2D Registration Using Deep Learning

After obtaining 3D models required for image registration, 3D-2D registration techniques are used to register the 3D implanted knee model onto the fluoroscopy images, as an example in Figure 21. 3D to 2D registration is the process of aligning a 3D model to a 2D image of the same object or scene. For many years, this process is usually completed manually, which quickly becomes most labor-intensive and time consuming [22]. The proposed approach aims to automate this procedure by using machine learning to achieve the high accuracy and efficiency of 3D to 2D registration.

The most common approach for 3D to 2D registration using machine learning is to train a deep neural network (DNN) model on a dataset of 3D models and their corresponding 2D images. The DNN model learns to map 3D points of a CAD model to their corresponding 2D image locations based on the input image and the 3D model. This allows the DNN model to estimate the pose and alignment of the 3D model relative to the 2D image. There are several popular DNN architectures for 3D to 2D registration, such as PointNetLK [23], DPOP [24], and PRNet [25]. These models use different techniques to capture spatial information and context in the input images and 3D models, such as point cloud processing, convolutional layers, and geometric constraints.

Among the existing machine learning models, EfficientPose [26] was chosen and constructed as the deep neural network model for registering implanted models for several reasons. This model is exceptionally fast and provides the highest accuracy compared to existing learning models, making it an ideal candidate for quickly registering implanted

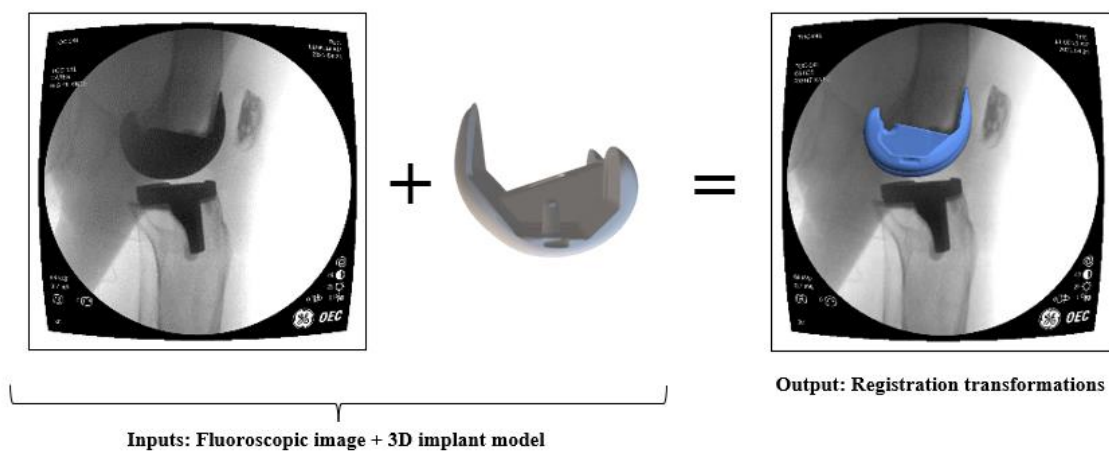


Figure 21. Input and output of 3D-2D registration step.

models while maintaining accuracy. Moreover, it can simultaneously predict multiple objects, including the femur and tibia. However, due to the variability of implant types from different companies and brands in the dataset, the machine learning results may contain errors. Therefore, after the predicted registration results from machine learning model are achieved, they are refined using a non-convex optimization problem to search for better registration fit around the initial pose. In Chapter VII Automated 3D-to-2D Registration, various optimization methods and non-learning registration methods will be compared to identify the optimal approach for implanted knees.

4.6 Kinematics Calculation

The ultimate goal of this study is for a subject to perform a specific activity and the algorithm derives the kinematics for an implanted knee. The analysis focuses on specific parameters such as AXR (Axial Rotation), LAP (Lateral Anterior/Posterior Position), MAP (Medial Anterior/Posterior Position), LSI (Lateral Superior/Interior Position), and MSI (Medial Superior/Interior Position), which are critical in describing the position and movement of the implant components with respect to the patient's anatomy. AXR refers to the rotational movement of the knee joint, and it is crucial to restore this natural rotation in TKA to prevent implant wear and instability. LAP and MAP, as well as LSI and MSI, describe the implant position in the anterior-posterior direction relative to the patient's anatomy, which is essential for proper patellar tracking and to prevent stress on surrounding tissues.

With respect to a TKA, it has been suggested many times that it is important to restore this natural motion of the knee joint to prevent excessive wear and instability of the implanted knee [27, 28]. Measuring AXR can help surgeons ensure that the implant is aligned properly to allow for natural rotational movement. Both LAP and MAP, as well as LSI and MSI, describe the position of the implant in the anterior-posterior direction relative to the patient's anatomy. Accurate placement of the implant in this dimension is important for proper tracking of the patella and to prevent excessive stress on the surrounding soft tissues. Misalignment in this dimension can lead to patellar instability and premature implant failure.

The above-described kinematics parameters have been used routinely in evaluating the kinematics of a TKA [29, 30, 31]. To calculate these interested parameters, a kinematic knee model in Chapter VIII Implanted Knee Kinematics Analysis was developed by incorporating the image segmentation, 3D reconstruction, and 3D-2D registration techniques discussed in above sections. The resulting model provides accurate measurements of the implant's position and movement in relation to the patient's anatomy, which is critical for the successful outcome of TKA.

Chapter V

Image Segmentation

The foundation for the upcoming chapters of 3D reconstruction and 3D-2D registration is image segmentation, which is a key image processing step required for automated 3D-2D registration. Specifically, image segmentation is used to process the fluoroscopic image to generate the 2D silhouette to which the corresponding 3D CAD model will later be registered. Utilizing a deep neural network technique with a large human-supervised dataset from the CMR research lab, knee image segmentation on every image in the database is performed. The simulation results demonstrate the model's robustness, achieving both high accuracy and fast running time.

5.1 Introduction

Image segmentation is the process of dividing an image into multiple parts or regions based on certain characteristics such as color, texture, or shape [11]. It is an essential task in computer vision and image processing, and it plays a crucial role in various applications, including object recognition and tracking. By segmenting an image into different objects, it becomes easier for a computer algorithm to recognize and track those objects in subsequent frames of a video or image sequence. In medical imaging such as CT scans or MRIs, image segmentation is used to identify and isolate specific structures or organs in the human body, which helps in diagnosing and treating diseases. For our 3D registration process, the raw fluoroscopy image contains unnecessary information such as bone texture, soft tissue texture, and generic excess image noises. Therefore, it is important

to remove this noise and isolate only the information of the implanted components, typically the femur and tibia, for kinematics analysis purposes, as demonstrated in Figure 22. Due to the size of the current database, developing a robust and fast image segmentation technique is important and becomes a foundation for the other problems such as 3D reconstruction and 3D-2D registration.

The traditional approach to image segmentation involves manual tracing of regions of interest by an expert, which can be time-consuming and prone to human error. Thus, automated image segmentation algorithms have been developed to overcome these limitations and provide faster, more accurate segmentation. There are several techniques for automated image segmentation, including thresholding, edge detection, clustering, and machine learning-based methods. Thresholding involves setting a threshold value to distinguish between the object of interest and the background [32]. While this approach is simple and easy to implement, it is subject to errors from image noise. Edge detection algorithms identify boundaries between different structures based on changes in intensity [33]. Clustering algorithms group pixels with similar characteristics into distinct regions [34]. The ability of those methods to segment blurry or low-quality fluoroscopic images is very limited. With the explosion of computation power, machine learning-based methods can use artificial intelligence algorithms efficiently to learn how to segment images based on training data.

Due to the dynamic nature of patient data collection, fluoroscopic videos and images used for 3D-2D registration can pose several challenges and can result in blurry

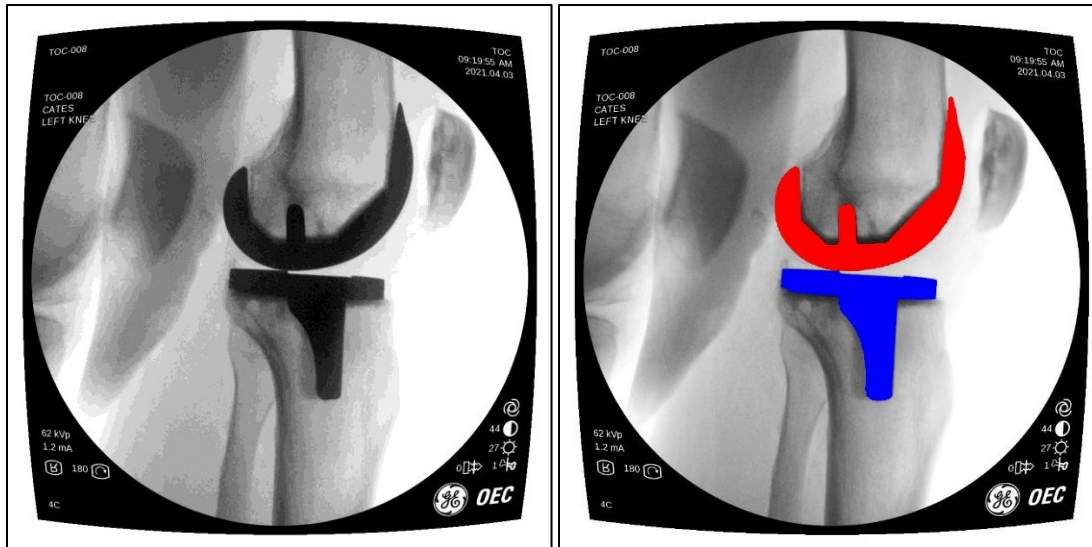


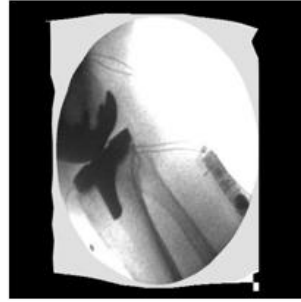
Figure 22. Example of the implanted knee fluoroscopic image (left) and its segmentation with femur (red) and tibia (blue).

images. For example, Figure 23 (a-i) illustrates several typical worst-case scenarios that can occur during data collection. The first case (a) is the most popular and characterized by blurry fluoroscopic images due to the low-rate video produced by the machine to ensure patient and technician safety. The next cases (b-c) frequently occurs when the knee is outside the captured range of the small fluoroscopic machine. Cases (d) and (e) arise when the implanted femur, implanted tibia, and resurfaced patella overlap, making it challenging to distinguish the metal implants in the fluoroscopy. The next case (f) occurs when both legs have been implanted, further complicating image capture. Sometimes the fluoroscopic data contains noise as case (g) with a sub-screen inside the image. The most challenging cases (h) and (i) combine all the above issues, such as blurry edge, and overlapping images. Usually, the traditional approaches require careful manual oversight in cases such as these, so resolving these cases automatically will be an important foundation for the next chapters.

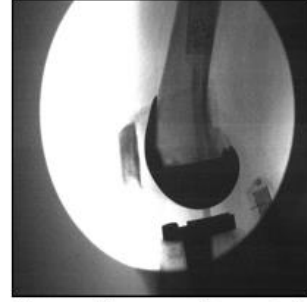
Deep learning is the most widely used machine learning approach for image segmentation. Among advanced deep neural networks available for this task, such as FCN [35], CRF-RNN [36], DPN [37], GCN [38], and Yolact [12], Yolact stands out for its high accuracy and speed, making it a logical choice for knee image segmentation. Firstly, because each fluoroscopy knee video contains multiple frames, which varies based on the frames-per-second of the individual fluoroscopy unit, the learning model needs to be fast enough to segment many images. Yolact is built based on ResNet model [39] that has excellent speeds, around 5 ms, to evaluate the mask. Additionally, the quality and



(a) Blurry image



(b) Out of the captured range (edge)



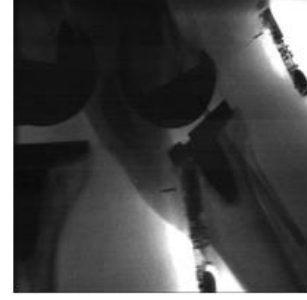
(c) Out of the captured range (edge)



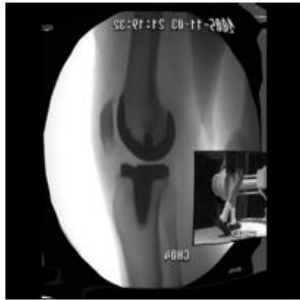
(d) Overlap (femur & tibia)



(e) Overlap (femur & resurfaced patella)



(f) Two implanted knees



(g) Noise (sub-screen)



(h) Combinations (blurry & overlap)



(i) Combinations (edge & blurry & overlap)

Figure 23. Examples of several representative challenging cases of implanted knee image segmentation that might be impossible by manual human vision.

robustness of Yolact in evaluating multiple subjects has been validated with higher accuracy in comparison with other methods using COCO's dataset with 2.5 million labeled subjects of 91 different objects in 328 thousand images [40]. Thus, Yolact model is very promising for segmenting implanted knee regions, including both femur and tibia, if the training supervised implanted knee dataset is big and accurate enough to train the model.

The remaining biggest challenge in using the above deep learning model is how to build the dataset for both training and testing. The CMR research lab has amassed a vast dataset of fluoroscopy data, comprising approximately more than 5,000 implanted knee images, which are further associated with corresponding 3D implanted models that were overlaid manually onto their fluoroscopy images for the conventional kinematics analysis purposes. By using a voxelization projection of the 3D models with known transformations onto the fluoroscopic image plane, the segmented image of the corresponding fluoroscopy image is created to become the training data for the image segmentation deep learning model. Overall, this chapter proposes a methodology to convert the existing registration dataset into effective training data for machine learning and build a robust segmentation deep learning model.

5.2 Method

A machine learning-based approach was selected for knee image segmentation because this method employs artificial intelligence algorithms that can learn how to segment images by utilizing training data. The method consists of two primary subsections:

preparation of the training data and development of the deep neural network model, which are both summarized as the framework in Figure 24.

5.2.1 Data Preparation: Convert Registration Data to Segmentation Data

Per Figure 24, the deep learning model dataset is comprised of three inputs: fluoroscopic images, corresponding CAD models, and 3D-2D registration transformations. The CMR lab has manually registered around 5,000 subjects with an accuracy of 0.65 mm and 1.5 degrees [5], providing a reliable dataset for a 3D-2D registration technique that is aimed at matching the projection of the 3D model with the silhouette of the implant in the 2D image.

By using voxelization projection, the known registration transformation can produce segmented images of both femur and tibia, as depicted in Figure 25. This process generates a raw segmentation dataset consisting of 5,000 pairs of the original fluoroscopic images and corresponding segmented and labeled images of either femur or tibia.

5.2.2 Data Preparation: Convert to COCO's format

The COCO (Common Objects in Context) dataset is a widely used, large-scale dataset for image recognition, segmentation, and captioning, which is particularly useful for developing and evaluating computer vision algorithms for object detection, segmentation, and image captioning [41]. By converting the current implanted knee dataset to the COCO format, it is possible to compare and verify the accuracy and performance of the proposed deep learning model against other existing learning methods. This involves

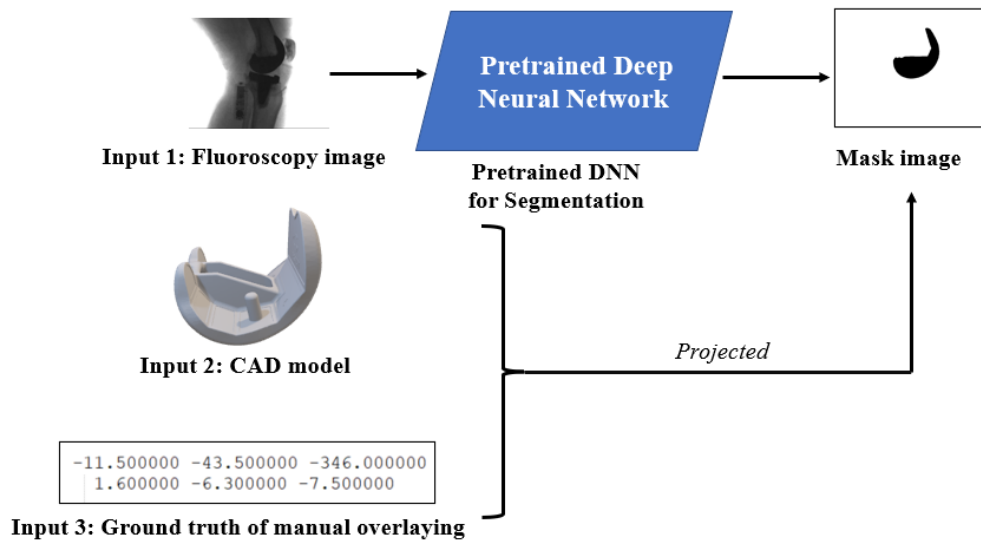


Figure 24. Data preparation of one subject for training the deep neural network of image segmentation.

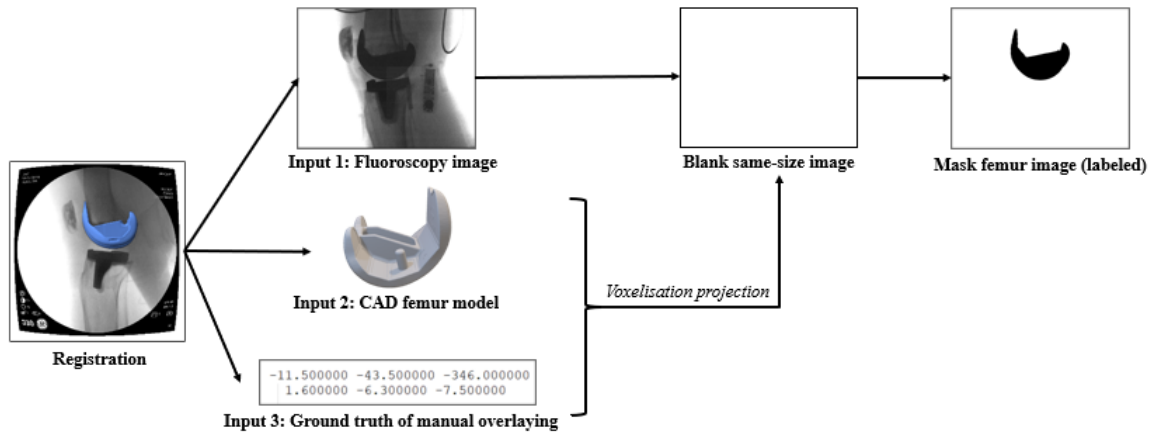


Figure 25. Example of converting registration data into segmentation data.

dividing the TKA dataset into training and testing subsets, with each subject containing the original fluoroscopic image in jpg format as well as additional parameters such as image size, bounding box, area, boundary, and category of the object (either femur or tibia), all saved in a json format file. The dataset is then verified using the open-source COCO viewer software, as seen in Figure 26, ensuring the accuracy and readiness of the dataset for training purposes.

5.2.3 Deep Learning: Training

The training data for the proposed deep learning model consists of 4,500 subjects, each containing both femur and tibia data. These subjects are fed into the model built based on the open-source Yolact [12] which is a development of ResNet model [42]. The model takes the original fluoroscopic images as inputs and produces segmented images with segmentation parameters of a bounding box, boundary, and label as outputs. The training procedure uses the backpropagation method, accelerated by GPU acceleration multiplication. This procedure is performed on the Alienware Aurora R13 Desktop, equipped with 32 Gb RAM and NVIDIA GeForce RTX 3080 GPU. It runs for 1 million epochs with a batch size of 8 and takes approximately 3 days to complete the training process.

5.2.4 Deep Learning: Testing and Evaluation

Upon completion of training the deep neural network, the model was evaluated on a testing set comprising 500 subjects, and subsequently validated in the following sections.

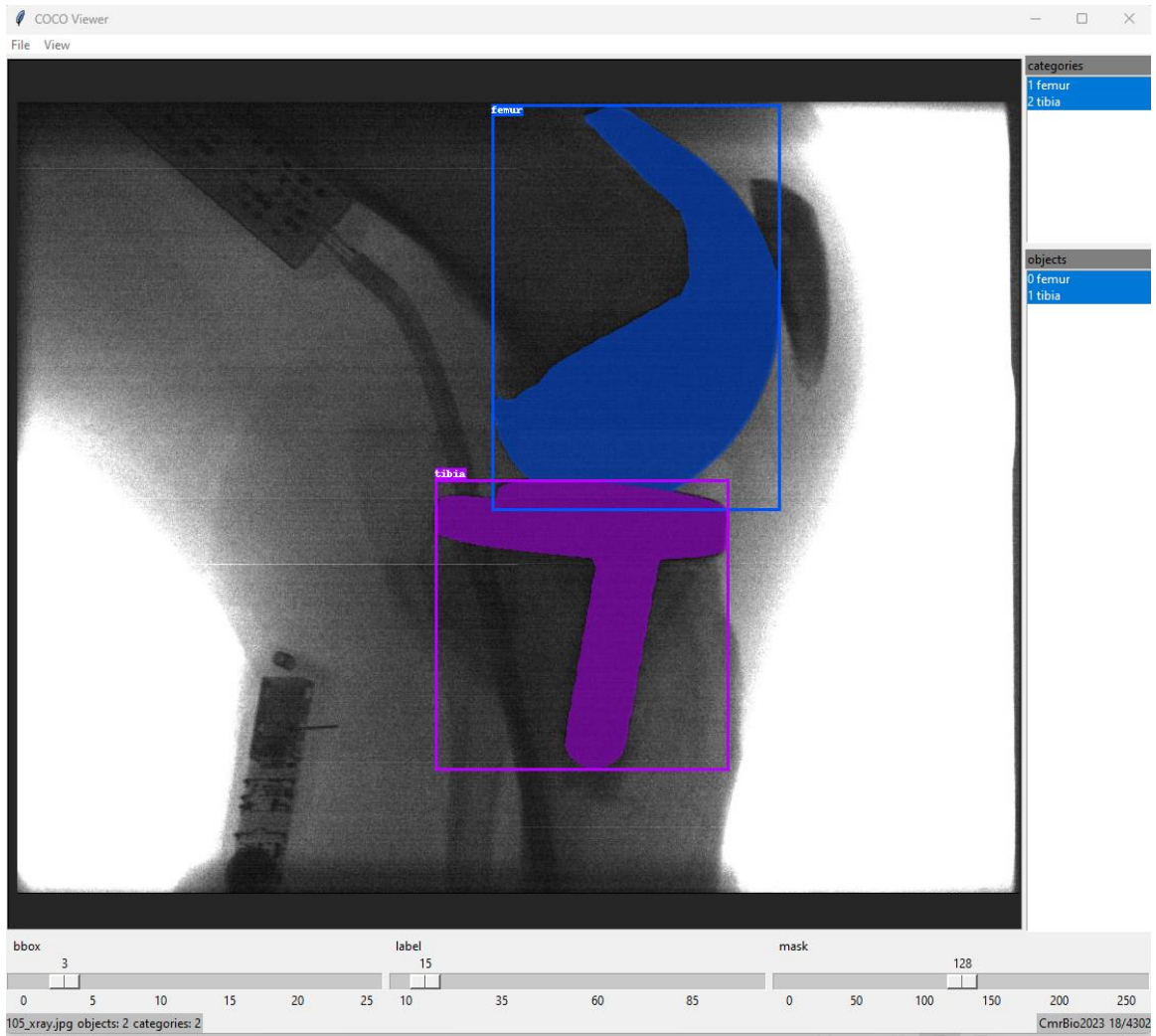


Figure 26. The TKA dataset is verified by the open-source COCO viewer.

As illustrated in Figure 27, the trained model was employed to predict the mask images of the femur and tibia in the implanted knee image. The DNN, after receiving the fluoroscopy image as input, outputs a segmented image or mask image. Notably, this DNN has the ability to predict both femur and tibia simultaneously. The predicted mask image obtained can be utilized to compare with the ground truth in the testing set, or to assess the accuracy in comparison with other segmentation or learning methods in subsequent sections.

5.3 Results

This section describes the evaluation of the proposed deep learning method in comparison to the other methods as well as the robustness of the deep learning method by segmenting a fluoroscopic video. The 5,000 implanted knee images in the database yield approximately 10,000 known segmented images of both femur and tibia, which were separated into a training set (90%) and testing set (10%). The simulation results of the trained deep learning model revealed impressive accuracy via comparison of the predicted segmentation images with the ground truth using Mean Average Precision (mAP) criteria (a metric for evaluating object detection models with a range of [0; 100]) [43]. With Intersection-over-Union (IoU) = 0.5, the mAP for predicting the mask of the femur is excellent at 98.99, while this number for tibia is approximately 100. The proposed method is also compared with the manual approach and other automatic methods as thresholding and edge detection. The deep learning model revealed impressive results when segmenting all frames of a fluoroscopic video, including blurry images that might not be able to be

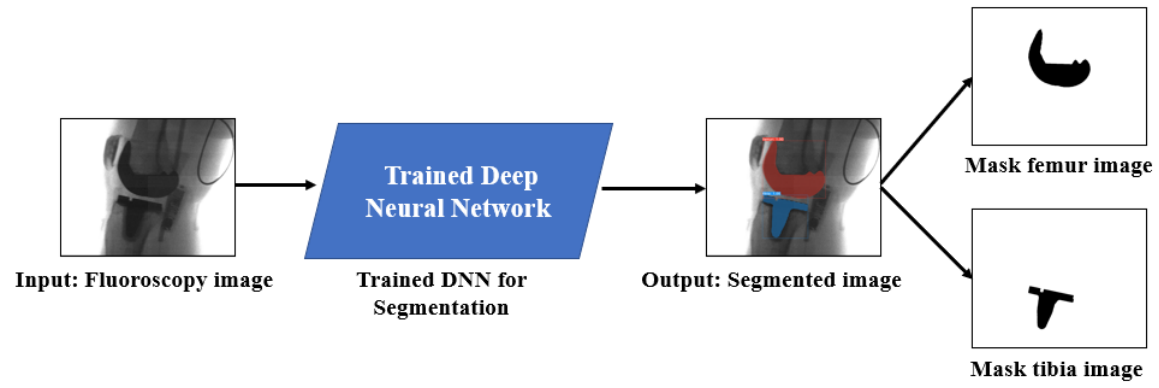


Figure 27. The framework of using the trained segmentation deep neural network.

segmented my human vision. The details of the results will be discussed in detail in the next sections.

5.3.1 Evaluation Criteria

The image segmentation algorithm is evaluated based on several metrics to measure overall accuracy and effectiveness. Some common metrics used for evaluating image segmentation include Intersection-over-Union (IoU), precision and recall, F1 score, and mean-Average-Precision (mAP) [43]. IoU measures the overlap between the predicted segmentation and the ground truth segmentation, as shown in Figure 28. It is calculated as the ratio of the area of the intersection between the predicted and ground truth segmentation to the area of their union.

For precision and recall, precision measures the proportion of true positives in the predicted segmentation, while recall measures the proportion of true positives in the ground truth segmentation, which are both useful metrics for evaluating the performance of the segmentation algorithm in detecting an object or region of interest. The F1 score is a weighted average of precision and recall that considers both metrics. It is calculated as the harmonic mean of precision and recall.

Finally, Average Precision (AP) is computed by sorting items or predictions based on their confidence scores, computing the precision and recall values for each item, plotting the precision-recall curve, and then computing the area under the curve using numerical integration. Specifically, the formula for AP involves summing up the product of precision and the difference between adjacent recall values for each item. The resulting value

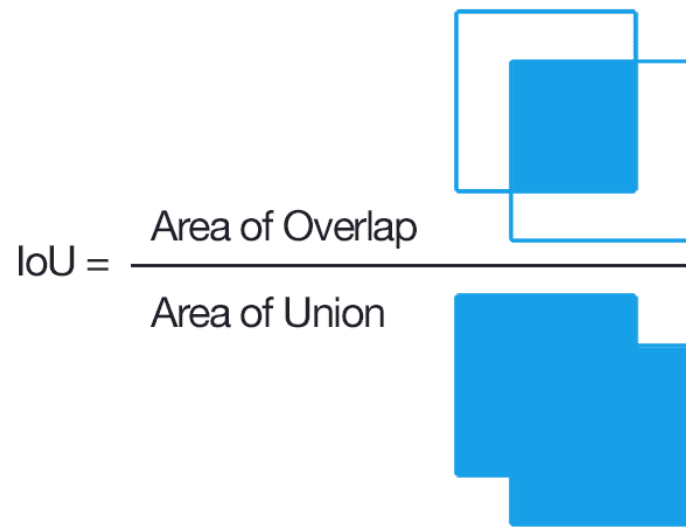


Figure 28. Illustration of IoU (pyimagesearch.com).

provides a measure of the quality of the ranked list of items or predictions, with higher values indicating better performance.

$$mAP = \frac{1}{N} \sum_{k=1}^N AP_k$$

Mean Average Precision (mAP) is therefore the average of the Average Precision (AP) values across multiple classes or categories. To compute mAP, AP is computed for each class separately using the steps outlined for AP. Then, all of the derived AP values are averaged to get the mAP. This approach is commonly used in object detection and image retrieval tasks, where there are multiple categories of objects or images that need to be ranked and evaluated. mAP provides an overall measure of the performance of the system across all categories and is a widely used metric for evaluating object detection models [43]. The mAP metric is popular because it's very simple to compare the performance of many models by only a single number. Thus, we chose mAP as the criteria for evaluating the outputs of the proposed deep learning method.

5.3.2 Result of Several Representative Cases

Figure 29 and Figure 30 depict the original fluoroscopy images, along with their corresponding segmented images, generated using a single deep neural network. The implants used in the images are PS and CR TKA, as seen in Figure 29a and Figure 30a, respectively, and the neural network model classifies the objects and predicts the corresponding masks for both femur (red) and tibia (blue), as illustrated in Figure 29b and Figure 30b. The masks of both femur and tibia are perfectly overlaid with the

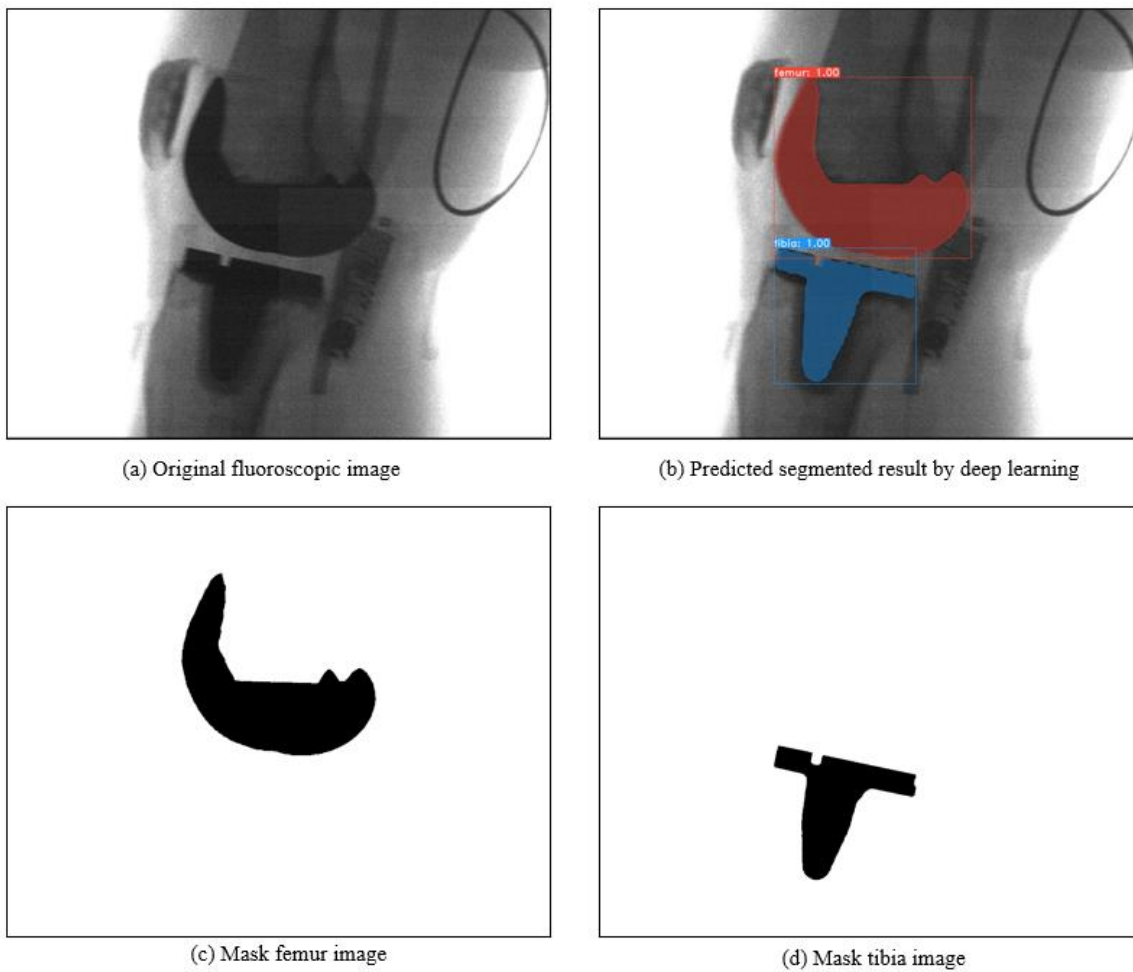


Figure 29. A sample of an original fluoroscopy image of the PS implant patient and the corresponding of the segmented images.

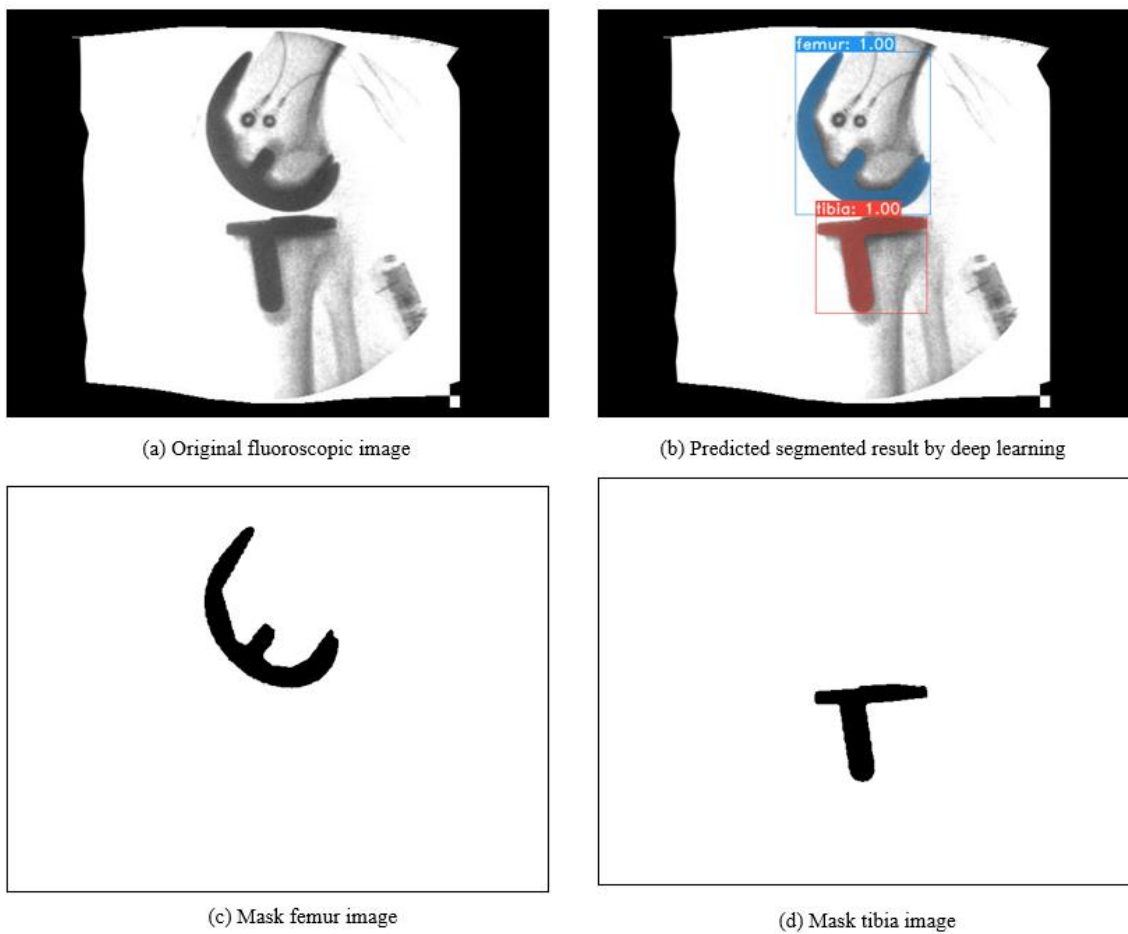


Figure 30. A sample of an original fluoroscopy image of the CR implant patient and the corresponding of the segmented images.

corresponding silhouettes, as evident from the visualization. The segmentation outputs are presented as mask images for femur (Figure 29c and Figure 30c) and tibia (Figure 29d and Figure 30d), which will be utilized in the later steps for automatically registering 3D CAD models to 2D fluoroscopy images. While there are no quantitative metrics calculated from Figure 29 and Figure 30, these qualitative results overall suggest the efficacy of the proposed deep learning approach in accurately segmenting the implants in fluoroscopy images.

5.3.3 Comparison with the Ground Truth

The results for the evaluation of the deep learning model for predicting the implanted knee are performed on 500 testing images with the known ground truths, presented in Table 1. A segmentation model with a higher mAP value indicates greater accuracy. The best three deep learning models for segmenting the COCO test-dev dataset [44] are Mask R-CNN with mAP=34.7 [45], PA-Net with mAP=36.6 [46], and MS R-CNN with mAP =38.3 [47]. The mAP for the segmented mask images of the implanted knee by deep learning compared to the ground truth was 88.83, while this number of predicting the 2D box was 97.57. These findings demonstrate the high accuracy of the deep learning model for segmenting the femur images compared to the ground truth, and the high-quality training dataset. Therefore, the mask of the femur obtained by deep learning can be considered accurate and can be utilized in the subsequent steps.

Table 1. mAP comparison between the ground truth and the predicted registered knee image.

IoU	all	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
box	97.57	100.0	100.0	100.0	100.0	100.0	100.0	98.96	98.87	97.59	80.33
mask	88.83	100.0	100.0	100.0	99.35	99.35	98.14	98.01	96.30	86.76	10.34

5.3.4 Comparison with Other Methods

The qualitative comparisons of the segmented results with manual and thresholding methods are shown in Figure 31 and Figure 32. The thresholding method uses the pixel color value to distinguish the darker silhouette of the femur and tibia metal implants from other components in the fluoroscopic image, resulting in a mask for both implants as shown in the second column of the figure. The manual approach uses Adobe Photoshop to detect the regions of the femur (red) and tibia (blue) separately, which is demonstrated in the third column of the figure. The results obtained through deep learning are presented in the fourth column, where the femur (red) and tibia (blue) are distinguished. In the regular testing cases presented in the figure, the results obtained through deep learning outperform those achieved using the benchmark manual methods. Furthermore, the proposed deep learning approach handles several cases that failed the benchmarks remarkably well, as discussed in the subsequent sections.

5.3.5 Challenging Cases: Blurry Images

In line with the introduction section, the predicted segmentation results are presented for challenging cases that can be occasionally encountered during the collection of fluoroscopic data and may pose a significant challenge for traditional segmentation methods. These challenging cases are shown in Figure 33 through Figure 42. Specifically, Figure 33 and Figure 34 depict the blurry case that results from patient movement. Typically, these cases contain only around a half number of the frames in each fluoroscopic video and are particularly difficult to segment using manual approaches due to the inability

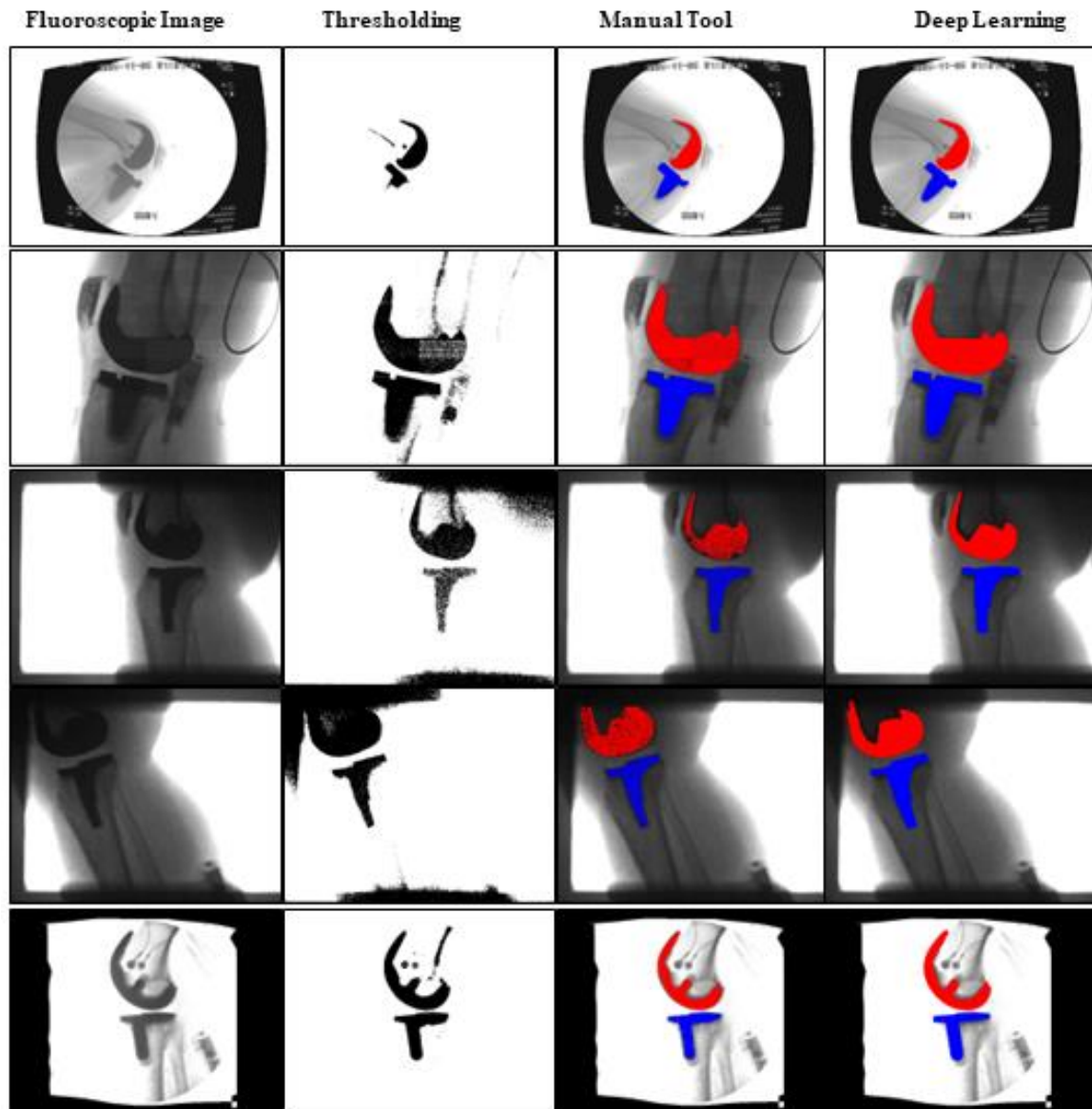


Figure 31. The comparison of the deep learning method with other traditional methods.

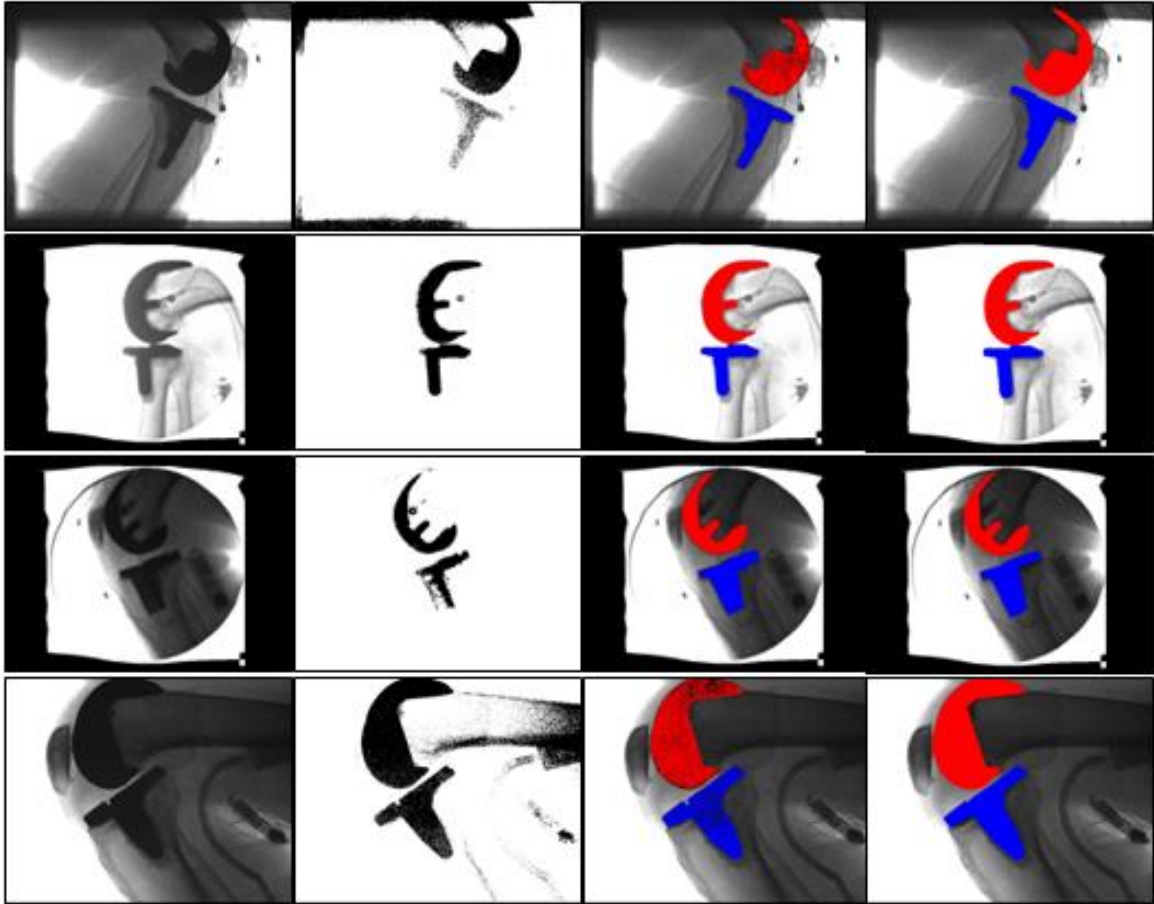


Figure 32. The comparison of the deep learning method with other traditional methods (continued from the previous page).

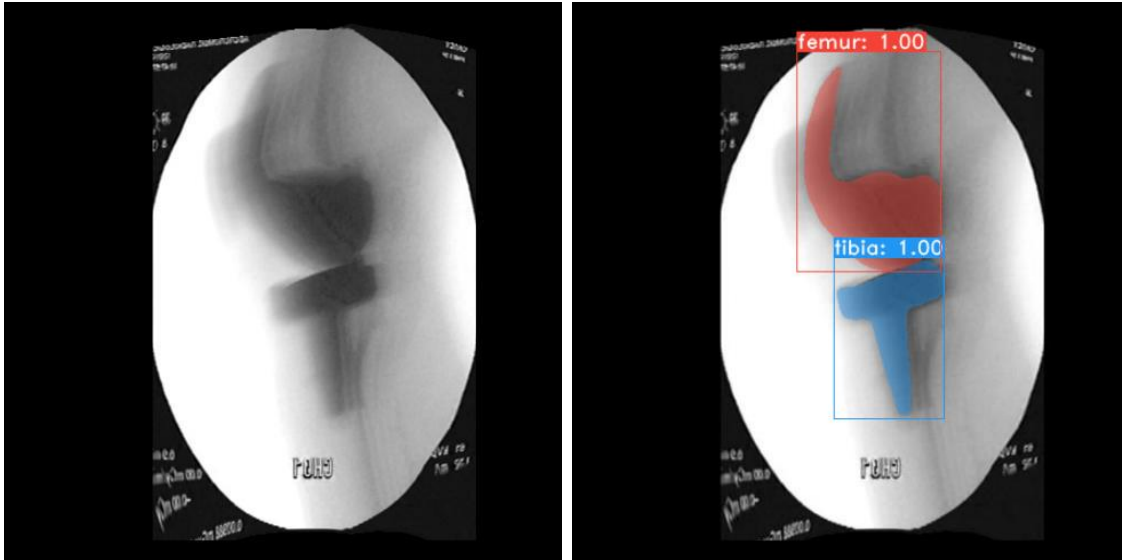


Figure 33. A challenging case of the blurry image (left) and the predicted segmentation result (right).

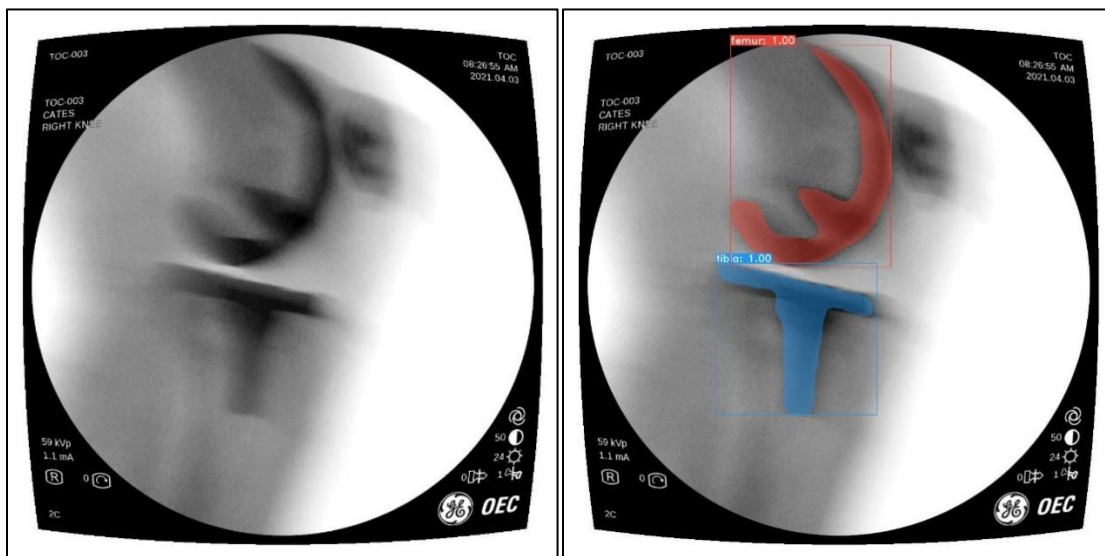


Figure 34. A challenging case of the blurry image (left) and the predicted segmentation result (right).

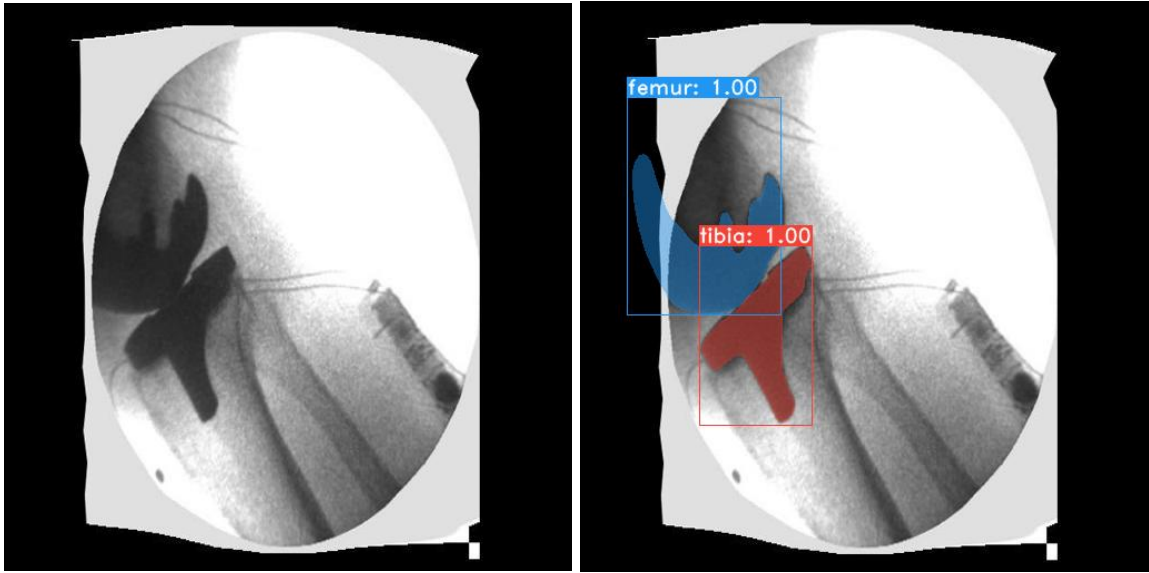


Figure 35. A challenging case of the out-of-captured range image (left) and the predicted segmentation result (right).

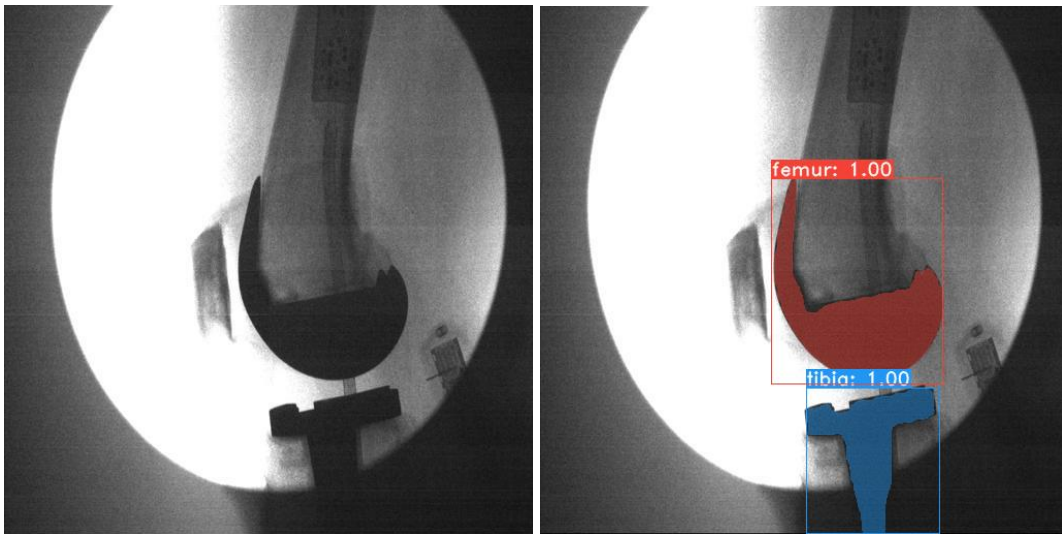


Figure 36. A challenging case of the out-of-captured range image (left) and the predicted segmentation result (right).

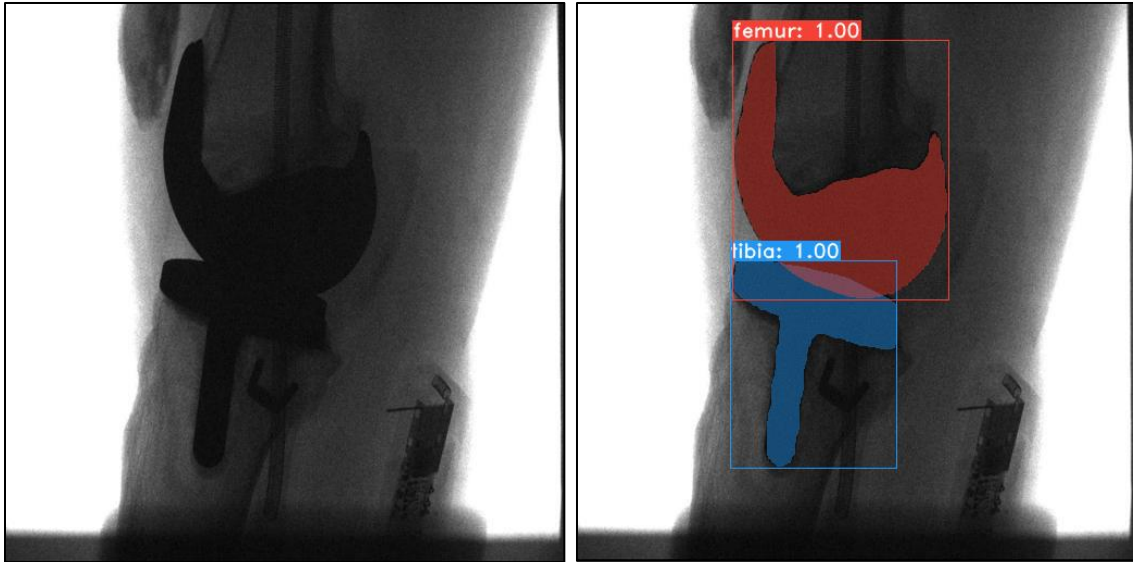


Figure 37. A challenging case of the overlap case image (left) and the predicted segmentation result (right).

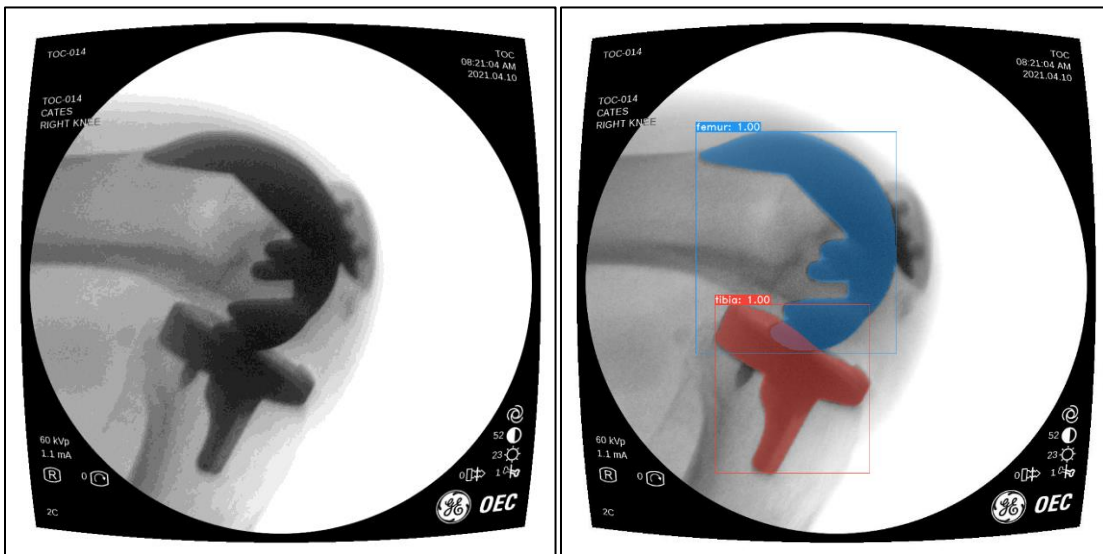


Figure 38. A challenging case of the overlap case image (left) and the predicted segmentation result (right).

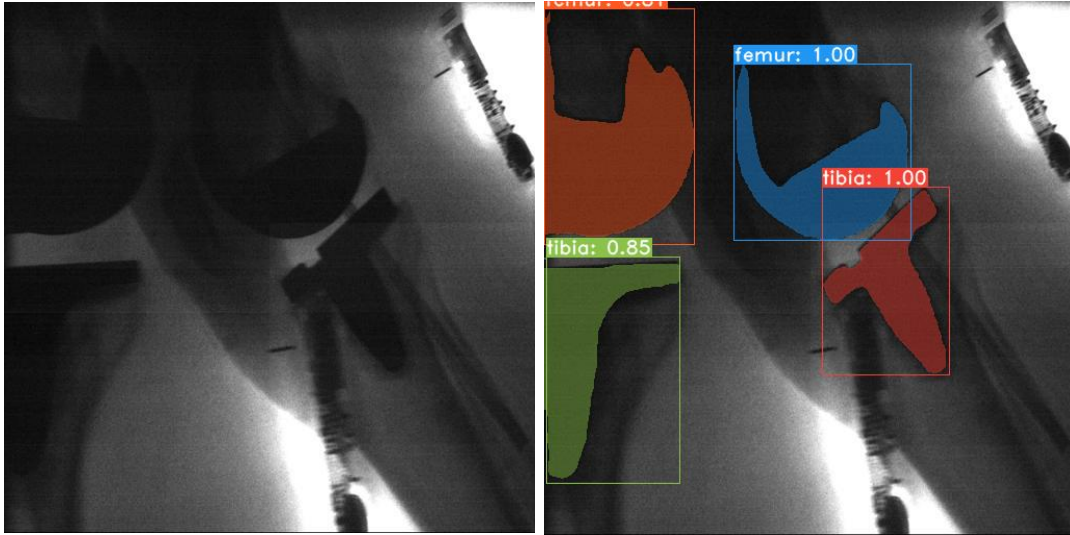


Figure 39. A challenging case of the multiple implants image (left) and the predicted segmentation result (right).

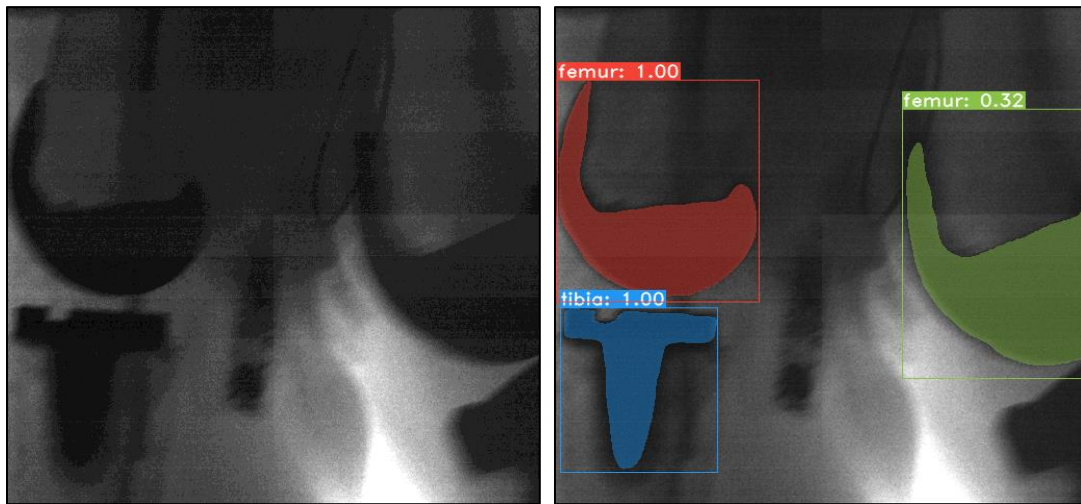


Figure 40. A challenging case of the multiple implants image (left) and the predicted segmentation result (right).

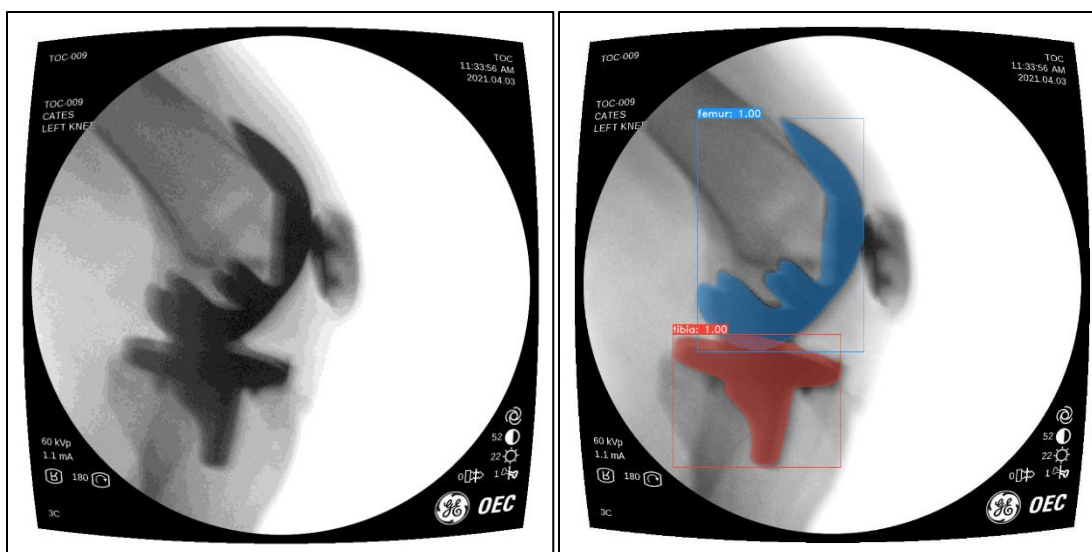


Figure 41. A challenging case of the image (left) with multiple challenges (overlap & blurry) and the predicted segmentation result (right).

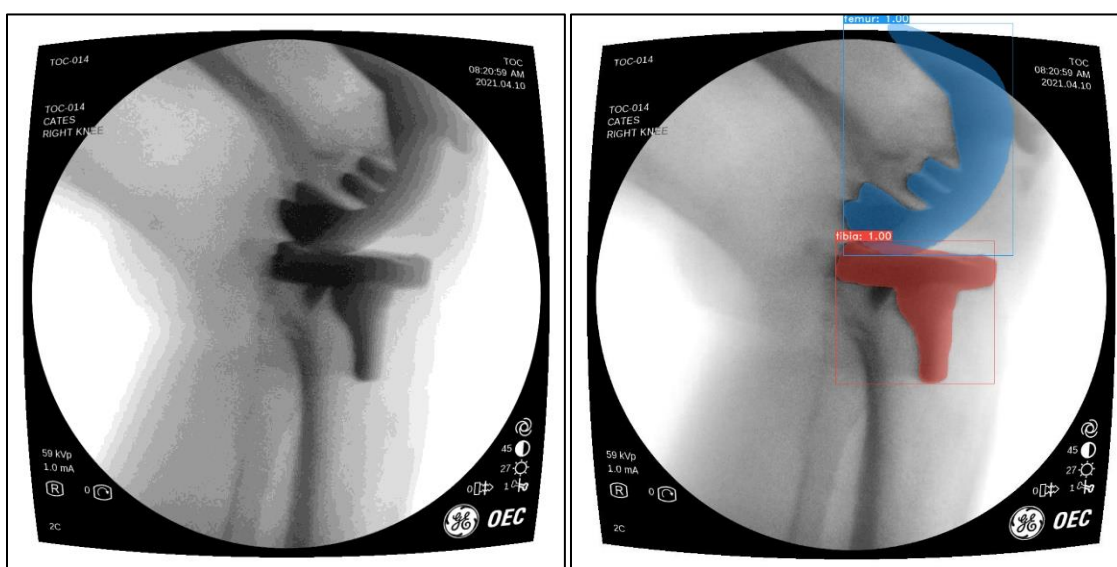


Figure 42. A challenging case of the image (left) with multiple challenges (edge & overlap & blurry) and the predicted segmentation result (right).

of human vision to detect the edge of the implanted models in the fluoroscopic images. However, the right images in both Figure 33 and Figure 34 demonstrate the remarkable effectiveness of our deep learning approach in segmenting blurry images.

5.3.6 Challenging Cases: Edge Cases Images

The out-of-captured range cases, which are often referred to as edge cases are shown in Figure 35 and Figure 36. In Figure 35, the femur falls outside the captured range of the camera, while in Figure 36, the tibia is not within the captured range. These cases can also occur during data collection, as the fluoroscopy unit is collimated, and the image intensifier is typically small to ensure patient and technician safety from radiation exposure. The results presented in these figures again demonstrate the ability of our deep learning model to not only detect masks within the image range but also accurately predict the shapes of both femur and tibia outside of the capture range. This capability of our model is particularly noteworthy as it enables the segmentation of even the most challenging cases, ultimately leading to more accurate and reliable results.

5.3.7 Challenging Cases: Overlap Cases Images

The overlap cases when two or more of the femur, tibia, and resurfaced patella overlap each other are shown in Figure 37 and Figure 38. Usually, the femoral and tibias TKA components are separated by a thin polyethylene bearing, so depending on the angle the patient's knee makes with the X-ray source, their silhouettes may overlap with each other on the fluoroscopic image, seen in Figure 37. Similarly, a metal resurfaced patella

can be considered noise for image segmentation when it overlaps with the femur in Figure 38. This all makes the segmentation more challenging when using the conventional threshold method. The results on the right images show that the deep learning approach can still predict the segmented regions of both femur and tibia for these difficult cases, as well.

5.3.8 Challenging Cases: Two Implanted Knees

Another challenging case where patients have both knees implanted leads to difficulties visualizing the knee implant of questions as shown in Figure 39 and Figure 40. Although fluoroscopic unit is used to capture a single implanted joint for evaluation, sometimes the fluoroscopic images can include two implants when the patient is a bilateral patient, as seen in the multiple implants case. Nonetheless, the right images in both figures again demonstrate the ability of our deep learning model to detect the segmentations of multiple implants in the image, showcasing the robustness of our approach.

5.3.9 Challenging Cases: Combinations

Finally, in Figure 41 and Figure 42, multiple challenges are evident when a combination of the aforementioned cases occurs. In each figure, the left image displays the original fluoroscopic image, while the right image exhibits the predicted segmentation result for both femur and tibia. Figure 41 poses a challenge with a blurry and overlapping image, while Figure 42 contains edge, overlap, and blurry cases, making them both difficult images to segment manually. Nonetheless, the segmentation results in the right images

demonstrate the exceptional performance of our deep learning model in accurately segmenting both femur and tibia. These results testify to the robustness of our proposed learning approach, especially when dealing with challenging images.

5.3.10 Application in Segmenting a Fluoroscopic Video

The manual registration approach can also be limited in its ability to register all frames of a fluoroscopy video. In other words, we typically only register at intervals of 0, 30, 60, 90, and 120 degrees. While this range can be expanded, it does require more manual work. However, the deep learning model offers a more robust and versatile solution, allowing for segmentation of every frame within a fluoroscopic video. In Figure 43, the machine learning model is demonstrated to be highly robust in segmenting all frames of a fluoroscopy video. It means that the method can capture every single degree of flexion. Due to the limitation of a visual layout, only 10 degrees increments of flexion are shown in Figure 43. This excellent performance makes the model highly promising for analyzing the kinematics of all frames in a fluoroscopic video. Moreover, the model's high processing speed suggests potential for real-time use.

5.4 Discussion

Automated image segmentation is an essential tool in 3D-to-2D image registration which can revolutionize the field by allowing for faster, more accurate analysis of images. These techniques can also be expanded to additional fields of medical imaging analysis, such as CT scan segmentation and Ultrasound analyses. While manual segmentation is time-consuming, prone to human error, and subject to inter-observer variability, automated

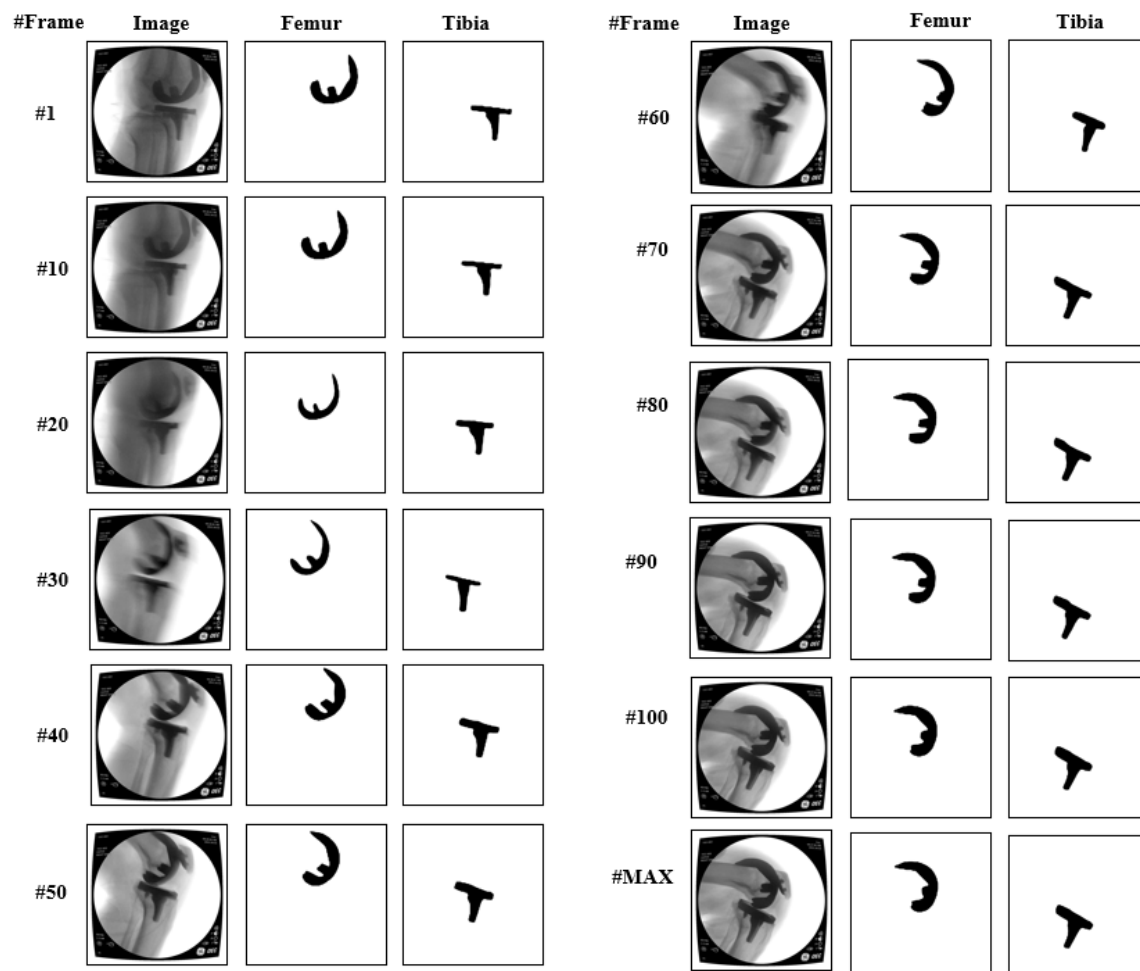


Figure 43. Application of deep learning in segmenting a fluoroscopic video.

image segmentation provides a more objective and reproducible analysis of medical images.

The various methods of automated image segmentation have both strengths and limitations. For example, thresholding is a simple method that is effective for segmenting objects with a well-defined intensity range, but it may not work well for objects with overlapping intensities or inhomogeneous backgrounds. On the other hand, machine learning-based methods can learn complex image features and can rapidly achieve state-of-the-art performance in various segmentation tasks. Thus, the proposed approach shows a fast and efficient deep neural network model for segmenting the implanted TKA images.

The proposed approach uses only a single deep neural network to create the masks of both femur and tibia with various implant types, such as PS and CR implants. Utilizing the registration dataset from the CMR lab, the segmentation dataset for training and testing purposes is created by using the voxelization projection [48] from the registration results. The simulation results show the great performance of the segmentation results by deep learning in comparison with the ground truth as well as the other benchmark methods. Additionally, the deep learning model is very robust by showing the great segmentation results by evaluating many challenging cases as well as via the entire frames of a fluoroscopic video that are impossible by doing manually.

While deep learning has demonstrated promising results in image segmentation, there are still several limitations to its application. One major challenge is the requirement for large amounts of labeled data to train the model, which can be both time-consuming

and expensive. Another limitation is the model's tendency to overfit the data, resulting in poor generalization for new images. Although techniques such as data augmentation can help address this issue, adding more training data has proven to be more effective in improving accuracy. Moreover, deep learning models may struggle with complex structures or variations in lighting, which can lead to segmentation errors. It should also be noted that perfect segmented results cannot always be achieved due to the training dataset containing a variety of implant models from different manufacturers. Despite these limitations, deep learning remains a powerful tool for image segmentation and will continue to be refined and improved through ongoing research and development efforts.

In summary, deep learning-based image segmentation has emerged as a game-changing technique in the field of medical imaging. It has brought significant improvements in accuracy, speed, and robustness compared to traditional methods, with potential applications beyond the knee joint. As medical imaging data continues to grow in quality and quantity, and deep learning algorithms continue to evolve, we can anticipate further breakthroughs and practical solutions for clinical problems. The successful implementation of deep learning-based image segmentation is a critical step in achieving automatic and efficient analysis of implanted knee kinematics through subsequent 3D reconstruction and 3D-2D registration.

Chapter VI

3D Reconstruction

This chapter pertains to the 3D reconstruction of a 3D CAD model that might be optional, if for instance study involves fluoroscopy of a TKA but the research does not have CAD models of the implanted components. Therefore, the 3D reconstruction step might be skipped if the implant CAD model available are obtained from the manufacturer. However, when the implant CAD model is not available, this chapter shows an optimal methodology to reconstruct it from four fluoroscopy images.

6.1 Introduction

3D-to-2D registration methodologies have been around for nearly 30 years [49, 50, 7, 5], where the purpose of this such methodologies is to register a 3-dimensional (3D) computer-aided design (CAD) model onto several 2-dimensional (2D) X-ray or fluoroscopic images to ultimately generate the in vivo weight bearing kinematics of a patient's implanted or non implanted joint [5, 22]. These techniques can be applied to a variety of activities to extract patient kinematics [51, 31]. Further using these kinematics and other patient parameters in combination with computational modeling, the kinetics of the implanted joints can be further computed and analyzed for pre- and post-operative surgical evaluation or to aid with improvements of the next implant generation design. These methods have successfully contributed to the improvement of both total knee replacement (TKA) [30, 52] and total hip replacement (THA) [53] systems, as well as other

implanted joints. To conduct an evaluation of an implanted joint, both fluoroscopic videos and the corresponding CAD models are required.

While 2D videos are easy to obtain by capturing the patient's joint motion via fluoroscopy, obtaining the correct corresponding CAD model can be more challenging. CAD models are typically obtained directly from the manufacturer, through 3D laser scanning, or through CT scan segmentation [54]. However, all three of these methods have limitations, and the CT approach seems particularly harmful to the patients due to the radiation exposure [55]. Thus, when CAD models are unavailable, reconstructing accurate 3D models from a limited number of fluoroscopy X-ray images (Figure 44) is a challenging task but important for both patient safety and data cohesiveness.

There are many current techniques for 3D reconstruction problems using X-ray/fluoroscopic images of the implant models. Previously, a technique using the rigid Iterative-Closest-Point (ICP) algorithm was used to reconstruct patient-specific 3D bone surfaces [56]. However, one major drawback of rigid registration techniques is that they are unable to reconstruct complex curves because rigid registration algorithms aim to solve translation, rotation, and scale information only. A modern technique uses statistical shape modeling (SSM) to register 3D knee joint model from single plane X-ray fluoroscopy using nonlinear shape priors [57], and it is also possible to utilize bi-planar fluoroscopic image registration and morphing, which can reduce error magnitudes but exposes the patient to additional radiation [58]. A recent study also used SSM to reconstruct lumbar spine by biplanar radiographs [59] with moderate success of 1.7 mm – 2.0 mm of average error

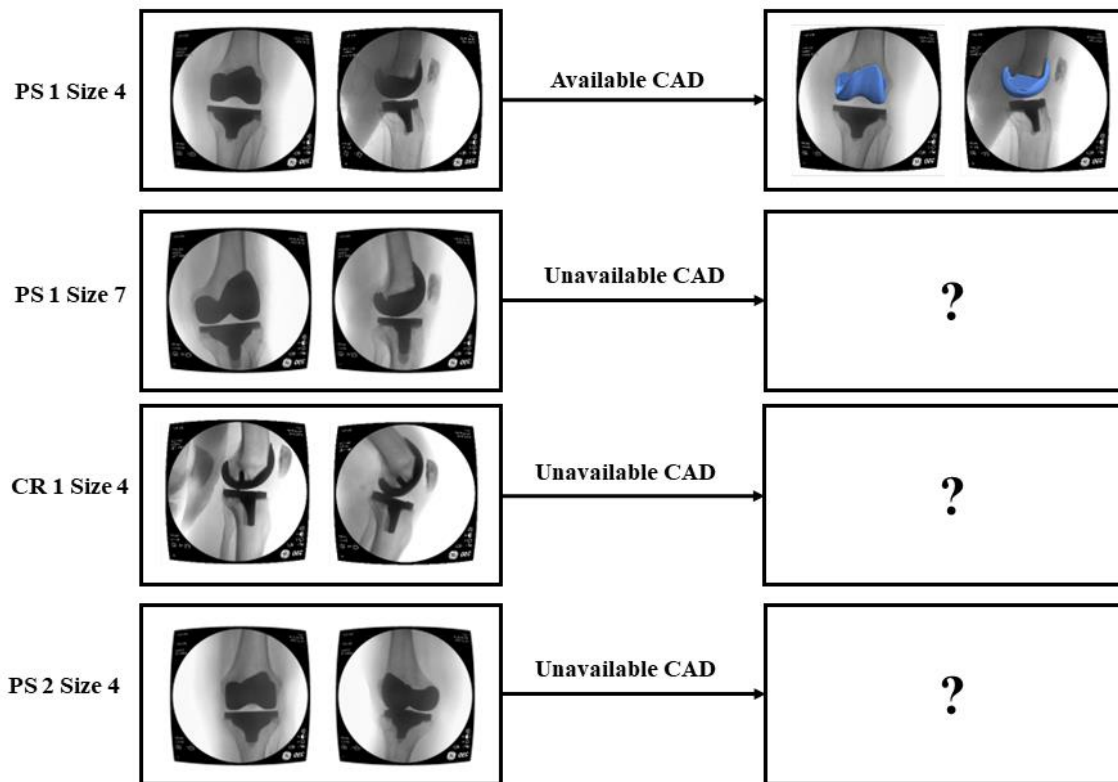


Figure 44. The proposed problem, wherein 2D-3D registration is needed for analysis, but no CAD models are available to generate results.

compared with the ground truths. Similarly, with the development of artificial intelligence (AI) techniques, a machine learning (ML) tool seems very powerful for handling 3D reconstruction of knee bones from bi-planar X-ray images using convolutional neural network [60]. Unfortunately, both the SSM and ML methods are logistically complex due to the requirements of massive training datasets that are not always available.

To overcome the above drawbacks of the current techniques, one of the main objectives of this study was to develop a non-rigid registration/morphing approach to reconstruct 3D implant component CAD models. Coherent-Point-Drift (CPD) is a non-rigid method that can morph and interact with complex curvatures [16, 61], and these algorithms can be combined with a limited number of 2D X-ray images to produce more accurate 3D model reconstructions. Utilizing this advantage, an enhancement of CPD is proposed as the core of this methodology to ultimately design a new 3D implant based on multiple 2D X-ray images.

6.2 Methods

The overall framework of designing a new 3D femoral implant CAD model from X-ray/fluoroscopy images is shown in Figure 45. In the beginning, the template model is scaled uniformly to the actual size of the designing model. The scaled template CAD model is then morphed onto the actual silhouettes of the implant on the X-ray/fluoroscopy images using iterative non-rigid registration algorithms. The following sections will describe in detail the steps for the environment setup, methodology, evaluation, and application of the morphing algorithm. With regard to maximum robustness, a single template CAD model

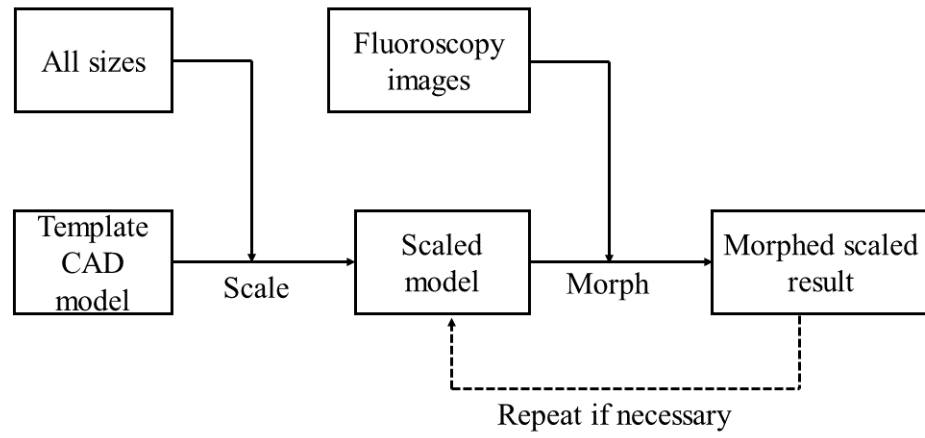


Figure 45. 3D TKA reconstruction framework.

is necessary, from which the rest of the catalog can be morphed, and this single CAD model can be easily created using 3D laser scanning.

Prior to data collection and algorithm implementation, to optimize the number of the fluoroscopy images needed for data collection, an analysis was conducted to determine the number of vertices that are "visible" to the virtual camera when rotating the template CAD model about a superior/inferior axis (Figure 46). This analysis determined that four positions (at 0, 45, 90, and 135 degrees) contain the highest number of vertices. Since many femoral implant models are usually not symmetric between the medial and lateral side, the above four fluoroscopic images provide the information on both medial and lateral articulating surfaces of the designing model for the morphing algorithm. Thus, the data collection set up shown in Figure 47 was used to capture four fluoroscopic/X-ray images, with the participant standing fully upright, which successfully optimizes data input and reduces radiation exposure for the patients.

After the fluoroscopic images have been collected and corrected for distortion [22], the next step of the proposed methodology involves multiple overlays of the template CAD model on to the four aforementioned fluoroscopic images [5]. However, these overlays will logically not fit perfectly because the template model is not the true CAD model (Figure 48). Nonetheless, the overlay is completed as accurately as possible, and the morphing algorithm then aims to transform these vertices to match with the silhouette of the fluoroscopy images as illustrated in Figure 49. Furthermore, the morphing algorithm aims to transform the template model's contour (initial bound) to the target contour (target

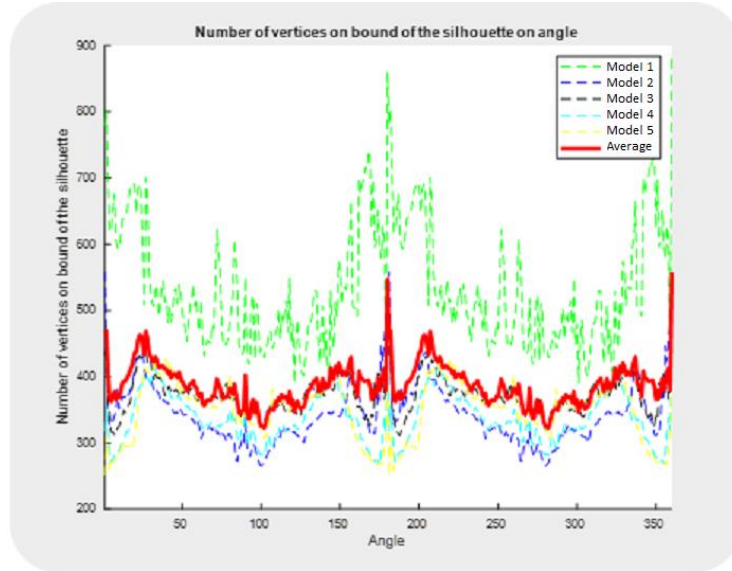


Figure 46. Evaluation of the number of vertices to determine the ideal positions.

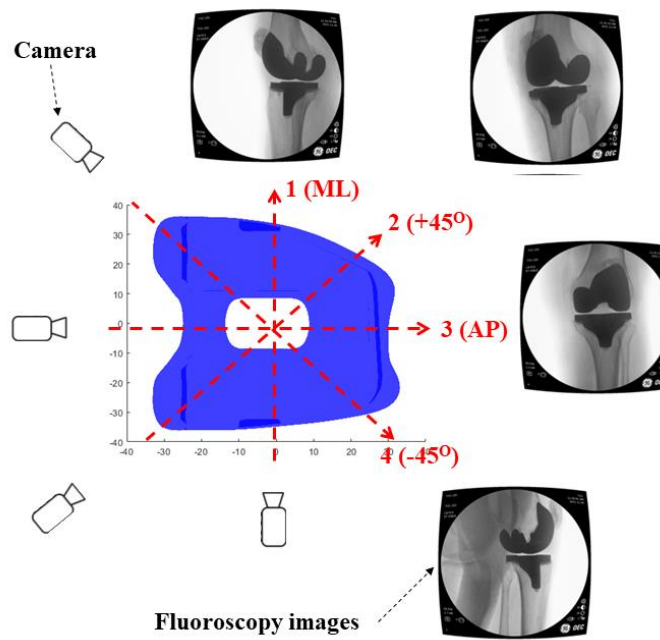


Figure 47. 3D TKA reconstruction framework with an appropriate number of fluoroscopy images to maximize data output and minimize patient risk.

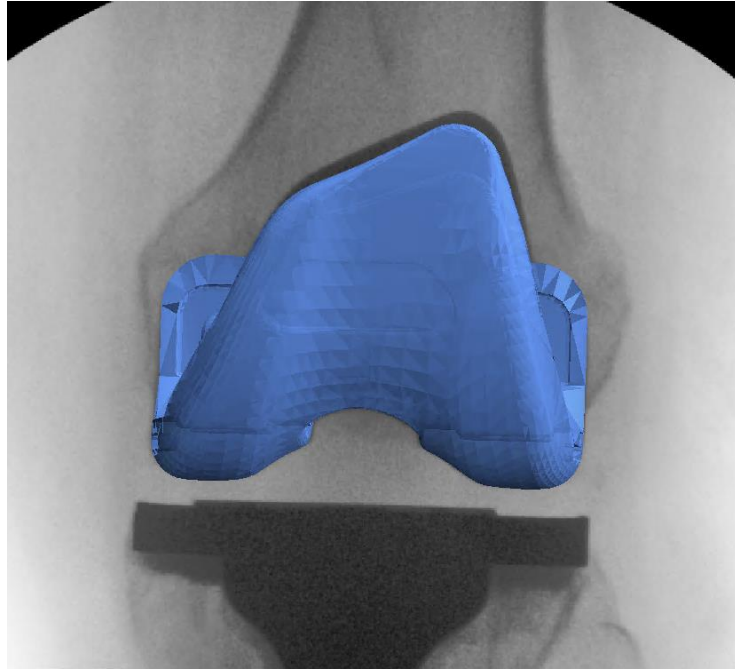


Figure 48. Registration of the template model with errors.

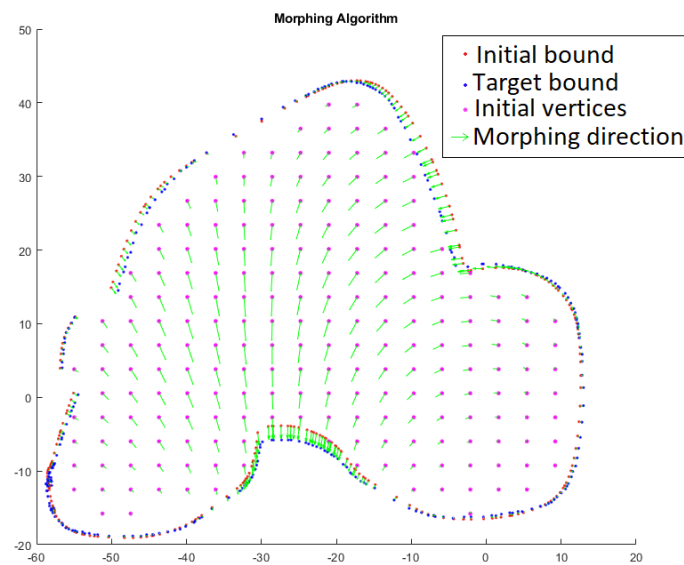


Figure 49. Schematic diagram showcasing the defined directions used for analysis of the kinematic variables.

bound) and keep the template model's vertices moving coherently. The next sections describe in more detail the processing of the CAD model and the fluoroscopic image to get those bounds.

6.2.1 Model Processing: Boundary Vertices Statistics

The goal of the model processing step is to find the bounding edge of the CAD model, represented by the vertices on the boundary of the CAD model, as seen in Figure 50 and Figure 51. The CAD model bounding edge is viewed as a silhouette in the simulation environment, thus allowing the bounding edge of the CAD model to be compared to the bounding edge of the fluoroscopy image. The vertex set of the CAD model is denoted as $Z_{P \times 3}$ and forms the yellow curves in Figure 50, where P is the number of vertices. The output of the model processing step are the boundary vertices $Y_{N \times 3} \subset Z_{P \times 3}$, shown as the red contour in Figure 51, where N is the number of vertices on this bound ($N < P$). This process is completed for all 4 registration images.

6.2.2 Image Processing: Edge Detection

Similarly, the goal of the image processing and edge detection steps are to find the contour of the knee in the fluoroscopy images as seen in Figure 52 and Figure 53. The fluoroscopy images contain unnecessary information such as bones, extra components, and other noise, as seen in Figure 52, that needs to be removed. This step can be completed with standard image processing tools due to the contrast level of the knee implant in comparison with other parts [62]. However, the quality of fluoroscopy images are not

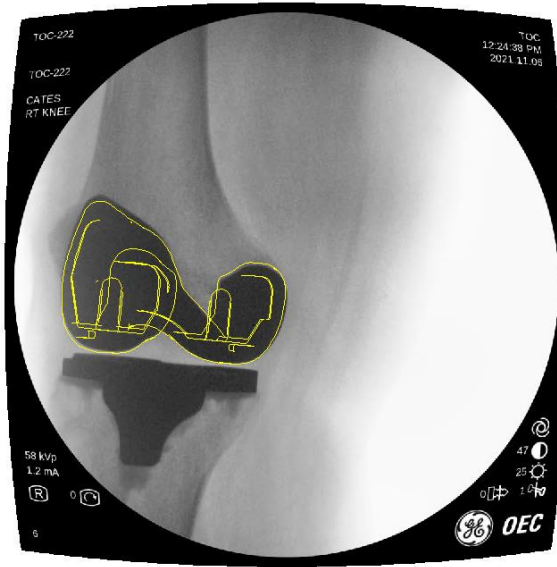


Figure 50. Boundary vertices and edge detection in the morphing algorithm: Model Processing
Input.

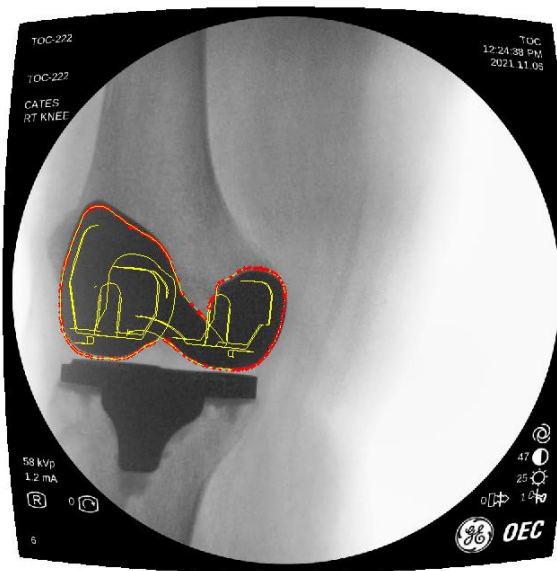


Figure 51. Boundary vertices and edge detection in the morphing algorithm: Model Processing
Output.

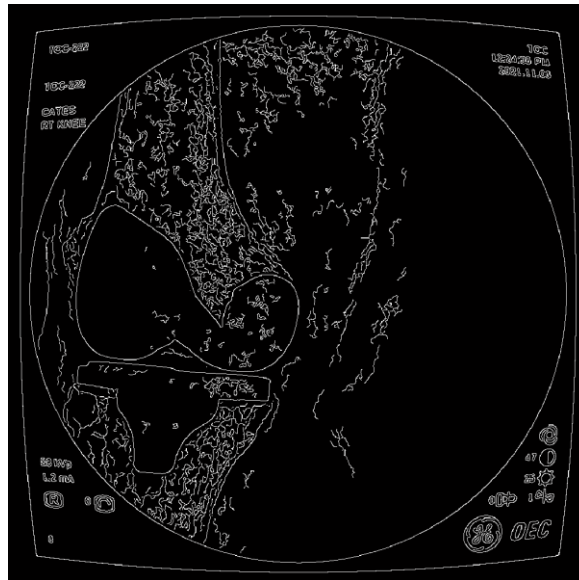


Figure 52. Boundary vertices and edge detection in the morphing algorithm: Image Processing

Input.

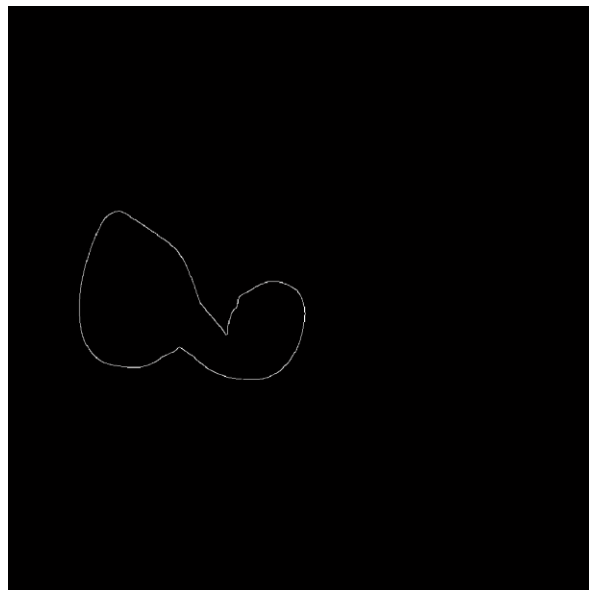


Figure 53. Boundary vertices and edge detection in the morphing algorithm: Image Processing

Output.

always clear enough for automated tools to work efficiently [5], so manual image processing can also be used to deal with low quality images. Once the image is processed to show only the outline/edge of the CAD silhouette, the output of the image processing step is a set of the white pixels as in Figure 53. Those pixels are reconstructed in 3D by uniformly searching the closet third coordination with $\mathbf{Y}_{N \times 3}$ to achieve the target bound of the CAD model, called $\mathbf{X}_{N \times 3}$. This process is completed for all 4 registration images.

6.2.3 Morphing Algorithm

After preparing the above data (template CAD model $Z_{P \times 3}$, initial bound $\mathbf{Y}_{N \times 3}$, and target bound $\mathbf{X}_{N \times 3}$), an Enhanced-Coherent-Point-Drift (ECPD) algorithm is proposed to design and morph a new CAD model based on the subjected requirements. As above mentioned, the conventional Coherent-Point-Drift (CPD) algorithm is a registration method that aligns two point clouds by iteratively optimizing a probability-based model to find the best transformation that maximizes the coherence between the points. While the standard CPD algorithm is able to morph between two bounds with the same 3D dimensions only, our ECPD can design a 3D model subjected to those two bounds. For each iteration, a section of the template CAD model is combined with the initial bound to morph to the target bound. At the end, the morphed model's vertices $\mathbf{R}_{P \times 3}$ is combined with the original template model's faces to be saved as a new CAD model. The details of the algorithm are presented as in Algorithm 1. To make the algorithm more robust, $D=3$ is the second dimension of the vertices set throughout the ECPD algorithm.

Algorithm 1 Enhanced-Coherent-Point-Drift (ECPD) algorithm

Inputs Template 3D model's vertices ($\mathbf{Z}_{P \times D}$), target bound ($\mathbf{X}_{N \times D}$), initial bound ($\mathbf{Y}_{N \times D}$)
Initialize the number of morphed vertices per iteration $M - N$ ($M > N$)
Computed the rate $\frac{M-N}{P}$ of the morphing algorithm
for $i = 1$ to $\lceil \frac{P}{M-N} \rceil$ **do**
 Combine initial bound with a section of $\mathbf{Z}$: $\mathbf{C}_{M \times D} = \{\mathbf{Y}_{N \times D}, \mathbf{Z}_{((i-1)(M-N)+1):(i(M-N)) \times D}\}$
 Initiate kernel matrix $\mathbf{G}_{M \times M}$, coefficients matrix $\mathbf{W}_{M \times D} = 0$, $\sigma^2 > 0$, $0 \leq w \leq 1$, $\beta > 0$, $\gamma > 0$
 while not convergence **do**
 Compute the posteriori probability distribution matrix $\mathbf{P}_{M \times N}$
 Compute $\mathbf{W} = (\mathbf{G} + \lambda \sigma^2 d(\mathbf{P}\mathbf{1})^{-1})^{-1} (d(\mathbf{P}\mathbf{1})^{-1} \mathbf{P}\mathbf{X} - \mathbf{C})$
 Compute $\mathbf{T} = \mathbf{C} + \mathbf{G}\mathbf{W}$
 Update $\sigma^2 = \frac{1}{D\mathbf{1}^T \mathbf{P}\mathbf{1}} (tr(\mathbf{X}^T d(\mathbf{P}^T \mathbf{1}) \mathbf{X}) - 2tr((\mathbf{P}\mathbf{X})^T \mathbf{T}) + tr(\mathbf{T}^T d(\mathbf{P}\mathbf{1}) \mathbf{T}))$
 end while
 Output of i -th iteration $\mathbf{R}_i = \mathbf{T}_{((N+1):M) \times D}$
end for
Output The morphed model's vertices $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{\lceil \frac{P}{M-N} \rceil}\}$, saved as a new CAD file
Smooth $\mathbf{R}$, if necessary

6.2.4 Model Refinement

To improve the accuracy of the morphing algorithm, the entire CAD model is not morphed as a single entity. Instead, the CAD model is divided into different sections to be morphed individually, per Figure 54, with different colors representing each section's vertices. In Algorithm 1, $\lceil \frac{P}{M-N} \rceil$ is denoted as the largest integer not greater than $\frac{P}{M-N}$, which represents the number of morphing sections, and in the end those sections are re-assembled after morphing. However, this leads to the rough surface at the intersection of the morphed sections (Figure 55). Laplacian smoothing methods [63] are then used to remove the rough surfaces of the morphed model. This is an optional step at the end of the morphing algorithm, only if necessary, and yields an improvement of the CAD model after smoothing, as seen in Figure 56.



Figure 54. An example of a worse morphed CAD model needed to be smoothed: Multiple sections in one model.

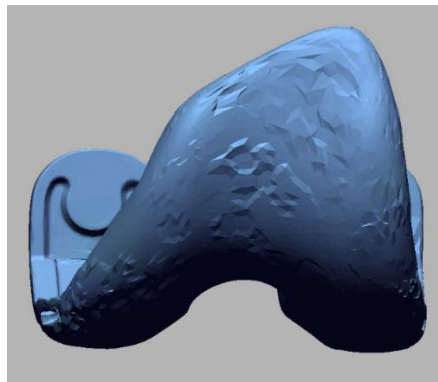


Figure 55. An example of a worse morphed CAD model needed to be smoothed: Before smoothing.

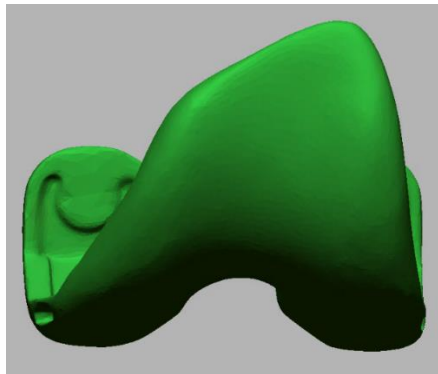


Figure 56. An example of worse morphed CAD model needed to be smoothed: After smoothing.

6.2.5 Posts Reconstruction

The purpose of the posts in knee implant design is to improve attachment with the bone, although not all knee implants feature them. Typically, CR implants include posts while PS implants do not. In our morphing algorithm, we excluded the posts because non-rigid registration can cause the posts to be positioned incorrectly due to differences in implant size, as shown in Figure 57. While kinematics primarily relies on the implant's outer surfaces, the posts may not significantly affect them. However, if desired, reconstructing the posts is a simple optional step that can be accomplished through rigid registration using the Iterative Closest Point (ICP) method, as demonstrated in Figure 57.

6.2.6 Model Comparison Evaluation

In the algorithm-evaluation study, the algorithm was used to morph into a smaller TKA template component model and compare it to multiple available known larger sizes. Within the scope of the evaluation for Algorithm 1, the actual larger CAD models were available for evaluation, and thus the resulting models from the morphing algorithm were compared directly to the known results. The morphed model and the actual model are aligned using ICP, then the error of each vertex of the morphed model is computed by its distance to the surface of the actual model. The accuracy of the proposed algorithm is analyzed by computing two types of errors: (i) the root-mean-square (RMS) error, and (ii) largest source error of all vertices.

6.3 Computational Results

Qualitatively, the results revealed that the new morphed model registered almost

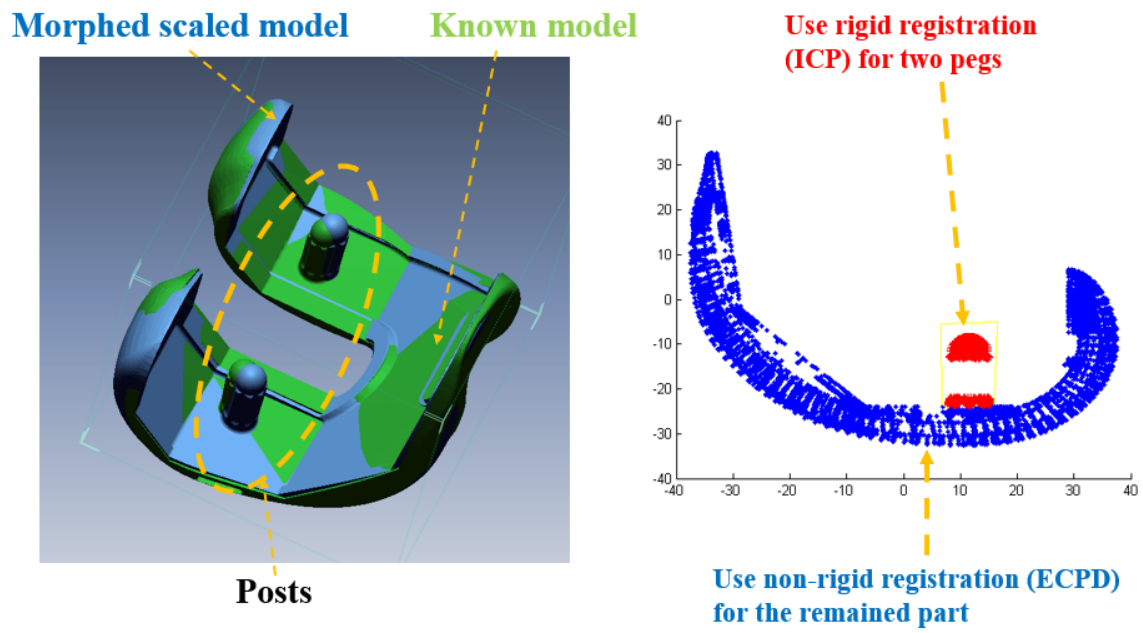


Figure 57 The wrong position of the posts after morphing (left) and the correction by using rigid registration (right).

perfectly with respect to the silhouette of the corresponding fluoroscopy images when compared to the unmorphed model, as seen in Figure 58. This comparison is evident in Figure 59 with four different views of a sample morphed model. Quantitatively, a comparison of the sample morphed model to the known target CAD model can be found in Figure 60. For the sample case, the RMS error was 0.43 mm, and the absolute largest source of error was 2.21 mm. Fortunately, the maximum largest error occurred on the internal surfaces of the model, as seen in Figure 60. Therefore, this is not considered a concern for future evaluations, as the outer articulation surfaces of the implants are of greater interest than internal fixation surfaces. In other words, while these surfaces can be valuable for implant fixation information, they are not articulating surfaces and therefore are generally not a concern for fluoroscopic studies. Similarly, the articulating surface of the J-curve shows substantially lower errors, on the order of -0.2 mm to 0.1 mm, which indicates that the articulating geometry will be accurate, as shown in Figure 61.

A series of 18 implant CAD models of different sizes from different companies were reconstructed using the proposed morphing algorithm. The results for seven morphed models are shown in Figure 62. The RMS error and largest source error compared between the morphed models and the ground truths are listed in Table 2. The average RMS error of those reconstructed models is 0.53 ± 0.12 mm in a range of 0.39 to 0.78 mm. The average of the largest errors is 2.78 ± 0.71 mm in a range of 1.64 to 4.37 mm, but again occurring on internal features of the implant. These errors magnitudes are acceptable for TKA kinematics analyses because 0.65 mm is the usual accuracy of 3D registration for 3D knee

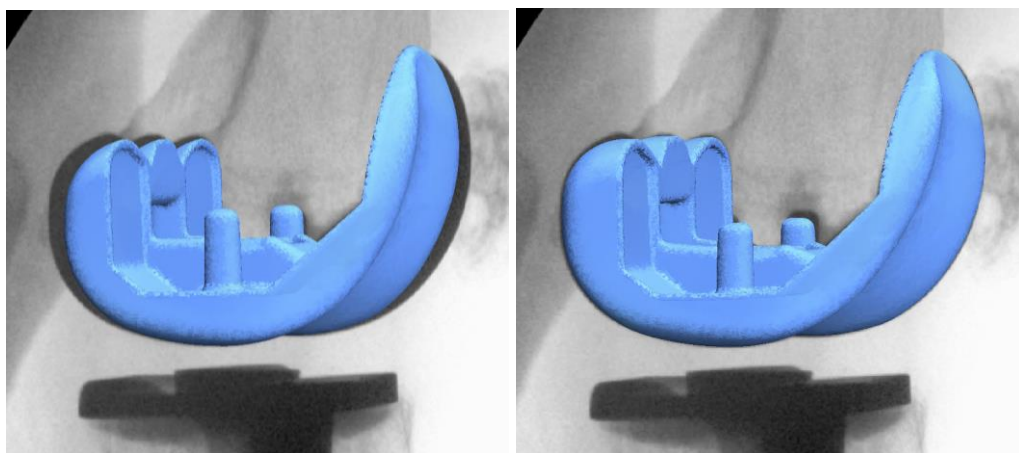


Figure 58. The results of overlaying the unmorphed model (left) and morphed model (right) onto the fluroscopic image.

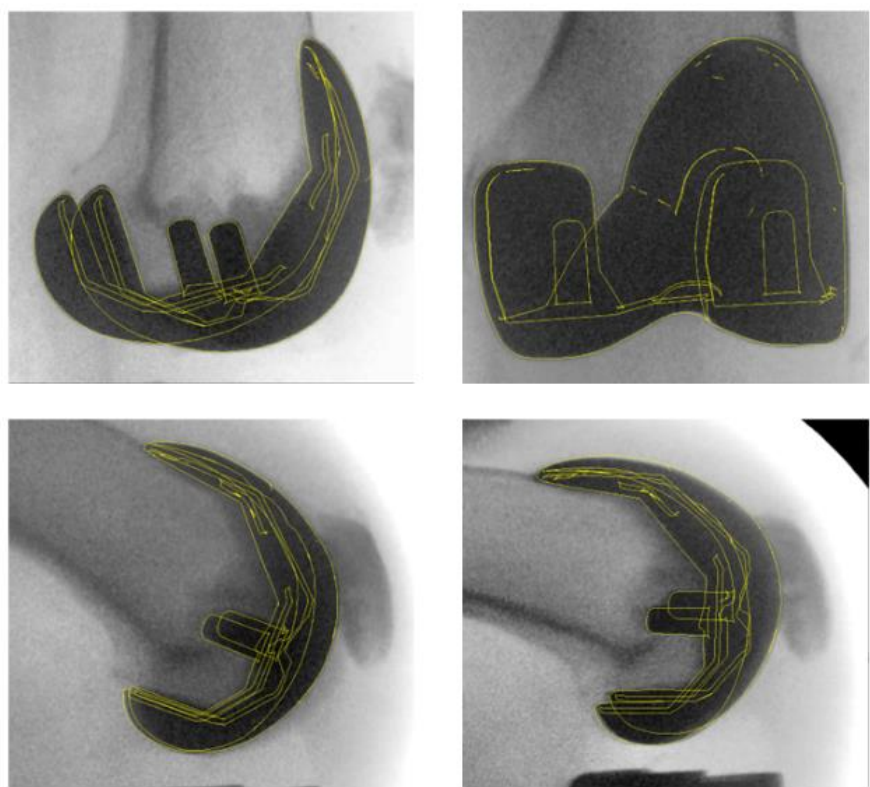


Figure 59. Morphed model overlays.

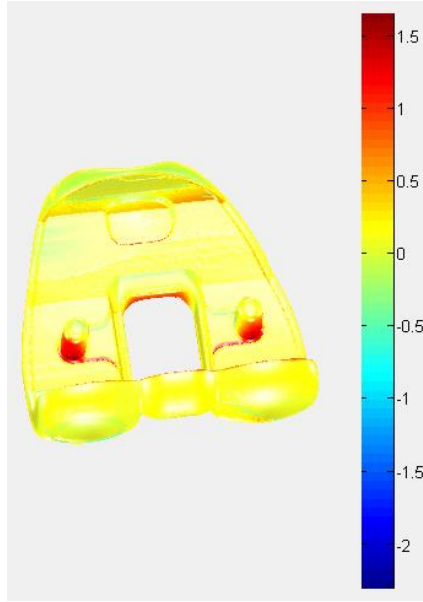


Figure 60. Heat maps show comparisons between the ground truth and the morphed model: Top view with largest error.

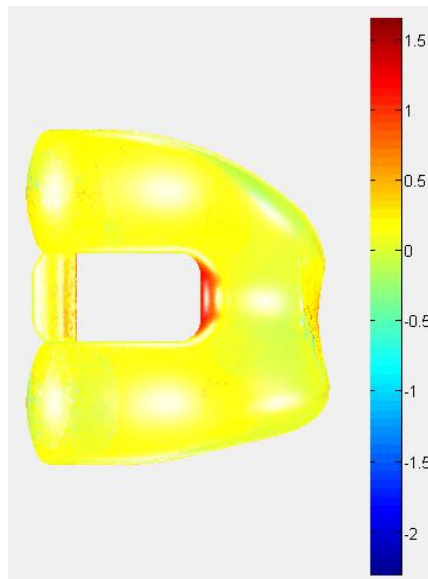


Figure 61. Heat maps show comparisons between the ground truth and the morphed model: Bottom view with smaller error articulating surfaces.

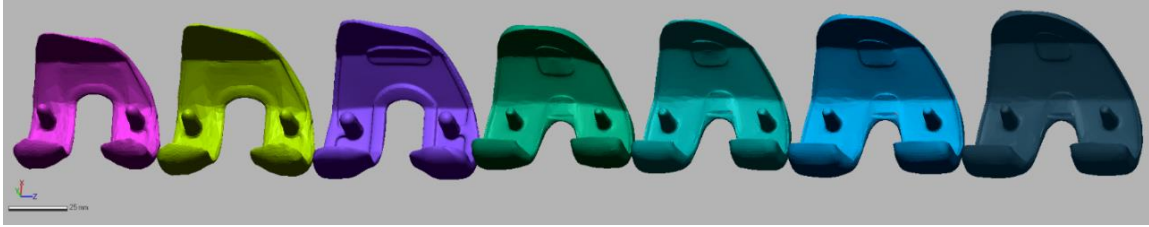


Figure 62. Several morphed models result.

Table 2. The evaluation of 18 reconstruction models.

Test Case	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	
RMS Error (mm)	0.49	0.49	0.57	0.50	0.68	0.41	0.42	0.42	0.63	
Largest Error (mm)	1.64	2.57	3.01	3.99	2.54	2.11	1.84	4.37	3.03	
Test Case	No. 10	No. 11	No. 12	No. 13	No. 14	No. 15	No. 16	No. 17	No. 18	Average
RMS Error (mm)	0.63	0.59	0.61	0.43	0.39	0.42	0.78	0.44	0.70	0.53
Largest Error (mm)	3.03	3.25	2.82	2.21	0.20	2.64	2.95	2.78	3.34	2.69

replacement kinematics evaluations and is often within the tolerances of the manufacturing process [64, 6]. Moreover, the largest sources of error consistently occurred on non-articulating surfaces, so the lateral anterior/posterior position (LAP) and medial anterior/posterior position (MAP) low point TKA kinematics are evaluated by the millimeter unit scale [65, 66, 29].

A comparison application of the implant morphing algorithm to recreate 3D femoral component catalogs of cruciate-retaining (CR) and posterior-stabilized (PS) implants is shown in Figure 63. With a template shown in grey on the left of Figure 63, the morphing algorithm outputs the other eight with the correct shape. For this mini-evaluation, the average RMS errors of CR and PS reconstructions are similar, 0.52 ± 0.10 mm and 0.54 ± 0.15 mm, respectively, as shown in Table 3. Even though the average RMS error of PS is higher than this error of CR, there is no significant statistical difference in the RMS errors between morphed CR models and morphed PS models ($p=0.347$). Thus, the proposed morphing algorithm can be used fairly for a variety of implant catalogs.

6.4 Discussion

In this chapter, a novel methodology was introduced to overcome common 3D-to-2D registration limitations by creating a family of 3D CAD models from a limited number of X-ray or fluoroscopy images. For recreating each implant CAD model, a morphing algorithm takes inputs of four fluoroscopic or X-Ray images and a template CAD model. After processing the images and the template model, the morphing algorithm uses an enhanced non-rigid registration algorithm to output a reconstructed model. The comparison

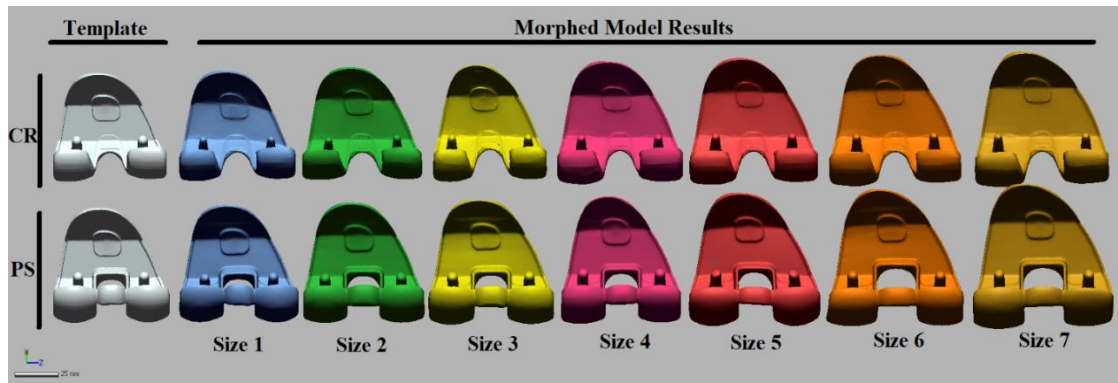


Figure 63. Implant morphing algorithm applications.

Table 3. Cruciate retaining and posterior stabilized TKA reconstructions evaluations.

Cruciate Retaining TKA	CR 1	CR 2	CR 3	CR 4	CR 5	CR 6	CR 7	Average
RMS Error (mm)	0.68	0.41	0.42	0.42	0.63	0.59	0.50	0.52
Largest Error (mm)	2.54	2.11	1.84	4.37	3.03	3.25	3.99	3.02
Posterior Stabilized TKA	PS 1	PS 2	PS 3	PS 4	PS 5	PS 6	PS 7	Average
RMS Error (mm)	0.61	0.43	0.39	0.42	0.78	0.44	0.70	0.54
Largest Error (mm)	2.82	2.21	0.20	2.64	2.95	2.78	3.34	2.42

between the reconstructed model and the ground truth shows excellent results, especially with low errors on the articulating surfaces.

The limitation of the proposed morphing algorithm is a long-running time proportional to the number of vertices of the designing model because the template model is divided into multiple sections to be morphed. While a CAD model of 2,000 vertices requires around 30 minutes to be designed, another CAD model of 8,000 vertices might take 8 hours, subject to the use of Core i7-8700 CPU @3.20 GHz. This drawback might be handled by using GPU to compute parallel morphing sections. Another limitation of the proposed algorithm is the availability of the template CAD model. If the template CAD model is similar to the ground truth, the morphed model will have smaller errors compared to the ground truth, and vice versa. Thus, a CR template model is recommended for designing a CR model instead of using a PS template model, and vice versa. Finally, the accuracy of the morphing algorithm relies on the initial overlay, so the reconstructed results might be slightly different due to small changes of the initial registration.

While many existing morphing algorithms focus on reconstructing the bones [56, 59], the proposed morphing methodology is novel and offers many advantages over machine learning or SSM methodologies for designing the implant knee CAD models. Bone reconstructions present new difficulties due to the diversity of the bone shapes, but the proposed morphing algorithm maintains the J-curve design of the articulating surfaces for the TKA models evaluated herein [67]. The evaluations revealed that the reconstructed implant CAD models meet a strict requirement of high accuracy for the implant's

kinematics analyses. Designing 3D models using machine learning often requires a high-performance computer [60], yet the proposed methodology is simple and can be efficiently implemented with the lower-performing computers according to Algorithm 1. Similarly, reconstruction methods using SSM often require huge data sets [57, 68], but the proposed methodology is possible with a single template model only. Also, fluoroscopic studies are presently limited to those supported by manufacturers so studies involving multiple TKAs for a large subset of surgeons have not been conducted because researchers cannot gain access to the required CAD models. In the future, there will be no limitation to the scope of fluoroscopic studies.

The proposed algorithm has demonstrated high levels of accuracy with multiple potential extended applications. While the evaluation focused on the femoral component model, our algorithm is robust enough to be extended to additional joints, and this algorithm can also be extended to bones to eliminate the need for CT scan segmentation. Also, this morphing algorithm guarantees convergence with a feasible solution because the proposed non-rigid algorithm can morph with any kind of shape from the initial CAD model to the target images. Having the ability to morph models for any joint can be beneficial for both designing a new model and revising an existing implant.

Additionally, this algorithm can also be extended to bones to eliminate the need for CT scan segmentations. While the results from scanning and segmenting 3D models from CT is undeniably impressive with an average error of 0.3 mm compared with the ground truth [69], the errors from our proposed algorithm are not substantially higher, only around

0.53 mm. However, a CT scan of a metal implant such as a TKA is not only difficult but also time-consuming due to the manual segmentation procedure. Metal scatter from the CT scan is a cause for concern regarding segmentation. Thus, the proposed methodology increases the speed of 3D model generation in comparison with CT scans. Moreover, the radiation of a CT scan for a knee is around 0.1 mSV [70] while taking four fluoroscopic images for the morphing algorithm is only around $0.005 \times 4 = 0.02$ mSV [71], yielding a smaller amount of radiation exposure for patients compared to traditional CT scans.

Creating more personalized knee or hip implants for total joint replacements also becomes possible using the proposed methodology, and personalized designs can theoretically improve surgical outcomes [72]. By looping between kinematics analyses, combining these methodologies with theoretical mathematical modeling [73], and then iteratively adjusting the articulating surfaces, the proposed morphing methodology has the potential to find the best customized implant shapes for each individual patient. This methodology would help companies save both in cost and in time when designing and testing implant models. In the long-term development of the proposed methodology, the CAD model is expected to be no longer needed for fluoroscopy studies.

In summary, this methodology can be a useful tool when applied to TKA design and evaluation given the high accuracy of 3D model outputs and removes all limitations for future studies that can involve multiple implants and numerous surgeons. These algorithms demonstrate the potential for detailed implant evaluations given minimal initial information, specifically CAD models, and can therefore dramatically improve the

robustness of 3D-to-2D image registration methodologies. Our algorithms also have many different extended applications, including other fields of engineering, software development, and computer vision. While this advancing technology is applied successfully to design the orthopedics implant model, it can be further extended to contribute to creating 3D objects virtual reality (VR) and augmented reality (AR) simulations that are becoming popular trends in orthopedics education and surgical practice.

Chapter VII

Automated 3D-to-2D Registration

After getting the 2D segmented image from Chapter V and 3D model from the Chapter VI. This chapter aims to show an automated approach to the 3D-to-2D registration problem, which is a proposed advanced step for deriving joint kinematics under in vivo conditions.

7.1 Introduction

Artificial Intelligence (AI) and Machine Learning (ML) are advancing technologies that support and enhance countless aspects of our lives, including autonomous vehicles [74], facial recognition [75], chatbots, fraud detection, operating rooms [76], and more. These technologies can also help reduce the workload and improve overall efficiency by automating tasks that have typically been done slowly and manually [77]. One such example is 3D-to-2D image registration, which has existed for over 30 years and is still the “gold standard” for in vivo orthopedics analyses, helping determine postoperative total knee arthroplasty (TKA) and total hip arthroplasty (THA) kinematic outcomes [4], as seen in Figure 64 and Figure 65.

Unfortunately, current techniques for 3D-to-2D registration typically rely on manual overlaying, which is extremely time-consuming and labor-intensive, especially when conducting larger studies on hundreds of participants [78, 79, 80, 81]. Thus, the application of machine learning for automating 3D-to-2D registration is potentially extremely impactful.

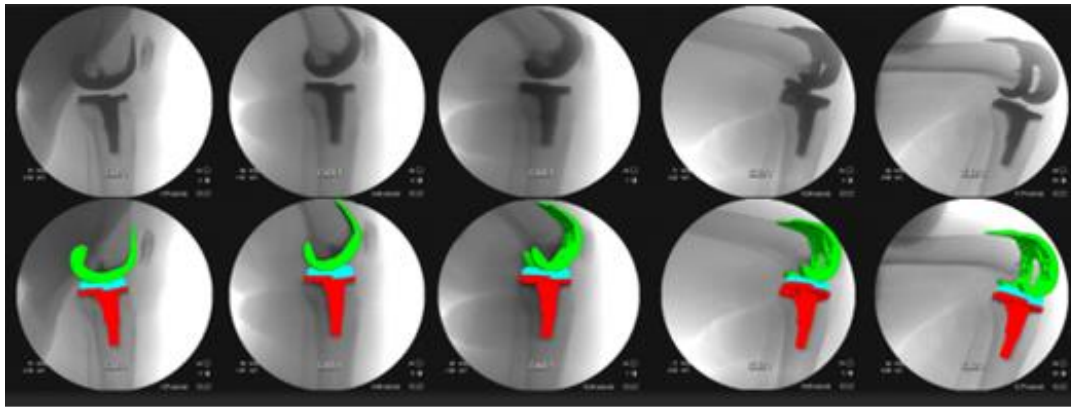


Figure 64. A sequence of fluoroscopic images for a DKB from full extension to maximum flexion at 30° increments is show in the top row, and the same sequence of images with the respective CAD models overlayed (after completion of the 3D to 2D image registration process) is shown on the bottow row.



Figure 65. Filmstrip showcasing the implant during DKB activity from full extension to maximum flexion.

Before the explosion of ML, many “non-learning” algorithms were used to try to solve the automated 3D-to-2D registration problem, such as edge-based optimization, intensity-based optimization, feature-based optimization, etc. Intensity-based registration has been used for registering both spine models [82] and femur/tibia models [17]. Also, a brain model is registered by using the intensity-based algorithm [83]. 3D-to-2D registration of non-implanted femurs was achieved using edge-based registration [84], and 3D-to-2D registration of TKA femur components was previously completed using edge-based registration via Lipschitzian optimization techniques [85]. Hence, it is evident that optimization methods have been extensively employed in addressing the 3D-2D registration problem in the past.

However, even though these non-learning optimization approaches have been successful in the past and have yielded adequate accuracy, they unfortunately still require either a good initial starting pose or human interventions during the process. Regarding our specific 3D-to-2D image registration technique, optimization methods have proven effective when applied to high-quality images, such as the left image depicted in Figure 66. However, due to the dynamic nature of fluoroscopic videos capturing in vivo movement of subjects, as illustrated by the right image in Figure 66, some fluoroscopic frames may be blurry and challenging, leading to a significant decrease in registration accuracy. Therefore, this can make both manual approaches and conventional optimization methods more difficult to analyze these cases.

There are many examples of using deep learning (DL), a popular branch of ML, for registering the 3D models within a specified space, with very high performance [86, 26, 87]. However, the applications of these methods in orthopedics are currently still limited. A non-implanted hip model was registered onto biplanar X-ray images using DL, but this approach requires the CT scan of the hip bone model [88]. The segmented images of the implanted femur and tibia are processed from the fluoroscopic image by using two convolutional neural networks, and then they are used for estimating the initial pose by Fourier descriptor-based estimation [89]. However, the drawback of Fourier descriptor-based method is their sensitivity to variations in object boundary, such as irregularities and noise. Even though deep learning models can handle substantial data very well, their accuracy is always good enough for analyzing in vivo TKA kinematics [89]. There are many reasons for this drawback such as the necessity of massive training data, variations of fluoroscopic imaging conditions, expensive hardware for training.

While there have been many attempts at automating these processes in the past, to our knowledge, all such attempts have fallen short because they are too unstable or have low accuracy. Fortunately, it is evident that combining the strengths of both non-learning and learning approaches has the potential to address historic limitations and overcome the drawback of each methodology in isolation. Additionally, most previous approaches have been done by using standard datasets such as Linemod [90], Occlusion [91], and T-LESS [92]. Our paper demonstrates a more targeted approach by using orthopedic-specific data. Thus, utilizing the large database of fluoroscopy images and manually registered

transformations owned by the Center for Musculoskeletal Research (CMR) lab, we propose a novel machine learning framework to register three-dimensional computer-aided design (CAD) implant models automatically and accurately onto two-dimensional X-ray or fluoroscopic images.

7.2 Methods

In the past decades, CMR has collected and manually registered more than 10,000 TKA images, including both femurs and tibias, which yields a large data set that can be used for algorithm training and can also be used for algorithm accuracy evaluations. Therefore, utilizing this data set, a novel methodology with two main phases is proposed for automatic 3D-to-2D TKA registration. In summary, the algorithm starts with all models oriented in their initial zero position, with no translations and no rotations. The predicted initial pose is then computed using a deep learning neural network, and from there the initial pose is refined using optimization techniques. This framework is demonstrated in Figure 67.

7.2.1 Camera Calibration

Prior to implementing any automated algorithms, a virtual environment for the camera must be accurately set up to mimic the actual fluoroscopy environment, as seen in Figure 68. It is important for the intrinsic and extrinsic camera parameters to be calibrated so that the computer 3D space is identical to the actual fluoroscopic space. Camera calibration in fluoroscopy for TKA patient involves the process of calibrating the fluoroscopic imaging system to ensure accurate and precise 3D-to-2D registration of the

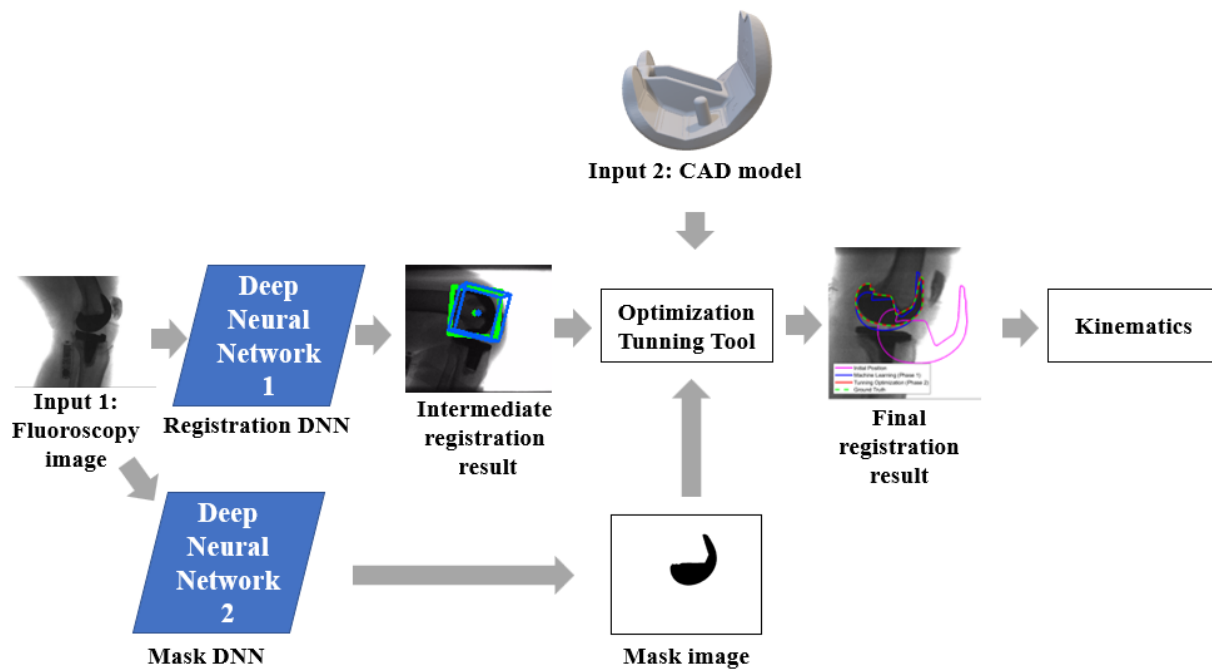


Figure 67. 3D-2D registration methodology using deep learning.

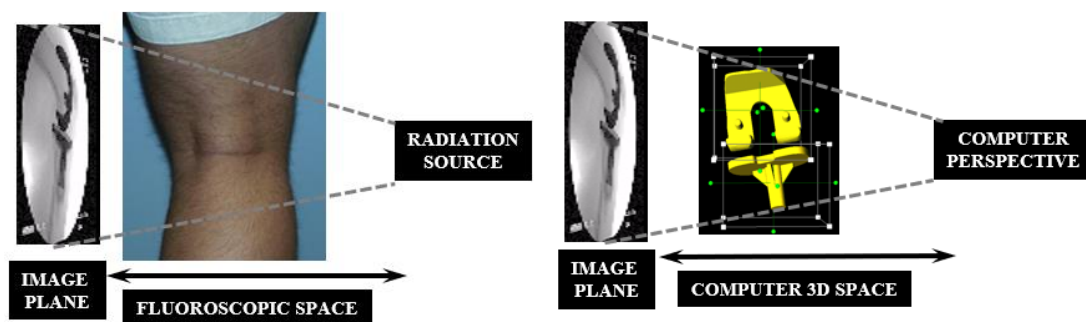


Figure 68. Virtual camera setup (left) to simulate the fluoroscopy (right).

images. The calibration process involves several steps. First, a calibration beadboard is used to unwarp and remove any fish-eye distortion that the fluoroscopy unit may inherently introduce. The calibration beadboard typically consists of a flat, planar surface with highly precise markers placed in known positions throughout. The fluoroscopic imaging system captures images of the phantom from various angles, allowing for the calculation of specific 3D-to-2D registration parameters, such as the focal length of the camera, the position and orientation of the camera relative to the beadboard, and the distortion characteristics of the imaging system [93].

While the intrinsic parameters are discussed in the 4.2 Current Processing Methodology: Distortion Correction section, this section takes care of the extrinsic parameters. One of the most important camera parameters is the camera matrix P [93], which is a 3×4 matrix. The matrix P is important because it maps from world points X to image points x by $x=PX$. According to [93], P is decomposed as

$$P = K[R \mid -R\tilde{C}],$$

where R is 3×4 rotation matrix, $\tilde{C}$ is the camera location vector, K is the form of

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where $(x_0; y_0)$ is the coordinate of the principal point, s is the skew number (usually 0), α_x and α_y are the scale factors in the x and y coordinate directions respectively (usually $\alpha_x = \alpha_y$ for the square pixel of the image).

Camera calibration in fluoroscopy for TKA is a critical step in ensuring the accuracy and precision of the kinematics analysis. During this period, a unique camera setup was employed, as depicted in Figure 69, with a fixed camera distance of 533.4 mm from the fluoroscopy image plane. Additionally, the camera view angle between two parallel edges of the image was consistently set at 0.22 radians. Thus, the camera matrix P in this setup is a constant matrix, called P_{Karbs} .

$$P_{Karbs} = xX^{-1} = \begin{bmatrix} 0.5773 & 0 & 0 & 0 \\ 0 & 0.5773 & 0 & 0 \\ 0 & 0 & -0.0011 & 0.5773 \end{bmatrix}$$

Due to this setup, the image was often scaled down, with one side of the image equivalent to 118 pixels, which $533.4 \times \tan \frac{0.22}{2} \times 2 \cong 117.8 \cong 118$ (mm). For example, an original fluoroscopy image size of 1024×1024 (mm) would be scaled down by a factor of ten.

To standardize all current and future fluoroscopy images for a robust machine learning algorithm, each image in the past dataset needs to be converted into a new standard camera setup. Specifically, the new setup was calculated to translate from Figure 69 to the new setup in Figure 70 by three simultaneous steps. First, the coordinate system is translated on the z-axis by 533.4 (mm) to move the camera position to the original point (0;0;0), which is standard methodology [93]. Secondly, the image plane is translated on the z-axis with a new distance of $533.4 \times \text{actual_image_size}/118$ so that the fluoroscopy image is now representative of its actual size in the data collection room. After those two steps, the camera matrix P becomes a new value P_1 which is computed by Direct Linear Transformation (DLT) method [94]. Thirdly, the coordinate system is rotated around the

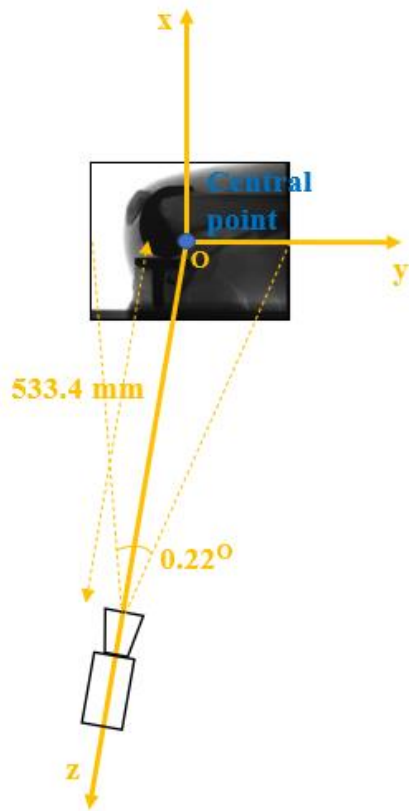


Figure 69. Camera setup in the data collection system (Karbs).

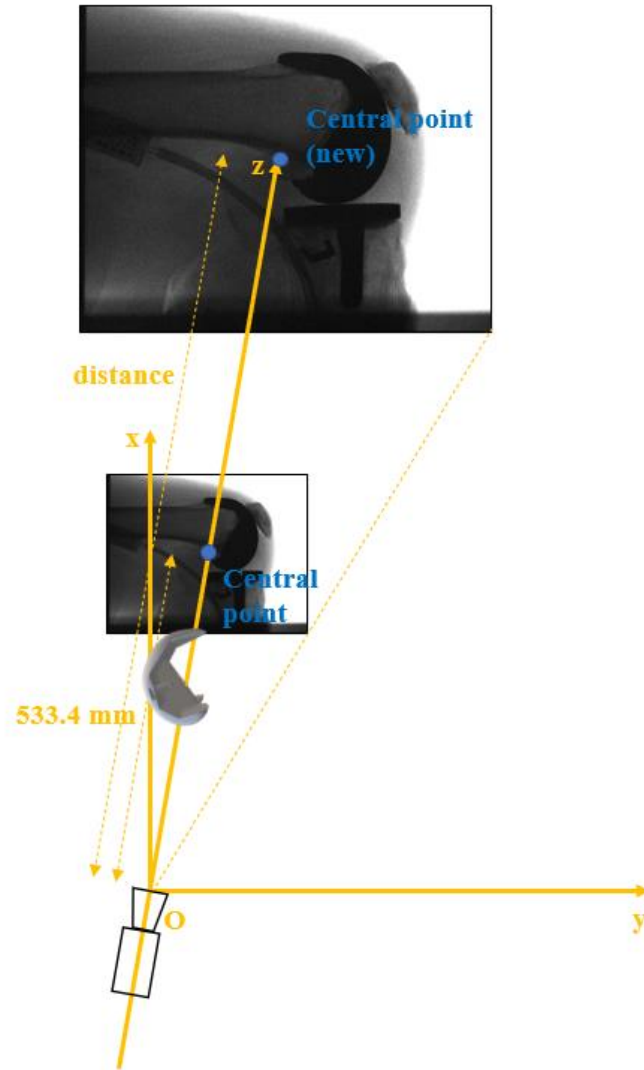


Figure 70. Translated camera setup in the data preparation for training of the deep learning model.

original point to make the camera viewpoint toward the positive direction [21]. To prevent the complexity of computing, the fluoroscopic image is flipped in the second step to change the axis direction.

This approach allows for not only the opportunity to use the original fluoroscopic images for algorithm development purposes but also analyze the kinematics with the original fluoroscopic images or video reference frames. During the above changes, the corresponding 3D implant model with the fluoroscopic image will be fixed to maintain its transformation. At this point, both the fluoroscopic images and the camera parameters have been corrected and standardized and are ready to be used in the next machine learning method. Based on the above changes of the camera position, while the old camera matrix is decomposed as $P_1 = K_1[R_1 | -R_1\tilde{C}_1]$, here

$$K_1 = \begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\tilde{C}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The new intrinsic parameter matrix P_{new} is computed as follows.

$$P_{new} = K_{new}[R_1 | -R_1\tilde{C}_1],$$

where K_{new} is computed as

$$K_{new} = \begin{bmatrix} \alpha_x & 0 & \frac{x_{image}}{2} \\ 0 & \alpha_y & \frac{y_{image}}{2} \\ 0 & 0 & 1 \end{bmatrix},$$

where x_{image} and y_{image} are the image's length and width. With this approach, the camera matrix is customized to image sizes and maintains the original correct images for analysis. By using this camera calibration matrix, the next section shows the approach to create the training data for the deep learning model using the voxelization projection.

7.2.2 Deep Learning Model: Training with Fluoroscopic Images Data

With the above environment setup, the machine learning transformation data is standardized and collected for 6835 TKA femur subjects. The distributions of those transformations of tx and ty show that most of the translations of the x and y are around (0,0), called in-plane difference, because the TKA object appears within the fluoroscopic image. However, the range of z translations is broad due to the depth of the TKA object, called out-plane difference. Similarly, the distributions of the rotations show that most of the rotations of x and y are around (0,0), while the rotation of z is like linear. This makes sense because most of the TKA femur fluoroscopic data is the medial-lateral (ML) images [27, 95]. The transformations obtained from the previously conducted manual registration are considered as the ground truths for this dissertation. The entire database is then divided into two subsets of 80% for training purposes and 20% for testing purposes.

Using 80% of the transformation for the training dataset, we generated the training data in the Linemod format [87]. Because the known transformations, fluoroscopic images,

and the CAD models are already available within the database, the remaining necessary data is the segmented image masks. As discussed in Chapter V, those segmented images are usually generated by using machine learning [96, 12] or images processing tools such as thresholding [34], edge-based segmentation, region-based segmentation [33]. In this research, we incorporated deep learning and the voxelization method [48]. Using this methodology, the resultant segmented images, called mask images, provide the final necessary data for model training.

A deep learning neural network (DNN) model is used for classifying and registering both the femur and tibia [87], as outlined in Figure 71. This DNN takes both fluoroscopic images and TKA CAD models as inputs and then outputs the predicted TKA poses. This DNN is trained by supervised registrations between the fluoroscopic images and the TKA CAD models. After training, the DNN can be used for automatically and quickly predicting the poses of new TKA CAD models for future subjects.

The original fluoroscopic images are used as inputs to train a deep neural network model based on the open-source EfficientPose [87], which is an extension of the EfficientDet model [97]. The model predicts the class (femur or tibia), 2D bounding box, and transformation (rotation and translation) of the corresponding CAD model from the direct fluoroscopic images. Training employs the backpropagation method, accelerated by GPU acceleration multiplication, on a Dell XPS 17 9700 equipped with 16 Gb RAM and NVIDIA GeForce GTX 1650 Ti GPU. The training process takes approximately 5 days to complete with a batch size of 2 over 500 epochs.

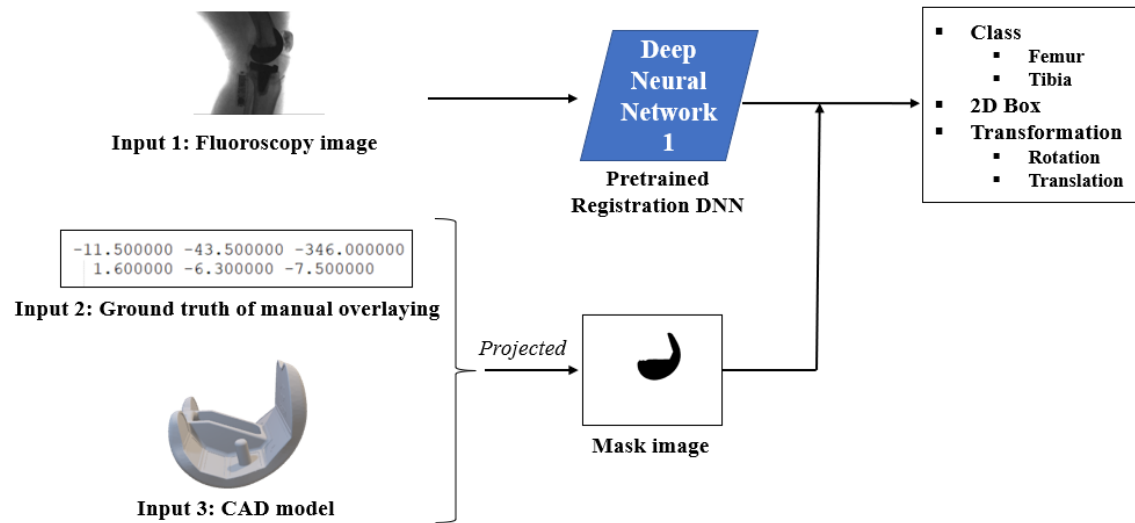


Figure 71. Training step for deep learning neural network structures with fluoroscopic data.

7.2.3 Deep Learning Model: Training without Fluoroscopic Images Data

In certain cases, the training approach mentioned above may not be feasible if the dataset is unavailable. For instance, the introduction of an entirely novel implant model which has never been analyzed before could cause errors in the learning model, and the training data may not cover the entire transformation domain, which can result in output prediction bias. As a result, an alternative deep learning approach without training data is proposed, as illustrated in Figure 72. The only difference between the two approaches is the image input. The fluoroscopic data approach uses the fluoroscopic image as input for the DNN, while the mask image approach employs the mask image as input for the DNN. The two approaches have similar structures, which are based on the EfficientPose model [87].

The segmented images are utilized to train a deep neural network model based on the open-source EfficientPose [87], a derivative of the EfficientDet model [97]. The model predicts the class (femur or tibia), 2D bounding box, and transformation (rotation and translation) of the corresponding CAD model from the segmented images. Training employs the backpropagation method, accelerated by GPU acceleration multiplication, on a Dell XPS 17 9700 equipped with 16 Gb RAM and NVIDIA GeForce GTX 1650 Ti GPU. The training process consists of 500 epochs with a batch size of 2. The training time of this approach is significantly increased due to the larger dataset in comparison to the previous section. This is primarily attributed to the quantization of the transformation domain of 6 degrees of freedom (DoFs), which encompasses translation and rotation.

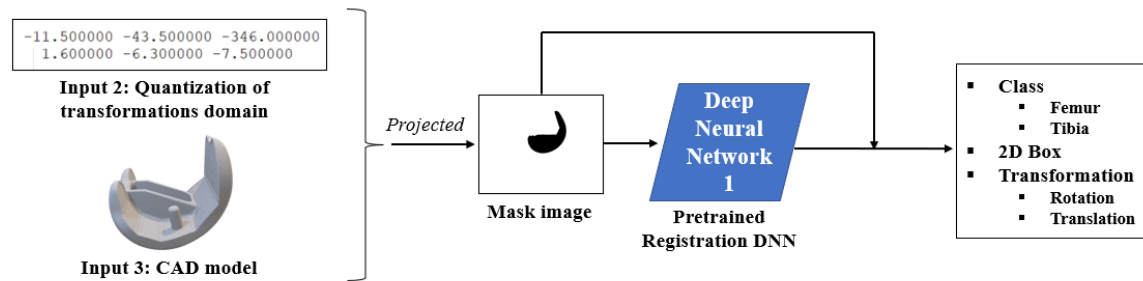


Figure 72. Training step for deep learning neural network structures without the fluoroscopic data.

7.2.4 Tuning Optimization: Edge-based Registration

The DNN's outputs may not always be optimal due to the diversity in implant models and sizes, as well as low-quality fluoroscopic images in the training data. To improve the results, the DNN's outputs are then used as initial poses that are, then refined using a tuning optimization method in the framework outlined in Figure 73. While the intensity-based registration method has high accuracy with fluoroscopic data [17], it is slow to converge due to numerous local minimums. In contrast, the edge-based registration method converges faster. To find the best solution within an appropriate running time, the optimal approach is to combine the advantages of both methods, as illustrated in Figure 73. The edge-based approach takes in two inputs: the mask image and the 3D model, which are the outputs of the preceding chapters. The objective of the optimization process is to find a six Degree-of-Freedom (DoF) transformation that aligns the projection of the 3D model onto the 2D image with the object silhouette in the mask image. The optimization problem involves six variables, including three translations and three rotations, with the constraint being the initial pose obtained from the deep learning results discussed earlier, whereas the objective function is derived as follows.

$$I(x) = c_1R + c_2(1 - R)Q + c_3(1 - Q),$$

The variables R and Q represent the boundaries of the 3D model projected onto the 2D image and the dilated edges of the segmented fluoroscopic image, respectively. Thus, the optimization problem will be written as follows.

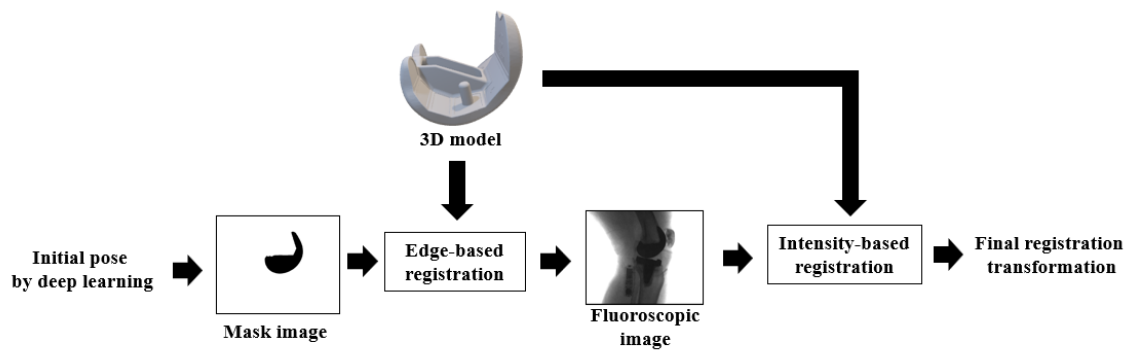


Figure 73. Tuning optimization framework.

$$\min_x I(x)$$

subject to

$$x_0 - \varepsilon \leq x \leq x_0 + \varepsilon$$

Here, x_0 denotes the initial position, while ε represents the search range.

7.2.5 Tuning Optimization: Intensity-based Registration

Upon obtaining the result using the edge-based registration approach, the registration outcome will be in close proximity to the global point. However, to further enhance the result, it is finally passed through an intensity-based registration method. The optimization problem used in this approach is similar to that of the edge-based approach, but the objective function differs. The objective function of the intensity-based method is a combination of normalized mutual information and gradient differences [98, 99], as shown below.

$$I(x) = MI + GD$$

Where Mutual-Information (MI) and Gradient-Difference (GD) are represented as follows.

$$MI = \sum_{x,y} p(x,y) \cdot \log \frac{p(x,y)}{p(x) \cdot p(y)}$$

$$GD = \sum_{i,j} \frac{\sigma_v}{\sigma_v + (D_v(i,j))^2} + \sum_{i,j} \frac{\sigma_h}{\sigma_h + (D_h(i,j))^2}$$

$$D_v(i,j) = \frac{\partial I_X(i,j)}{\partial i} - s_v \frac{\partial I_D(i,j)}{\partial i}$$

$$D_h(i,j) = \frac{\partial I_X(i,j)}{\partial j} - s_h \frac{\partial I_D(i,j)}{\partial j}$$

Here, σ_v and σ_h were the variances of the vertical and horizontal gradient images of fluoroscopic images, respectively.

7.3 Results

7.3.1 Comparison with the Ground Truths

Figure 74 displays the registration results of the femur or tibia, represented by a 3D box surrounding the object, for both a clear and a blurry fluoroscopic image. The 3D yellow boxes illustrate the registration results achieved through the proposed methodology, which successfully automatically registered 3D TKA CAD models to 2D fluoroscopic images. Several examples with different flexions are depicted in Figure 75. The green boxes represent the ground truth achieved through manual registration, while the blue boxes are the predicted poses generated by machine learning, which are in close agreement with the manual overlaying based on visual inspection.

By applying optimization techniques, Figure 76 demonstrates a substantial improvement in the registration results. Although deep learning produced a registration result that closely resembled the fluoroscopic image silhouette, optimization resulted in a pose of the implanted femur that matches the silhouette more accurately. The entire registration progress of the proposed methodology is shown in Figure 77, with the component position after the ML phase shown in blue bound projection and the final component position after optimization shown in red. When compared to the known final position obtained via manual registration (shown in green), the algorithm's predicted pose is clearly good but still has room for improvement. These values need to be improved by

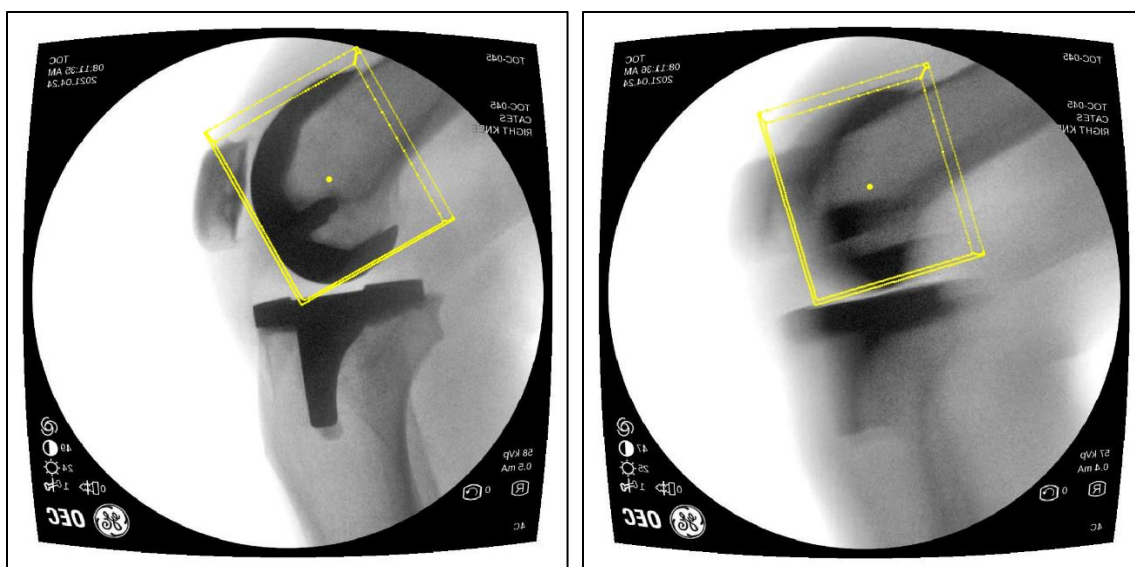


Figure 74. Registration results represented by the 3D box for a clear image (left) and blurry image (right).

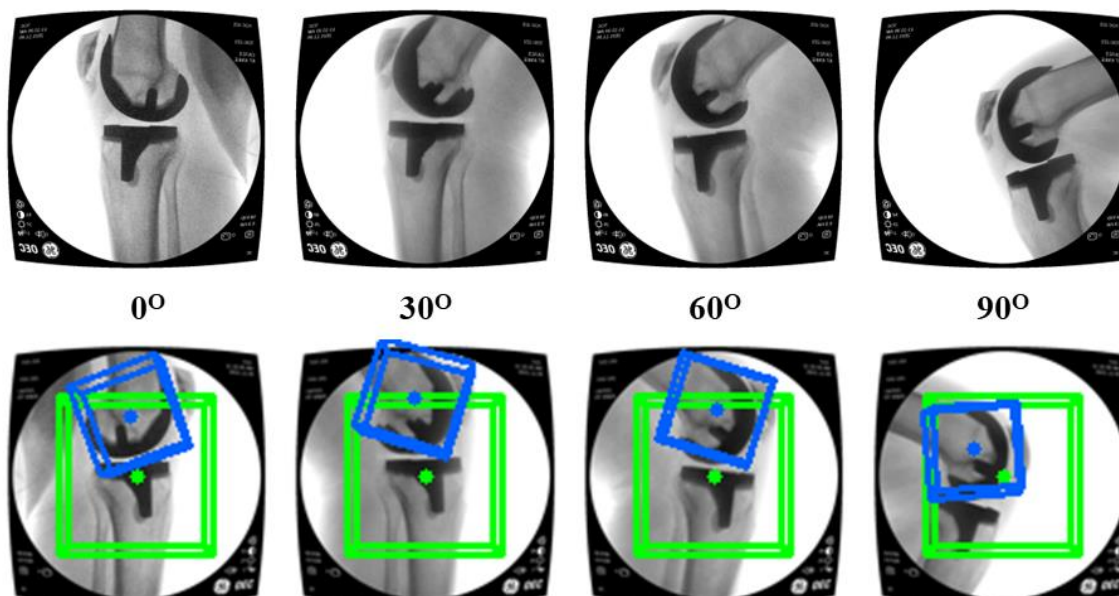
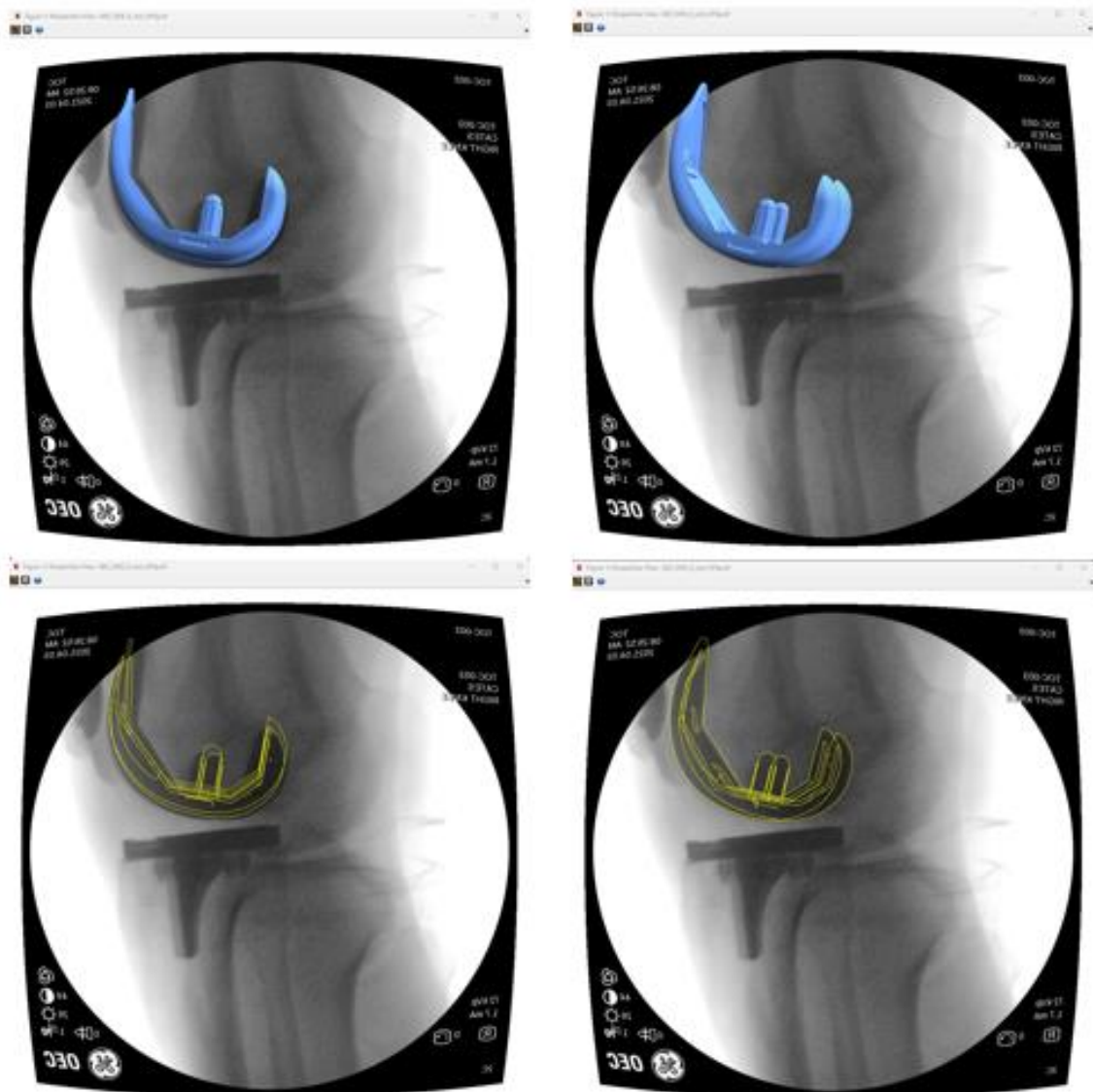


Figure 75. Registration results of a fluoroscopic sequence using machine learning.



Deep Learning

Optimization

Figure 76 Registration results with deep learning (left) and then applied optimization (right).

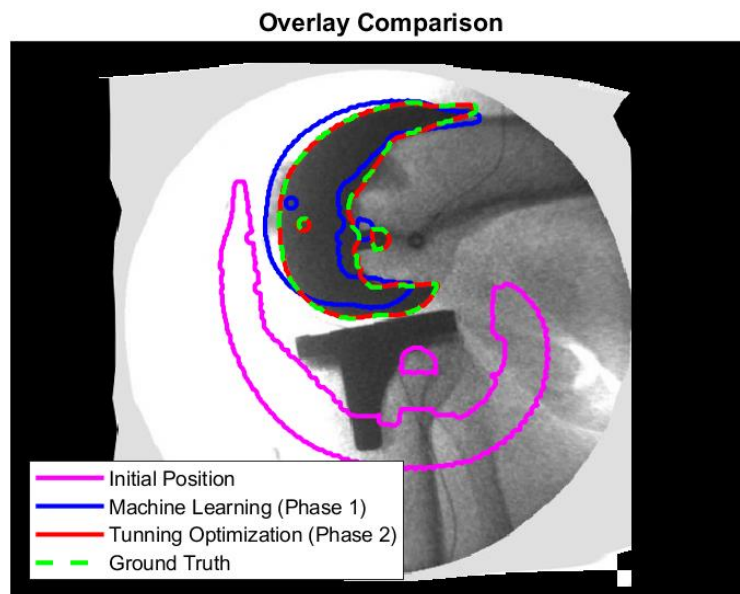
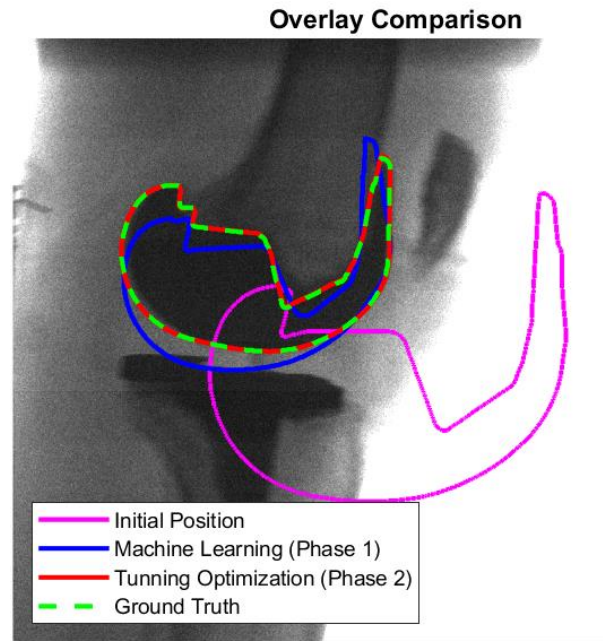


Figure 77. Registration results with initial position (pink), machine learning result (blue), optimization result (red), and the ground truth (green) of the PS implant case (top) and CR implant case (bottom).

enhancing the algorithm further, but nonetheless they show the potential of the proposed methodology.

Quantitatively, after the first phase, the average rotation error of the DNN's outputs is around 5.7 degrees, as seen in Table 4. Additionally, the average translation errors of the DNN's outputs in-plane and out-of-plane are 4 and 10 mm, respectively. After the tuning optimization step, the RMS errors of the translation and rotation between the predicted pose and the ground truth are lowered to 0.23 mm of in-plane difference, 0.98 mm of out-of-plane difference, and 0.90 degrees of rotation difference compared to the manual registration method as seen in Table 4. Among the three steps outlined in the proposed methods, Phase 2a, despite being slightly less accurate compared to Phase 2b, has the advantage of being completed within just 1 minute. Therefore, it is possible that Phase 2a alone could be sufficient. In practical situations, particularly when time constraints are a factor, Phase 2b may not be necessary.

7.3.2 Comparison with other Methods

In comparison with the results from a previous study as seen in Table 5, our proposed methodology is very competitive and fall within the same range of errors, with manual registration errors previously established as 0.65 mm of in-plane difference, 4.7 mm of out-of-plane difference, and 1.47 degrees of rotation difference [80]. Previously, Jensen et al. [89] developed a highly accurate machine learning algorithm that follows a similar framework to automatically produce segmented images from fluoroscopy images. While the methodologies may be similar, the previous research does not extend into the

Table 4. The average errors of the registration comparisons with the ground truths.

Unit (mm)	Method	Search range (mm)	Time (min)	In-plane xy error (mm)	Out- plane z error (mm)	Rotation error (°)
Phase 1: Deep learning	Automatic	$\pm 1,000$ mm	Real-time	5 mm	10.35 mm	5.71°
Phase 2a: Edge-based Optimization	Automatic	± 200 mm	1 min	0.33 mm	4.47 mm	1.56°
Phase 2b: Intensity- based Optimization	Automatic	± 50 mm	45 mins	0.23 mm	0.98 mm	0.90°

Table 5. The average errors of the benchmarks for the registration.

Unit (mm)	Method	Search range (mm)	Time (min)	In-plane xy error (mm)	Out- plane z error (mm)	Rotation error ($^{\circ}$)
Benchmark 1 (Banks, 1996)	Manual	± 250 mm	–	0.50 mm	6.6 mm	1.10°
Benchmark 2 (Mahfouz, 2003)	Manual	–	–	0.65 mm	3.21 mm	1.50°
Benchmark 3 (Fregly, 2005)	Automatic	± 50 mm	(long)	0.42 mm	5.60 mm	1.30°
Benchmark 4 (Tsai, 2010)	Automatic	± 100 mm	5 mins	0.72 mm	4.61 mm	1.21°
Benchmark 5 (Flood, 2018)	Automatic	± 55 mm	1 min	0.40 mm	1.40 mm	0.13°
Benchmark 6 (Postolka, 2020)	Automatic	± 305 mm	3 mins	5.89 mm	5.89 mm	1.35°
Benchmark 7 (Jensen, 2022)	Automatic	$\pm 1,000$ mm	-	0.68 mm	1.75 mm	1.61°

realm of manipulating CAD models in 3D space to generate precise condylar motion patterns. Our algorithm extends these applications further to provide the users with even more information, such as medial and lateral condylar rollback patterns, condylar liftoff patterns, etc., which can be used to provide valuable information regarding the behavior and functioning of the knee joint.

7.4 Discussion

In this study, we proposed a novel machine learning framework to register three-dimensional computer-aided design (CAD) implant models automatically and accurately onto two-dimensional X-ray or fluoroscopic images. This novel methodology can be used to overcome common 3D-to-2D registration limitations of manual overlaying and can be easily implemented with regular computer performance. The computational results show the competitive accuracy of the proposed methodology with the previous studies.

One of the major challenges of using deep learning in 3D-to-2D registration is that it typically requires a large amount of data to train the model effectively. The reason for this is that deep learning models have many parameters and require a significant amount of data to learn these parameters accurately. However, in many real-world applications, collecting a large dataset can be difficult or expensive. This can lead to overfitting, where the model memorizes the training data rather than learning to generalize to new data. To overcome this challenge, the proposed approach has incorporated additional techniques such as optimization and data augmentation, which can help to improve the performance of deep learning models even when only a small amount of data is available.

However, the above errors of the algorithm still occasionally need human revision in the end. The algorithm performs worse in some cases, for example when implant components overlap each other or are out of the fluoroscopic image. Another drawback of this methodology is the requirement of a massive training data, so future studies of other joints or component designs may be limited without extensive manual data for training. Regarding speed, the algorithm is developed on a Desktop with 16 Gb RAM and GPU GTX 1650, so this algorithm can be used on the regular computer performance. The training time for 200 epochs is around 2 days, but time for predicting the poses of 6,589 images based on the proposed methodology was recorded as 813 seconds, meaning the average speed of the algorithm is 8 images/second, which is interestingly the same framerate as traditional, commercial fluoroscopic units.

In terms of clinical applications, it is clear that this study could be very meaningful for post-operative evaluation of TKA kinematics. These algorithms could allow us to quickly and automatically analyze large data already contained within hospital or clinic databases, and it also provides us with benchmark directions for making 3D-to-2D registration evaluation methods available to clinicians around the world, without the need for additional manual labor. Furthermore, with the development of the AR/VR trend in orthopedics [100], these machine learning algorithms can also be implemented to yield more realistic and accurate poses of objects in the simulated environment, which can be beneficial for both orthopedics training purposes and doing surgeries with higher accuracy.

In the end, our aim is to show the potential ability to use machine learning for automatically determining in vivo TKA kinematics and how the proposed algorithm can be used to reduce workload and increase efficiency. Even though there will always be room for improvement within the algorithms, the speed and accuracy of the algorithms show potential uses in expanding access to these analysis tools as well as determining TKA kinematics in real time, perhaps in the operating room or in a clinical setting. Our algorithm also demonstrates great extended applications in biomedical engineering, as it can also be used on other joints and even non-implanted joints.

Chapter VIII

Implanted Knee Kinematics Analysis

Following the technical chapters, this section aims to shift the focus towards applying the proposed framework to analyze the kinematics of implanted knees.

8.1 Introduction

The implanted knee kinematics refer to the study of the movement and function of the knee joint after knee replacement surgery. Studying the movement and function of the knee joint after knee replacement surgery, known as implanted knee kinematics, is crucial for assessing the success of the surgery and identifying any potential issues [101, 27]. This information is used by orthopedic surgeons to ensure long-term success and function of knee replacement surgeries, as well as to improve future implant generations. Chapter V, Chapter VI, and Chapter VII have focused on important techniques such as 2D image segmentation, 3D reconstruction, and 3D-to-2D registration. Building on these techniques, this chapter aims to apply the proposed framework to analyze the kinematics of implanted knees.

The first step of the proposed framework involves reconstructing the 3D implant models (femur and tibia) using the morphing algorithm, if they are not already available. Next, the machine learning algorithms register these 3D models onto a series of fluoroscopic images. The registration transformations are then optimized using a combination of the 2D image registration machine learning technique and the tuning registration optimization process. The output of this step is the motion of both the femur

and tibia in 3D space, which is used to compute the relative movements of the femur with the tibia. This information provides us with kinematics information about various parameters of interest such as femorotibial (LAP, MAP, LSI, MSI), and axial rotation (AXR), as seen in Figure 78. Those parameters have been used successfully for evaluating the TKA kinematics outputs for around 30 years. The results of this analysis are detailed in the following sections.

8.2 Kinematics Model

After the above 3D-to-2D registration steps, the motions or transformation of both femur and tibia are obtained in global space. Since the polyethylene was attached to the tibia, the polyethylene's transformation is assigned to the same with the tibia's transformation. To output the kinematics, all models are transformed into the tibia's space so that the related motion between the femur and tibia is obtained. From the motion of the femur, the kinematics parameters included are computed to help the surgeons and engineers to evaluate the femoral component motion relative to the tibial component, in the tibial reference frame, as seen in Figure 79.

In order to evaluate the kinematics of a patient's movement, a fluoroscopic video capturing the movement is composed of multiple frames. To analyze the kinematics of the patient's movement, the analysis is performed on each individual frame of the video. Using the aforementioned methodology, the transformation of both femur and tibia is computed for each frame. Subsequently, the kinematics of the patient's frame are computed using the following steps.

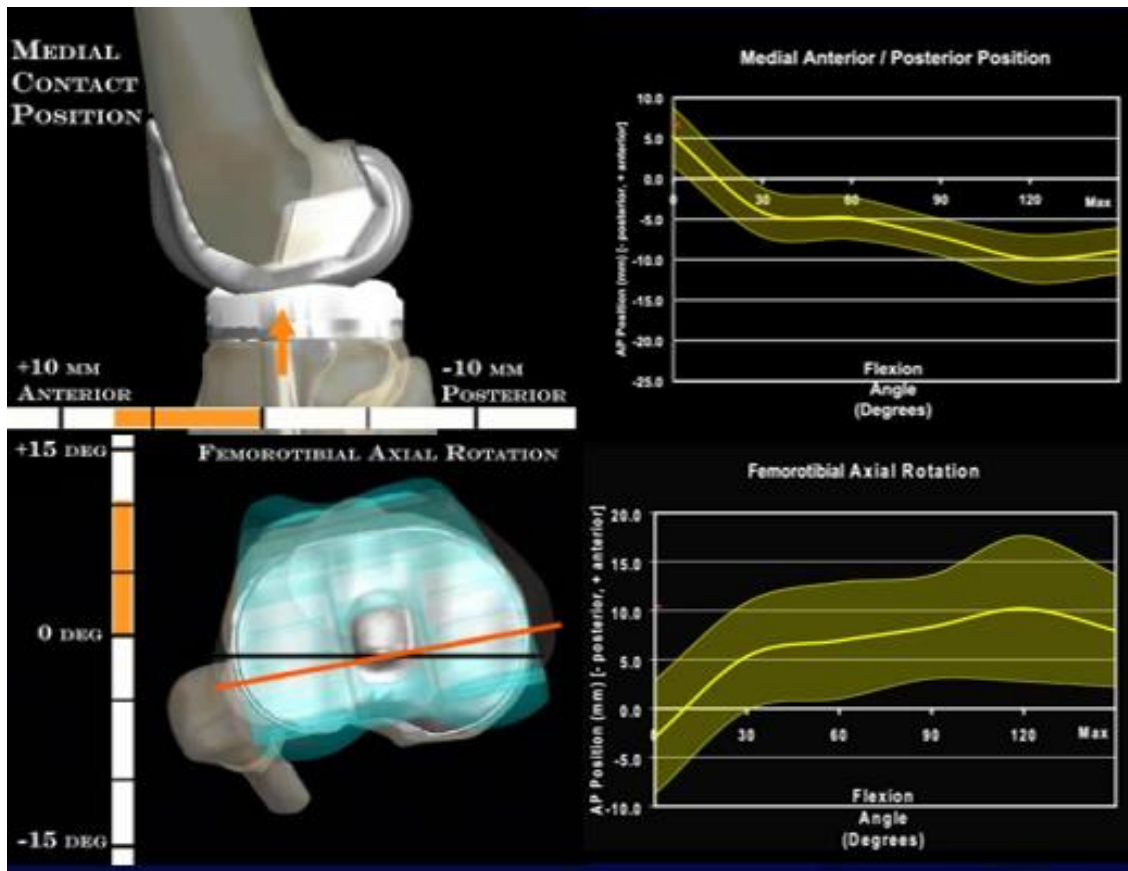


Figure 78. Parameters of interest: Femorotibial contact (LAP, MAP, LSI, MSI) and axial rotation (AXR).

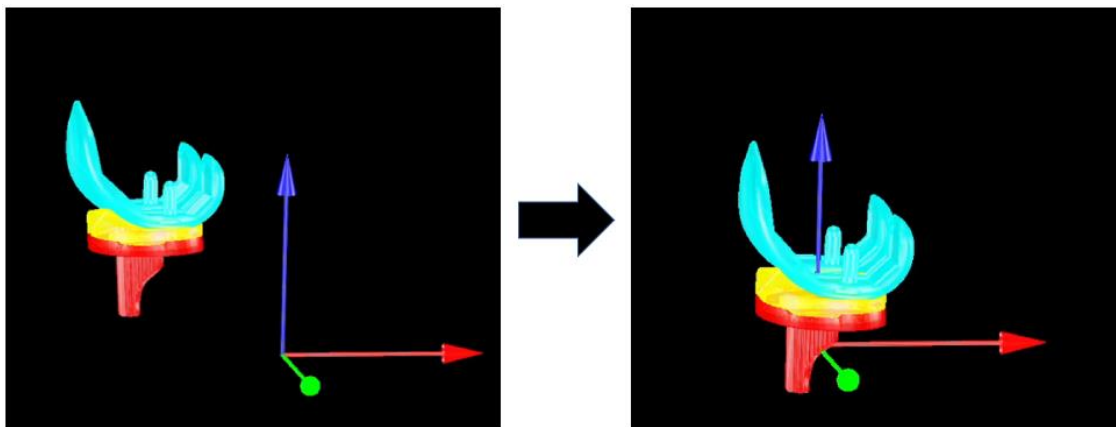


Figure 79. Converting the models from the global space to the tibia space.

1. Obtain the transformations of both the femur and tibia in 3D space using the proposed framework.
2. Calculate the relative motion of the femur with respect to the tibia.
3. Use the relative motion to compute the kinematic parameters of interest, including LAP, MAP, LSI, MSI, and axial rotation.
4. To calculate LAP and MAP, the relative motion of the femur with respect to the tibia in the anterior-posterior direction is measured. Positive values of LAP indicate lateral displacement of the femur, while negative values indicate medial displacement. Similarly, positive values of MAP indicate anterior displacement of the femur, while negative values indicate posterior displacement.
5. To calculate LSI and MSI, the relative motion of the femur with respect to the tibia in the medial-lateral direction is measured. Positive values of LSI indicate lateral displacement of the femur, while negative values indicate medial displacement. Similarly, positive values of MSI indicate medial displacement of the femur, while negative values indicate lateral displacement.
6. To calculate axial rotation, the rotation of the femur with respect to the tibia is measured. Positive values indicate external rotation of the femur, while negative values indicate internal rotation.

7. Finally, the computed values for each kinematic parameter can be used to assess the success of the knee replacement surgery and identify any issues that may require further treatment.

8.3 Results of Knee Kinematics Analysis

The ultimate goal of this dissertation is for a subject to perform a specific activity and the algorithm derives the kinematics for an implanted knee. The analysis focuses on specific parameters such as AXR (Axial Rotation), LAP (Lateral Anterior/Posterior Position), MAP (Medial Anterior/Posterior Position), LSI (Lateral Superior/Interior Position), and MSI (Medial Superior/Interior Position), which are critical in describing the position and movement of the implant components with respect to the patient's anatomy. The simulation results of the kinematics of one representative patient are shown in Figure 80.

Three different approaches were used in this study, and their results were presented in Figure 81. Each figure corresponded to a different set of parameters, including LAP, MAP, LSI, MSI, and AXR. The green line represents the ground truth, which is the known CAD model from the designing company and manual registration. The red line represents the known CAD model with the proposed automated 3D-2D registration method by machine learning. Finally, the blue line represents the fully automated approach with the 3D reconstructed model and automated 3D-2D registration. After comparing the results presented Figure 81, it was determined that both automated approaches performed well compared to the the known CAD & manual overlay ground truth. The maximum error for

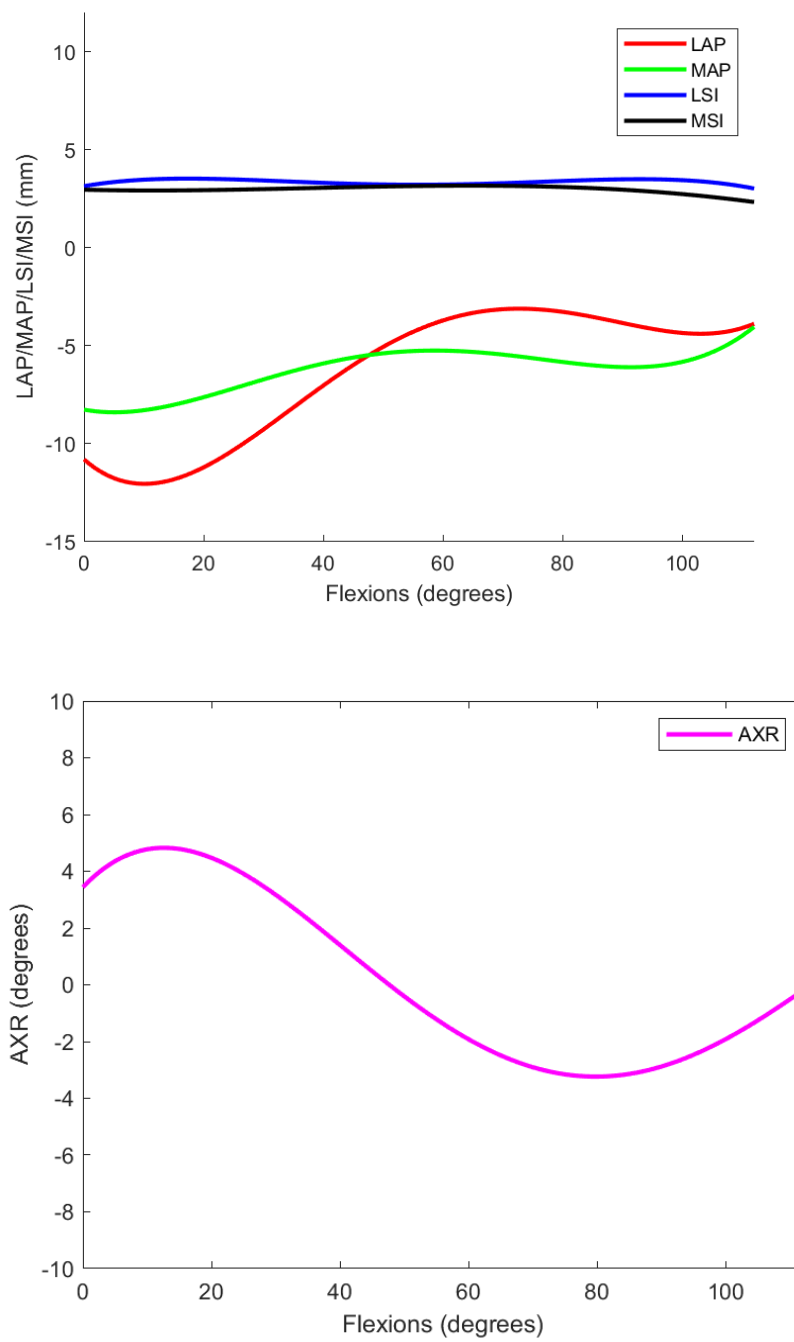


Figure 80. Femorotibial contact (top) and axial rotation (bottom).

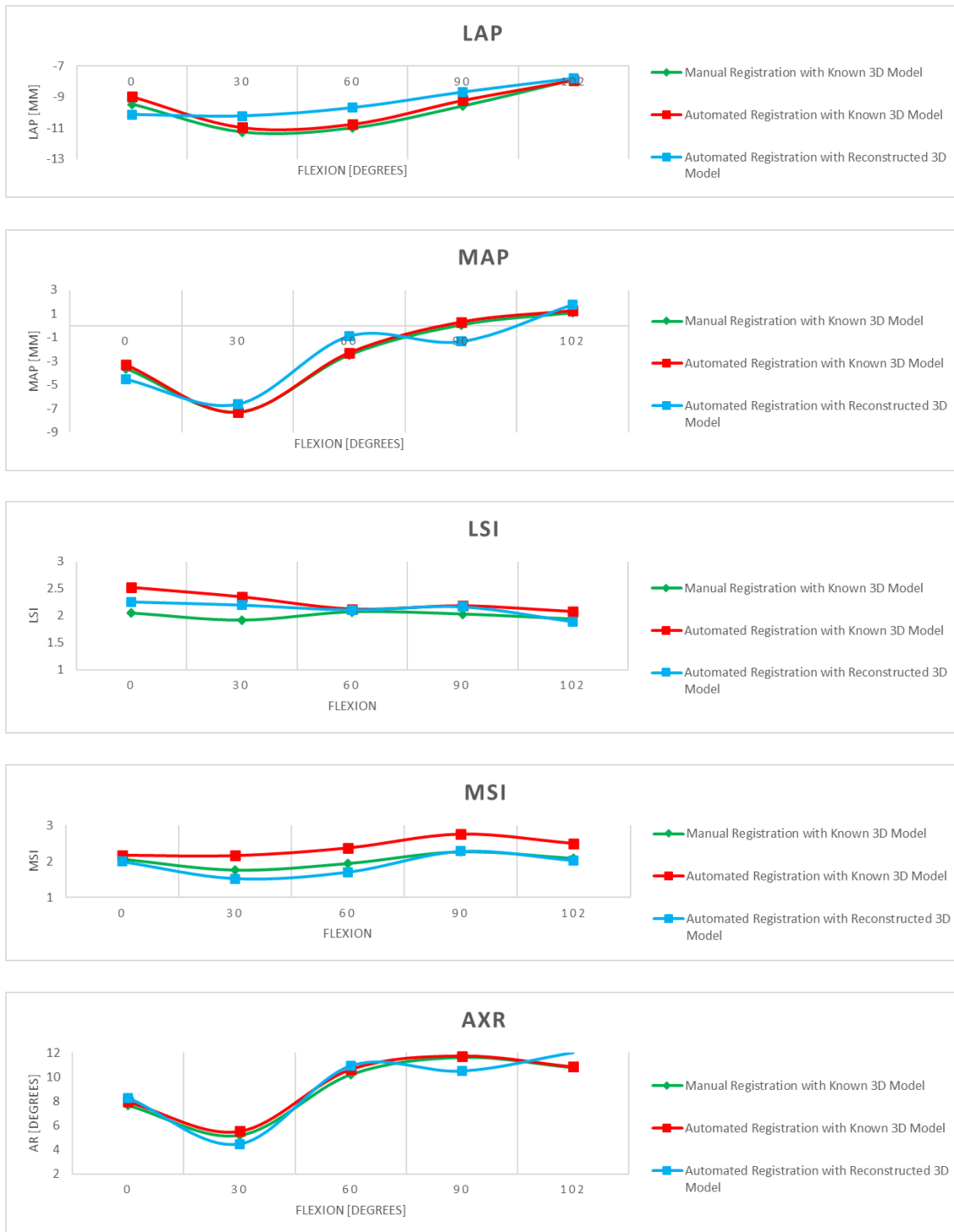


Figure 81. Kinematics outputs of different approaches.

all approaches were less than 1.5 mm for all translation parameters and 1.2 degrees for rotation measurements, with average errors of 0.53 mm for translation and 0.88 degrees for rotation. The known CAD & automated overlay approach exhibited the best performance between the two automated approaches. This is likely due to the errors from the morphed model reconstruction compounding with the errors from the automated overlay algorithms.

8.4 Discussion

The analysis of knee kinematics plays a vital role in evaluating knee replacement surgeries. It allows for the assessment of the movement and performance of the knee joint following the insertion of artificial components, ensuring the effectiveness and long-term functionality of the joint. The aforementioned kinematics modeling method proves to be robust and efficient in evaluating the motion of the femur and tibia within the implanted knee during flexion. This modeling approach offers a comprehensive and precise understanding of the joint's position and dynamics during dynamic activities, which is particularly valuable for evaluating knee implants designed for complex knee pathologies or active individuals. The analysis of knee kinematics not only aids in assessing the outcomes of knee replacement surgeries but also contributes to the development of enhanced surgical techniques and implant components. By examining knee kinematics, orthopedic surgeons can identify areas for enhancement and create new implant designs that more closely replicate the natural movement and functionality of the knee joint.

The simulation results showcased the high accuracy of the new model, particularly in relation to clinically significant outcomes like knee kinematics. Consequently, these

algorithms have the potential to facilitate swift and automated analysis of extensive data stored within hospital or clinic databases. They also serve as a valuable reference for implementing 3D-to-2D registration evaluation methods globally, benefiting clinicians and research groups without requiring additional manual effort. Moreover, these evaluations can be conducted rapidly, examining all images within a fluoroscopic video sequence, even without 3D CAD models. This breakthrough opens the door to determining in vivo implant kinematics for any patient, regardless of location or time.

In summary, the primary objective of this study was to showcase the potential of machine learning in automatically determining in vivo total knee arthroplasty (TKA) kinematics and how the proposed algorithm could enhance efficiency and reduce workload. While there is always room for algorithm improvement, the speed and accuracy demonstrated in this study highlight the potential for widespread utilization of these analysis tools, enabling real-time determination of TKA kinematics, such as in the operating room or clinical settings. Furthermore, this novel algorithm exhibits versatility in its applicability to other biomedical engineering domains, including analysis of joints beyond knee implants and even non-implanted joints. Unlike previous fluoroscopic studies that relied on sponsorships for funding and access to CAD models, this algorithm enables analysis of kinematics for any patient, anywhere, at any time, and is capable of handling larger cohorts by analyzing complete fluoroscopic sequences.

Chapter IX

Conclusions and Future Work

9.1 Conclusions

The aim of this research was to present a fresh technique for examining in vivo 3D kinematics using an automated method and to perform evaluations to verify its precision and efficiency in contrast to the manual methods utilized in the past. While the manual studies are expensive and are very time-consuming, often requiring months to complete, the proposed machine learning algorithm may prove to overcome previous hurdles by increasing the speed, reduce human error, and improve upon technical barriers to successfully derive kinematics even when 3D CAD were not available.

Historically, there have been many attempts at automating various aspects of medical imaging and biomechanics analyses. Many “non-learning” algorithms were used to try to solve automated 3D-to-2D registration problems, such as edge-based optimization, intensity-based optimization, feature-based optimization, etc. Intensity-based registration alone has been used for registering both spine models [82] and femur/tibia models [17]. Also, a brain model is registered by using the intensity-based algorithm, but this approach requires a CT scan of the human brain [83]. 3D-to-2D registration of non-implanted femurs was achieved using edge-based registration [84], and 3D-to-2D registration of TKA femur components was previously completed using edge-based registration via Lipschitzian optimization techniques [85]. However, these previous methods typically displayed moderate accuracy and unstable results, and to the author’s knowledge, no group has

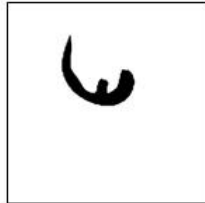
developed a robust algorithm nor extended these applications further to yield more direct clinical comparisons to practical kinematic fluoroscopic analyses.

There are many examples of using deep learning (DL), a popular branch of machine learning (ML), for registering the 3D models within a specified space, with very high performance [86, 26, 87]. A non-implanted hip model was registered onto biplanar X-ray images using DL, but this approach requires the CT scan of the hip bone model [88]. The segmented images of the implanted femur and tibia are processed from the fluoroscopic image by using two convolutional neural networks, and then they are used for estimating the initial pose by Fourier descriptor-based estimation [89]. Previously, a similar framework [89] developed a highly accurate machine learning algorithm produce automatically segmented images from fluoroscopy images. While these methodologies may be similar, the previous research does not extend into the realm of manipulating CAD models in 3D space to generate precise condylar motion patterns. The proposed algorithm in this present study extends upon these applications further by providing users with even more clinically-relevant information.

The entire flow of the framework as seen in Figure 82, from automated image segmentation to the generation of clinically applicable results, can be used to overcome common 3D-to-2D registration limitations of manual overlaying and can be easily implemented by non-technical users. While our computational results reveal highly competitive accuracy, one of the major challenges of using deep learning in 3D-to-2D registration is that it typically requires a large amount of data to train the model effectively.



Step 1: Data Collection



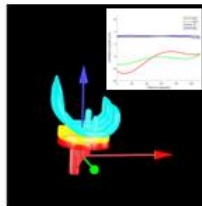
Step 2: Image Segmentation



Step 3: 3D Reconstruction



Step 4: 3D-to-2D Registration



Step 5: Kinematics analysis

Figure 82. The summarization of the entire framework of 5 main steps of the dissertation: (i) data collection, (ii) image segmentation, (iii) 3D reconstruction, (iv) 3D-to-2D registration, and (v) kinematics modelling.

Overall, this is simply a difficult problem to solve. To compound this challenge even further, specific training sets may not always be optimal due to the diversity in implant models and sizes. For example, a model can be trained on specifically femoral component, but then it may have difficulty distinguishing between PS and CR implants. Conversely, an algorithm can be trained on a multitude of different models, perhaps a different one for each different implant make/model, of all different sizes, from each different company, but that would be unfeasible and would yield small training datasets. This can also lead to overfitting, where the model memorizes the training data rather than learning to generalize to new data. To overcome this challenge, it has been demonstrated with this new approach that an optimized number of training models based on the sides (left or right) and the major implant types (PS or CR) can be utilized, while incorporating additional techniques such as optimization and data augmentation to help further improve the performance of our models.

The results from this study demonstrated that this new model is highly accurate with respect to clinically relevant results such as knee kinematics. Thus, the algorithms could allow a user to quickly and automatically analyze large data already contained within hospital or clinic databases, and it also provides us with benchmark directions for making 3D-to-2D registration evaluation methods available to clinicians and research groups around the world, without the need for additional manual labor. In addition, these evaluations can be done extremely quickly, analyzing all images within a fluoroscopic video sequence, even if 3D CAD models were not available. Essentially, it can now become

possible to determine in vivo implant kinematics for any patient, anywhere, at any time. While the intensity-based method proposed herein is time-consuming, the ML and edge-based methods finish within seconds and still maintains competitive accuracy, potentially allowing for real-time feedback and data that can be shown to patients in the clinic immediately after the procedure. Therefore, making this model and subsequent resources available to clinicians, engineers, and researchers alike, it is believed that these methods can help the orthopaedic community immediately evaluate kinematic patterns and potentially identify previously unknown complications (such as instability or ligament imbalancing) much more quickly than previously possible. This approach can also be readily adapted for use in the analysis of kinematics in other joints.

As in all studies, there are still several limitations in the proposed framework. While deep learning has demonstrated promising results in image segmentation, there are still several limitations to its application. One major challenge is the requirement for large amounts of labeled data to train the model, which can be both time-consuming and expensive, as the currently trained model may not be as accurate with other datasets. Another limitation is that various fluoroscopy units have different properties and camera calibration requirements. Additionally, it should be noted that image quality is a significant factor that can impact the accuracy of results. When training algorithms use controlled, high-quality static data for training, the accuracy may be compromised on older fluoroscopy units during dynamic activities, which may possess slower framerates and exhibit more pronounced image ghosting. Conversely, training models on data obtained

from older units that exhibit fish-eye warping effects may pose challenges for higher quality images obtained from flat panel units or those that present variations in the apparent size of knees on the screen.

In summary, one of the main aims was to demonstrate the potential ability for using machine learning to automatically determine in vivo TKA kinematics and how the proposed algorithm can be used to reduce workload and increase efficiency. Even though there will always be room for improvement within the algorithms, the speed and accuracy of the algorithms in this study revealed potential uses in expanding access to these analysis tools as well as determining TKA kinematics in real time, perhaps in the operating room or in a clinical setting. This new algorithm also demonstrated an extension to other applications in biomedical engineering, as it can also be used on other joints and even non-implanted joints. Although previous fluoroscopic studies were limited to securing sponsorships for funding and subsequently receiving CAD from the sponsors, this new algorithm allows for analysis of any patient, anywhere, at any time, and can analyze the full fluoroscopic sequence for much larger numbers of subjects.

9.2 Assumptions

There are several assumptions of the current work:

1. The manual registrations are considered as the ground truth for the training data for both the image segmentation and the image registration deep neural network models.

2. The CAD models from the company are considered as the ground truth for the 3D reconstruction problem in this dissertation.

9.3 Limitations and Future Work

While the dissertation demonstrates impressive results, it is important to acknowledge the presence of certain limitations within the current work. However, to address these limitations, future efforts will be concentrated on implementing necessary improvements.

1. Computational power is limited with the regular desktop and laptop performance (2 laptops: Core i7-10750H, CPU@2.6GHz, GeForce GTX 1650 Ti; 1 desktop: Core i9-12900KF, CPU@3.2Ghz, GeForce RTX 3080). The results can be improved by using a higher computer configuration with more running time.
2. Machine learning algorithms can become too complex and fit the training data too well, resulting in poor generalization of new data.
3. Deep neural networks are black boxes, making it challenging to interpret how they arrive at their predictions. This lack of interpretability can be a significant limitation in applications where human understanding is necessary.
4. Deep learning algorithms require domain expertise to design, train, and evaluate them effectively. Thus, the current model parameters for deep learning might not be optimized and need to test more to find the best model.
5. Deep learning algorithms can inherit biases from the data they are trained on, leading to unfair or discriminatory predictions. Thus, the deep learning model might

need to be updated for the new implant knee models that were not available in the current dataset.

6. Despite the notable achievements of machine learning in registration and segmentation, there are still instances where low-quality fluoroscopic images pose challenges and lead to suboptimal performance.
7. The algorithm's performance may not be optimal when applied to new implant types that significantly differ from the currently existing femur and tibia.
8. The training and running time of the algorithm may be hampered by limitations in computer power, resulting in slower performance.
9. The current number of testing cases is limited, and further testing is necessary to obtain a comprehensive and conclusive result.

List of References

- [1] Varacallo M, Luo TD, Johanson NA, "Total Knee Arthroplasty Techniques," *StatPearls Publishing*, 2022.
- [2] Xingbang Yang, Ruoqi Zhao, Dana Solav, Xuan Yang, Duncan R.C. Lee, Bjorn Sparrman, Yubo Fan, Hugh Herr, "Material, design, and fabrication of custom prosthetic liners for lower-extremity amputees: A review," *Medicine in Novel Technology and Devices*, vol. 17, no. 100197, pp. ISSN 2590-0935, 2023.
- [3] Angerame, M. R., Holst, D. C., Jennings, J. M., Komistek, R. D., & Dennis, D. A., "Total Knee Arthroplasty Kinematics," *The Journal of arthroplasty*, vol. 34, no. 10, p. 2502–2510, 2019.
- [4] Dennis DA, Komistek RD, Hoff WA, Gabriel SM, "In vivo knee kinematics derived using an inverse perspective technique," *Clin Orthop Relat Res.*, vol. 331, pp. 107-17, 1996 Oct.
- [5] Mahfouz et al., "A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images," *IEEE transactions on medical imaging*, vol. 22, no. 12, p. 1561–1574, 2003.

- [6] Banks et al., "Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy," *IEEE transactions on bio-medical engineering*, vol. 43, no. 6, p. 638–649, 1996.
- [7] Dennis et al., "In vivo knee kinematics derived using an inverse perspective technique," *Clinical orthopaedics and related research*, vol. 331, p. 107–117, 1996.
- [8] Mohamed Abd Elaziz, Songfeng Lu, "Many-objectives multilevel thresholding image segmentation using Knee Evolutionary Algorithm," *Expert Systems with Applications*, vol. 125, no. 0957-4174, pp. 305-316, 2019.
- [9] Yang Feng, Li Sze Chow, Nadia Muhammad Gowdh, Norlisah Ramli, Li Kuo Tan, Suhailah Abdullah, Sew Sun Tiang, "Gradient-based edge detection with skeletonization (GES) segmentation for magnetic resonance optic nerve images," *Biomedical Signal Processing and Control*, vol. 80, no. 1746-8094, p. 104342, 2023.
- [10] Tongyi Wei, Xiaopeng Wang, Jiaxin Wu, Shengyang Zhu, "Interval type-2 possibilistic fuzzy clustering noisy image segmentation algorithm with adaptive spatial constraints and local feature weighting & clustering weighting," *International Journal of Approximate Reasoning*, vol. 157, no. 0888-613X, pp. 1-32, 2023.

- [11] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, Demetri Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *arXiv:2001.05566*, 2020.
- [12] Daniel Bolya, Chong Zhou, Fanyi Xiao, Yong Jae Lee, "YOLACT: Real-time Instance Segmentation," *<https://arxiv.org/abs/1904.02689>*, 2019.
- [13] Jun Yu, Wenbin Yin, Zhiyi Hu, Yabin Liu, "3D Reconstruction for Multi-view Objects," *Computers and Electrical Engineering*, vol. 106, no. 0045-7906, p. 108567, 2023.
- [14] Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 2022.
- [15] Yoni Kasten, Daniel Doktofsky & Ilya Kovler, "End-To-EndConvolutional Neural Network for 3D Reconstruction of Knee Bones from Bi-planar X-Ray Images," *International Workshop on Machine Learning for Medical Image Reconstruction*, vol. 12450, no. Machine Learning for Medical Image Reconstruction , 2020.
- [16] Myronenko et al., "Point Set Registration: Coherent Point Drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262-2275, 2010.
- [17] Barbara Postolka, Renate List, Benedikt Thelen, Pascal Schütz, William R. Taylor, Guoyan Zheng, "Evaluation of an intensity-based

- algorithm for 2D/3D registration of natural knee videofluoroscopy data," *Medical Engineering & Physics*, vol. 77, pp. 107-113, 2020.
- [18] J. Wu, 2D-3D Registration of Knee Joint from Single Plane X-ray Fluoroscopy Using Nonlinear Shape Priors, Knoxville: PhD diss., University of Tennessee, 2016.
- [19] Cong Wang, Shuaining Xie, Kang Li, Chongyang Wang, Xudong Liu, Liang Zhao, Tsung-Yuan Tsai, "Multi-View Point-Based Registration for Native Knee Kinematics Measurement with Feature Transfer Learning," *Engineering*, vol. 7, no. 6, pp. 881-888, 2021.
- [20] Srikrishna Jaganathan, Jian Wang, Anja Borsdorf, Karthik Shetty, Andreas Maier, "Deep Iterative 2D/3D Registration," *MICCAI* , 2021.
- [21] E. Gronenschild, "The accuracy and reproducibility of a global method to correct for geometric image distortion in the x-ray imaging chain," *Med Phys*, vol. 24, no. 12, pp. 1875-88, 1997.
- [22] Mahfouz et al., "Effect of segmentation errors on 3D-to-2D registration of implant models in X-ray images," *Journal of Biomechanics*, vol. 38, no. 2, pp. 229-239, 2005.
- [23] Yasuhiro Aoki, Hunter Goforth, Rangaprasad Arun Srivatsan, Simon Lucey, "PointNetLK: Robust & Efficient Point Cloud Registration using PointNet," *arXiv:1903.05711*, 2019.

- [24] Sergey Zakharov, Ivan Shugurov, Slobodan Ilic, "DPOD: 6D Pose Object Detector and Refiner," <https://arxiv.org/abs/1902.11020>, 2019.
- [25] Yao Feng and Fan Wu and Xiaohu Shao and Yanfeng Wang and Xi Zhou, "Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network," *ECCV*, 2018.
- [26] Zelin Zhao, Gao Peng, Haoyu Wang, Hao-Shu Fang, Chengkun Li, Cewu Lu, "Estimating 6D Pose From Localizing Designated Surface Keypoints," *arXiv:1812.01387*, 2018.
- [27] Dennis, Douglas A. MD***†; Komistek, Richard D. PhD***†; Mahfouz, Mohamed R. PhD*; Haas, Brian D. MD*; Stiehl, James B. MD‡, "Conventry Award Paper: Multicenter Determination of In Vivo Kinematics After Total Knee Arthroplasty," *Clinical Orthopaedics and Related Research*, vol. 416, pp. 37-57, 2003.
- [28] Miner AL, Lingard EA, Wright EA, Sledge CB, Katz JN, "Knee range of motion after total knee arthroplasty: how important is this as an outcome measure?," *J Arthroplasty*, vol. 18, no. 3, pp. 286-94, 2003.
- [29] Khasian et al., "Effects of Posterior Tibial Slope on a Posterior Cruciate Retaining Total Knee Arthroplasty Kinematics and Kinetics," *The Journal of Arthroplasty*, vol. 36, no. 7, pp. 2379-2385, 2021.

- [30] Khasian et al., "In Vivo Knee Kinematics for a Cruciate Sacrificing Total Knee Arthroplasty Having Both a Symmetrical Femoral and Tibial Component," *The Journal of arthroplasty*, vol. 35, no. 6, p. 1712–1719, 2020.
- [31] Smith et al., "In Vivo Knee Kinematics: How Important Are the Roles of Femoral Geometry and the Cruciate Ligaments?," *The Journal of arthroplasty*, vol. 36, no. 4, p. 1445–1454, 2021.
- [32] M. Cheriet, J. N. Said and C. Y. Suen, "A recursive thresholding technique for image segmentation," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 918-921, 1998.
- [33] H. G. Kaganami and Z. Bei, "Region-Based Segmentation versus Edge Detection," *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 1217-1221, 2009.
- [34] Li Guo, Pengfei Shi, Long Chen, Chenglizhao Chen, Weiping Ding, "Pixel and region level information fusion in membership regularized fuzzy clustering for image segmentation," *Information Fusion*, vol. 92, pp. 479-497, 2023.

- [35] Jonathan Long, Evan Shelhamer, Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation," <https://arxiv.org/abs/1411.4038>, 2015.
- [36] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, Philip H. S. Torr, "Conditional Random Fields as Recurrent Neural Networks," <https://arxiv.org/abs/1502.03240>, 2015.
- [37] Ziwei Liu, Xiaoxiao Li, Ping Luo, Chen Change Loy, Xiaoou Tang, "Semantic Image Segmentation via Deep Parsing Network," <https://arxiv.org/abs/1509.02634>, 2015.
- [38] Ma F, Gao F, Sun J, Zhou H, Hussain A, "Attention Graph Convolution Network for Image Segmentation in Big SAR Imagery Data," *Remote Sensing*, vol. 11, no. 21, p. 2586, 2019.
- [39] Deep Residual Learning for Image Recognition, "Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun," <https://arxiv.org/abs/1512.03385>, 2015.
- [40] Microsoft COCO: Common Objects in Context, "Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár," <https://arxiv.org/abs/1405.0312>, 2015.

- [41] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár, "Microsoft COCO: Common Objects in Context," <https://arxiv.org/abs/1405.0312>, 2014.
- [42] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition," <https://arxiv.org/abs/1512.03385>, 2015.
- [43] Paul Henderson, Vittorio Ferrari , "End-to-End Training of Object Class Detectors for Mean Average Precision," *Computer Vision – ACCV*, vol. 10115, p. 198–213, 2016.
- [44] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays,, "Microsoft coco: Common objects in context," *ECCV*, 2014.
- [45] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick, "Mask r-cnn," *ICCV*, 2017.
- [46] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia, "Path aggregation network for instance segmentation," *CVPR*, 2018.
- [47] Zhaojin Huang, Lichao Huang, Yongchao Gong, Chang Huang, and Xinggang Wang, "Mask scoring r-cnn," *CVPR*, 2019.

- [48] Vojtech Kulvait, Georg Rose, "Cutting Voxel Projector a New Approach to Construct 3D Cone Beam CT Operator," <https://arxiv.org/abs/2110.09841>, 2021.
- [49] Brunie et al., "Pre- and intra-irradiation multimodal image registration: principles and first experiments," *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 29, no. 2, p. 244–252, 1993.
- [50] Alperin et al., "Retrospective registration of X-ray angiograms with MR images by using vessels as intrinsic landmarks," *Journal of magnetic resonance imaging : JMRI*, vol. 4, no. 2, p. 139–144, 1994.
- [51] Dessinger et al., "Can an OA Knee Brace Effectively Offload the Medial Condyle? An In Vivo Fluoroscopic Study," *The Journal of arthroplasty*, vol. 36, no. 4, p. 1455–1461, 2021.
- [52] Sekiguchi et al., "Varus alignment after total knee arthroplasty results in greater axial rotation during deep knee bend activity," *Clinical biomechanics*, vol. 77, p. 105051, 2020.
- [53] LaCour et al., "Development of a hip joint mathematical model to assess implanted and non-implanted hips under various conditions," *Journal of Biomechanics*, vol. 112, p. 110051, 2020.

- [54] Cyteval, "Imaging of knee implants and related complications," *Diagnostic and Interventional Imaging*, vol. 97, no. 7, pp. 809-821, 2016.
- [55] Furey et al., "Prevention of radiation injury from medical imaging," *Journal of Vascular Surgery*, vol. 53, no. 1, pp. 22S-27S, 2011.
- [56] Zheng et al., "Reconstruction of Patient-Specific 3D Bone Surface from 2D Calibrated Fluoroscopic Images and Point Distribution Model," *Medical image computing and computer-assisted intervention*, vol. 4190, pp. 25-32, 2006.
- [57] Wu et al., "Reconstruction of knee anatomy from single-plane fluoroscopic x-ray based on a nonlinear statistical shape model," *Journal of medical imaging*, vol. 8, no. 1, p. 016001, 2021.
- [58] Sadick et al., "Impact of biplane versus single-plane imaging on radiation dose, contrast load and procedural time in coronary angioplasty," *The British journal of radiology*, vol. 83, no. 989, p. 379–394, 2010.
- [59] Bennani et al., "Three-dimensional reconstruction of In Vivo human lumbar spine from biplanar radiographs," *Computerized Medical Imaging and Graphics*, vol. 96, p. 102011, 2022.
- [60] Kasten et al., "End-To-End Convolutional Neural Network for 3D Reconstruction of Knee Bones from Bi-planar X-Ray Images," *Machine Learning for Medical Image Reconstruction*, Springer, vol. 12450, 2020.

- [61] Peng et al., "Robust CPD Algorithm for Non-Rigid Point Set Registration Based on Structure Information," *PLoS ONE*, vol. 11, no. 2, p. e0148483, 2016.
- [62] Anil et al., "An improved edge detection algorithm for X-Ray images based on the statistical range," *Heliyon*, vol. 5, no. 10, p. e02743, 2019.
- [63] Desbrun et al., "Implicit Fairing of Irregular Meshes Using Diffusion and Curvature Flow," *ACM Press/Addison-Wesley Publishing Co.*, p. 317–324, 1999.
- [64] Scarvell et al., "New registration algorithm for determining 3D knee kinematics using CT and single-plane fluoroscopy with improved out-of-plane translation accuracy," *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, vol. 28, no. 3, p. 334–340, 2010.
- [65] Komistek et al., "In vivo fluoroscopic analysis of the normal human knee," *Clinical orthopaedics and related research*, vol. 410, p. 69–81, 2003.
- [66] Victor et al., "In vivo kinematics after a cruciate-substituting TKA," *Clinical orthopaedics and related research*, vol. 468, no. 3, p. 807–814, 2010.
- [67] Bontempi et al., "J-curve design total knee arthroplasty: the posterior stabilized shows wider medial pivot compared to the cruciate retaining

- during chair raising," *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA*, vol. 28, no. 9, p. 2883–2892, 2020.
- [68] Mahfouz et al., "A novel approach to 3D bone creation in minutes : 3D ultrasound," *The bone & joint journal*, Vols. 103-B, no. 6 Supple A, p. 81–86, 2021.
- [69] Kim et al., "Accuracy and reliability of length measurements on three-dimensional computed tomography using open-source OsiriX software," *Journal of digital imaging*, vol. 25, no. 4, p. 486–491, 2012.
- [70] Sanjit et al., "Computed Tomography Scan to Detect Intra-articular Air in the Knee Joint: A Cadaver Study to Define a Low Radiation Dose Imaging Protocol," *Journal of Orthopaedic Trauma*, vol. 27, p. 505–508, 2013.
- [71] Graulich et al., "Fluoroscopic Intraoperative Images Produce Higher Image Quality and Decrease Total Radiation Exposure Compared to Radiographic X-ray Images in Patients After Primary Total Knee Arthroplasty," *In vivo (Athens, Greece)*, vol. 34, no. 5, p. 2521–2526, 2020.
- [72] Moret et al., "Customised, individually made total knee arthroplasty shows promising 1-year clinical and patient reported outcomes," *Archives of orthopaedic and trauma surgery*, vol. 141, no. 12, p. 2217–2225, 2021.

- [73] Khasian et al., "A Validated Forward Solution Dynamics Mathematical Model of the Knee Joint: Can It Be an Effective Alternative for Implant Evaluation?," *The Journal of arthroplasty*, vol. 35, no. 11, p. 3289–3299, 2020.
- [74] H. Ning, R. Yin, A. Ullah and F. Shi, "A Survey on Hybrid Human-Artificial Intelligence for Autonomous Driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6011-6026, 2022.
- [75] Z. Wang, Y. Guo and W. Zuo, "Deepfake Forensics via an Adversarial Game," *IEEE Transactions on Image Processing*, vol. 31, pp. 3541-3552, 2022.
- [76] Birkhoff DC, van Dalen ASHM, Schijven MP, "A Review on the Current Applications of Artificial Intelligence in the Operating Room," *Surg Innov.*, vol. 28, no. 5, pp. 611-619, 2021 Oct.
- [77] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [78] Banks SA, Hodge WA, "Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy," *IEEE Trans Biomed Eng.*, vol. 43, no. 6, pp. 638-49, 1996 Jun.
- [79] S. Zuffi, A. Leardini, F. Catani, S. Fantozzi, and A. Cappello, "A model-based method for the reconstruction of total knee replacement

- kinematics," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 981-991, 1999.
- [80] M.R. Mahfouz, W.A. Hoff, R.D. Komistek, and D.A. Dennis, "A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images," *IEEE Transactions on Medical Imaging*, vol. 22, no. 12, pp. 1561-1574, 2003.
- [81] T. Yamazaki, T. Watanabe, Y. Nakajima, K. Sugamoto, T. Tomita, H. Yoshikawa, and S. Tamura, "Improvement of depth position in 2-d/3-d registration of knee implants using single-plane fluoroscopy," *IEEE Transactions on Medical Imaging*, vol. 23, no. 5, pp. 602-612, 2004.
- [82] Shao Z, Han J, Liang W, Tan J, Guan Y, "Robust and Fast Initialization for Intensity-Based 2D/3D Registration," *Advances in Mechanical Engineering*, vol. 6, 2014.
- [83] Liu S, Yang B, Wang Y, Tian J, Yin L, Zheng W, "Multimode Medical Image Registration Based on Normalized Cross-Correlation," *Applied Sciences*, vol. 12, no. 6, p. 2828, 2022.
- [84] Scarvell JM, Pickering MR, Smith PN, "New registration algorithm for determining 3D knee kinematics using CT and single-plane fluoroscopy with improved out-of-plane translation accuracy," *J Orthop Res.*, vol. 28, no. 3, pp. 334-40, 2010 .

- [85] P. D. L. Flood and S. A. Banks, "Automated Registration of 3-D Knee Implant Models to Fluoroscopic Images Using Lipschitzian Optimization," *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 326-335, 2018.
- [86] Yu Xiang, Tanner Schmidt, Venkatraman Narayanan, Dieter Fox, "PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes," *Robotics: Science and Systems*, 2018.
- [87] Yannick Bukschat, Marcus Vetter, "EfficientPose: An efficient, accurate and scalable end-to-end 6D multi object pose estimation approach," *arXiv:2011.04307*, 2020.
- [88] Van Houtte, J., Audenaert, E., Zheng, G. et al., "Deep learning-based 2D/3D registration of an atlas to biplanar X-ray images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, p. 1333–1342, 2022.
- [89] Andrew Jensen, Paris Flood, Lindsey Palm-Vlasak, Will Burton, Paul Rullkoetter, Scott Banks, "Joint Track Machine Learning: An autonomous method for measuring 6DOF TKA kinematics from single-plane x-ray images," <https://arxiv.org/abs/2205.00057>, 2022.
- [90] Colin Rennie, Rahul Shome, Kostas E. Bekris, Alberto F. De Souza, "A Dataset for Improved RGBD-based Object Detection and Pose

Estimation for Warehouse Pick-and-Place,"
<https://arxiv.org/abs/1509.01277>, 2016.

- [91] David Fleet, Tomas Pajdla, Bernt Schiele, Tinne Tuytelaars, "Computer Vision -- ECCV 2014," *13th European Conference*, vol. Part IV, 2014.
- [92] Tomas Hodan, Pavel Haluza, Stepan Obdrzalek, Jiri Matas, Manolis Lourakis, Xenophon Zabulis, "T-LESS: An RGB-D Dataset for 6D Pose Estimation of Texture-less Objects," <https://arxiv.org/abs/1701.05498>, 2017.
- [93] Hartley, R., & Zisserman, A. , Multiple View Geometry in Computer Vision (2nd ed.), Cambridge: Cambridge University Press, 2004.
- [94] Barone F, Marrazzo M, Oton CJ, "Camera Calibration with Weighted Direct Linear Transformation and Anisotropic Uncertainties of Image Control Points," *Sensors (Basel)*, vol. 20, 2020 .
- [95] Dennis, Douglas A. MD*,**; Komistek, Richard D. PhD*,**;
 Colwell, Clifford E. Jr MD†; Ranawat, Chitranjan S. MD‡; Scott, Richard D. MD||; Thornhill, Thomas S. MD||; Lapp, Mark A. MD||, "In Vivo Anteroposterior Femorotibial Translation of Total Knee Arthroplasty: A Multicenter Analysis. Cl," *Clinical Orthopaedics and Related Research*, vol. 356 , pp. 47-57, 1998 .

- [96] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, "Mask R-CNN," <https://arxiv.org/abs/1703.06870>, 2018.
- [97] M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10778-10787, 2020.
- [98] D.-J. Kroon, "B-spline Grid, Image and Point based Registration," *MATLAB Central File Exchange*, 2023.
- [99] Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Trans Med Imaging*, vol. 18, no. 8, pp. 712-21, 1999.
- [100] Junchen Wang, Hideyuki Suenaga, Kazuto Hoshi, Liangjing Yang, Etsuko Kobayashi, Ichiro Sakuma, Hongen Liao, "Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery," *IEEE Trans Biomed Eng.*, vol. 61, no. 4, pp. 1295-304, 2014.
- [101] Dennis DA, Mahfouz MR, Komistek RD, Hoff W, "In vivo determination of normal and anterior cruciate ligament-deficient knee kinematics," *J Biomech*, vol. 38, no. 2, pp. 241-53, 2005.

- [102] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, Martial Hebert, "PCN: Point Completion Network," <https://arxiv.org/abs/1808.00671>, 2019.

Appendices

I. Extended Applications to THA and Patella

Figure 83 presents the cup and stem segmentation outcomes from the THA dataset, utilizing the same deep learning methodologies employed for TKA. Correspondingly, Figure 84 showcases the registration results of the THA implants based on the segmentation outcomes. Furthermore, Figure 85 displays the segmentation and registration results of the non-implanted knee's patella. These simulation results exemplify the potential application of the proposed algorithm to various components in the body, such as bones or implants.

II. Software User Interfaces

In this dissertation, the training procedures for all deep learning simulations are executed within the Anaconda environment, utilizing Python commands. The user interface is developed using Matlab's GUIDE and functions are executed through Matlab and Python. The software user interface is seamlessly integrated into Karbs, the established software employed in the CMR lab, as shown in Figure 86. However, in Figure 87, the template 3D model is superimposed but does not align accurately with the fluoroscopic image due to sizing inaccuracies. The 3D reconstruction windows are depicted in Figure 88 to reconstruct the designing 3D implant model.

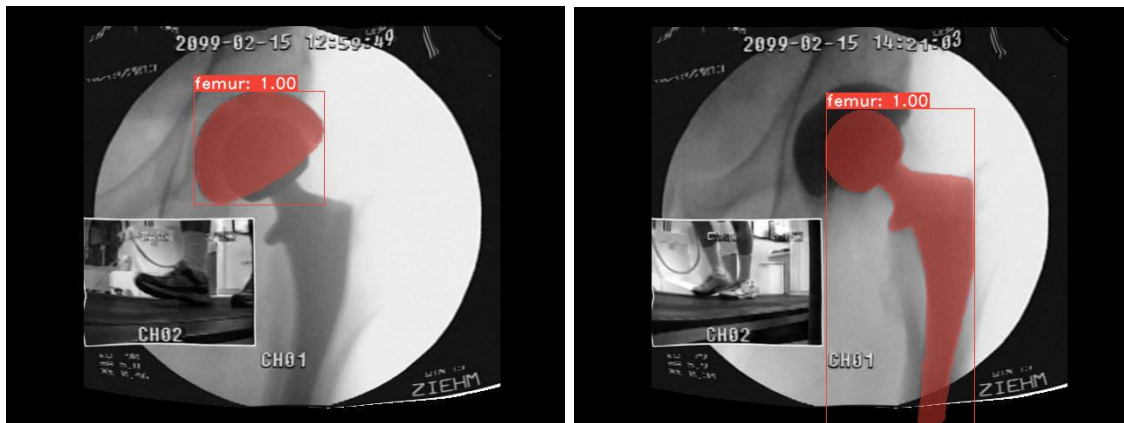


Figure 83. Segmentation results for the cup (left) and stem (right) of THA.

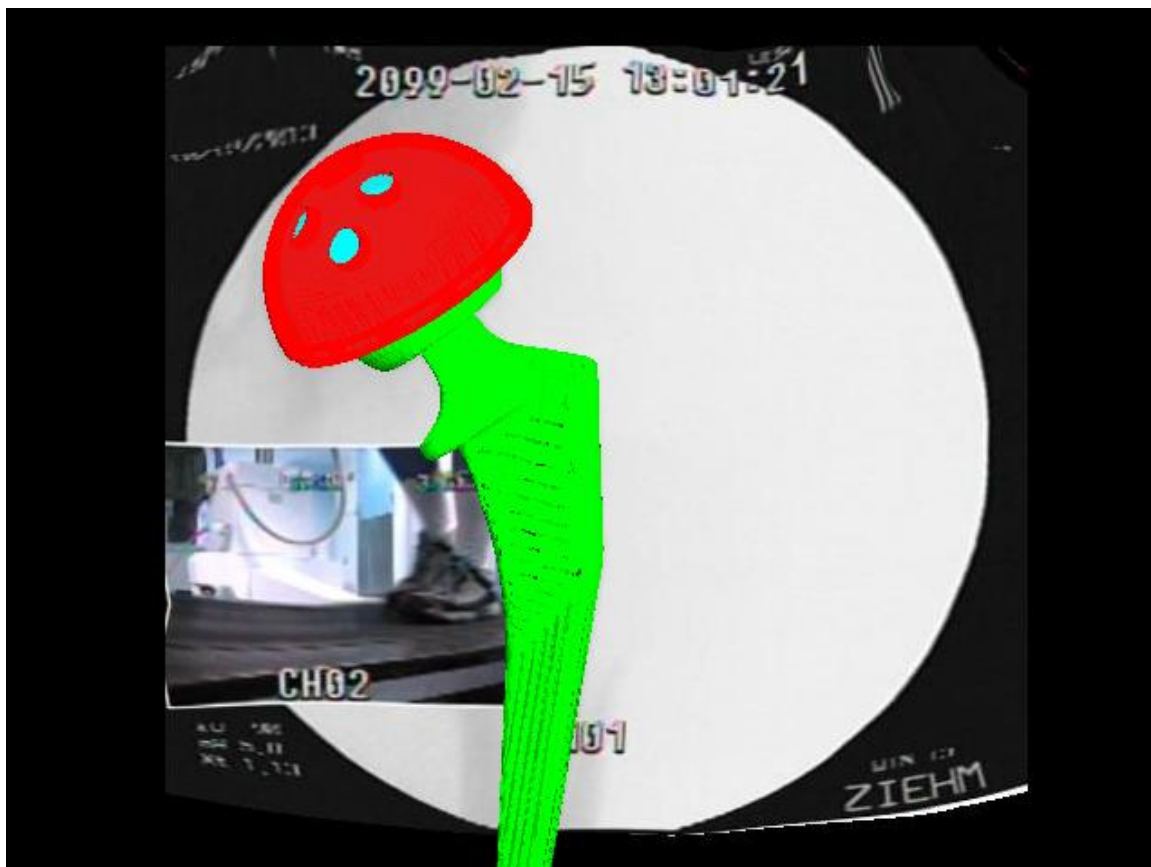


Figure 84. Registration results of the implants for THA.

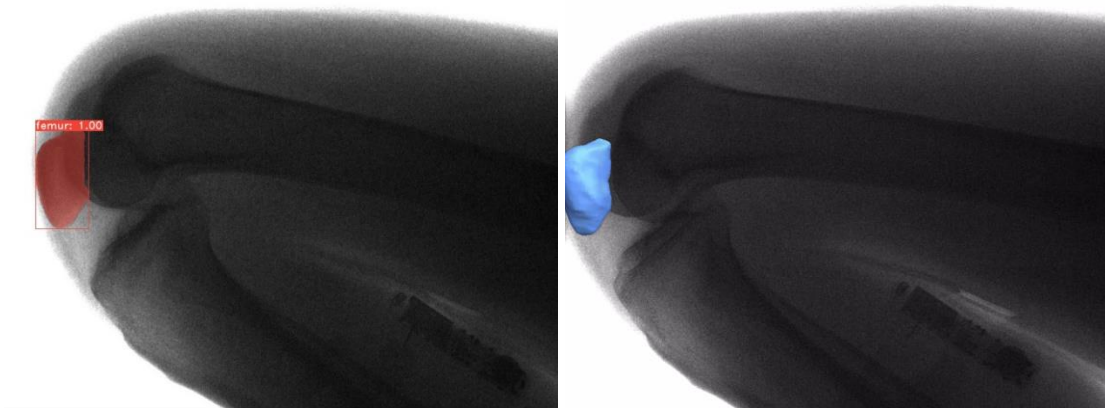


Figure 85. Segmentation and registration of the patella at the non-implanted knee.

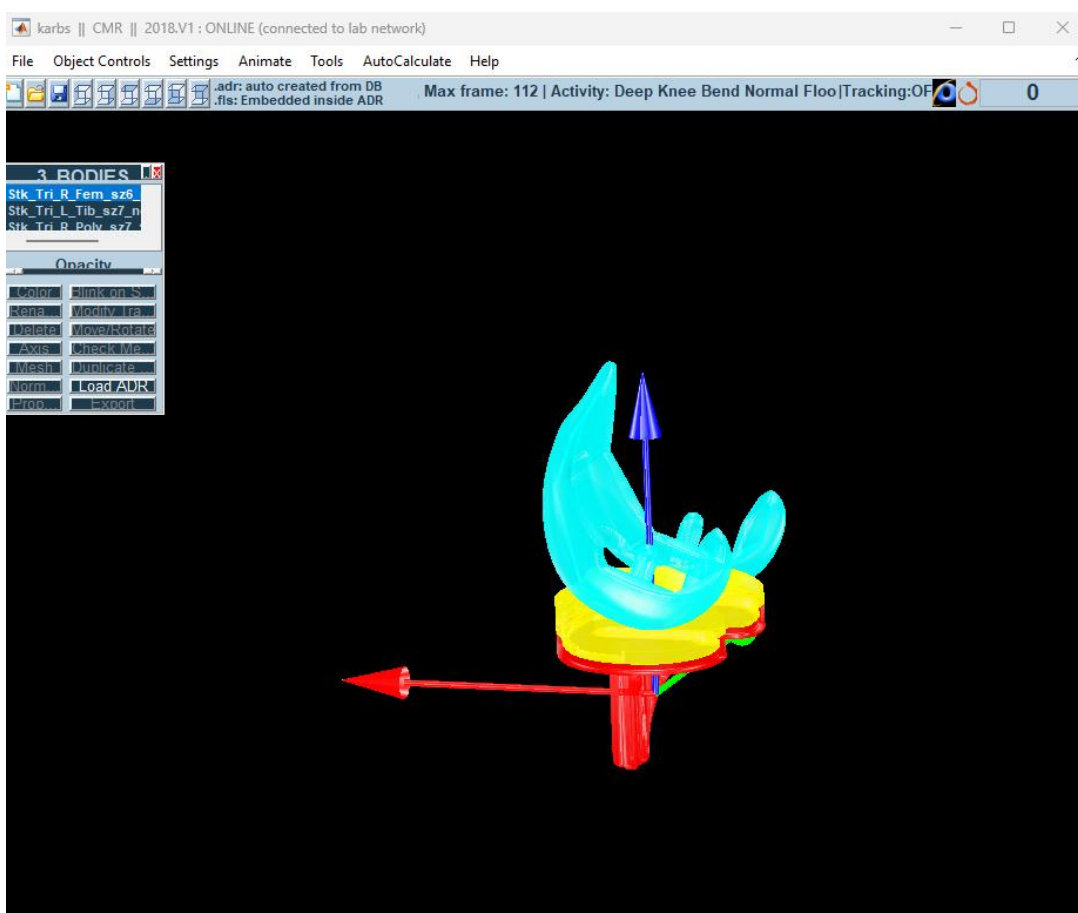


Figure 86. User interface of Karbs.

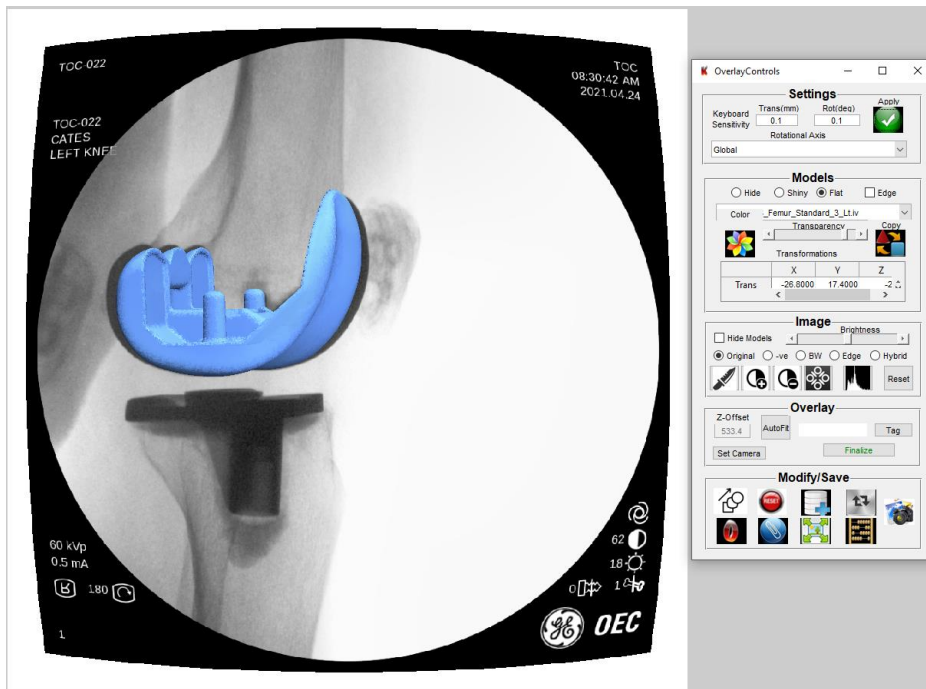


Figure 87. User interface for the registration in Karbs.

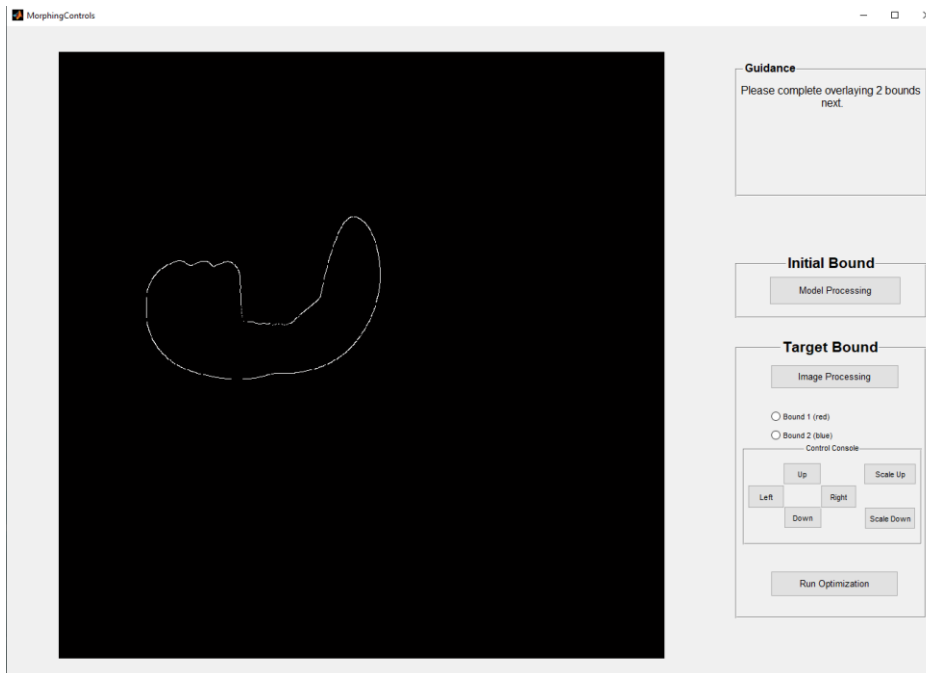


Figure 88. User interface for the 3D reconstruction tool.

Vita

Viet Dung Nguyen (Viet-Dung Nguyen) was born in Bac Ninh, Vietnam. As a child, he demonstrated his mechanical interests by disassembling and repairing his toys. During his study time at the school and university, he achieved a lot of olympiad awards in mathematics, mechanics, and informatics at the national level. In an academic journey, he obtained his bachelor's degree in the honors program in the Mechanical department at Hanoi University of Science and Technology in 2017. Following that, he gained valuable industry experience as a full-time research engineer for a year and a half, during which he secured a national patent and published a peer-reviewed journal article. Building on his success, he pursued his passion for learning and enrolled in graduate school at the University of Kansas, where he obtained his Master's degree from the Electrical Engineering and Computer Science department in 2020. Continuing his pursuit of knowledge, he completed his Doctor of Philosophy degree in Biomedical Engineering at the University of Tennessee at Knoxville in 2023.

A decade ago, during his sophomore years at Hanoi University of Science and Technology where he pursued mechanical engineering, Viet-Dung exhibited a keen interest in machine learning and its potential applications in the mechanical field. His Master's program at the University of Kansas provided him with a strong foundation in non-convex optimization and machine learning techniques. Pursuing his PhD at the CMR lab at the University of Tennessee equipped him with the necessary expertise to author a dissertation on the uses of machine learning and computer vision in biomechanics, thereby fulfilling

his long-held ambition. This accomplishment was made possible by the valuable mentorship of his advisors, the wealth of data and advanced equipment available in the research laboratory, the unwavering support of his family and friends, and his own tireless efforts. Undoubtedly, this dissertation marks a promising beginning towards the realization of his aspirations.