Automatic Personalized 3D model Reconstruction of the Distal Tibiofibular Joint from Biplanar Ankle X-Ray Images for Syndesmosis Assessment

by

Pejman HASHEMIBAKHTIAR

THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, NOVEMBER 18, 2025

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

Pejman Hashemibakhtiar, 2025

This Creative Commons licence allows readers to download this work and share it with others as long as the author is credited. The content of this work can't be modified in any way or used commercially.

BOARD OF EXAMINERS (THESIS PH.D.)

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Mr. Carlos Vázquez, Thesis Supervisor
Department of Software Engineering and IT at École de technologie supérieure

Mr. Jacques de Guise, Thesis Co-supervisor
Department of System Engineering at École de technologie supérieure

Mrs. Marie-Lyne Nault, Thesis Co-supervisor
Department of Surgery at Université de Montréal

Mr. Ismail Ben Ayed, President of the Board of Examiners
Department of System Engineering at École de technologie supérieure

Mr. Christian Desrosiers, Member of the jury
Department of Software Engineering and IT at École de technologie supérieure

Mr. Laurent Gajny, External Evaluator
Arts et Métiers Sciences et Technologies
Institut de Biomécanique Humaine Georges Charpak

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND PUBLIC

OCTOBER 29, 2025

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisor, Professor Carlos Vázquez, for accepting me as his student and for placing his trust in me throughout the completion of this project and my PhD. It has been a remarkable opportunity to participate in such a challenging and enriching project, which has significantly broadened my knowledge. I am truly grateful to have successfully completed it under his guidance. This work would not have been possible without the invaluable support of my project manager, Thierry Cresson, whose kindness and constant encouragement accompanied me through every step of this project. Their keen attention to detail, patience, and guidance not only directed me through the challenges but also taught me the true meaning of support.

I am also deeply thankful to my co-supervisors, Professor Jacques A. de Guise, who generously continued to guide me even after his retirement, despite having already contributed so much to my project and to our laboratory, the LIO, and Dr. Marie-Lyne Nault, for her valuable medical guidance and continuous support throughout this journey.

I would also like to thank the members of my thesis jury for dedicating their time to review my work and for their valuable comments and suggestions, which helped to strengthen this thesis.

I wish to acknowledge all the members of the LIO lab who supported me over these years: Lulu, Anaïde, Iris, Amiel, Yolène, Alice, Camille, Margaux, Lucas, Nahid, Ted, Marta, Marie, Melissa, Christine and Benjamin. I would also like to thank Karine at Hôpital du Sacré-Cœur for her valuable help in delivering the data. A very special thanks goes to Professor Nicola Hagemeister, the Head of the LIO, whose help in acquiring the data allowed me to overcome one of the greatest obstacles of my project.

I am very grateful to my dear friends, Nima and Bita, who became like family to me when I first arrived in Canada. Their presence and support made my early days much easier than expected and helped me begin my new life in Montreal smoothly.

I would also like to extend my heartfelt thanks to my family, my parents, who have supported me unwaveringly throughout these years of study, and my sisters, Nahid and Mitra, and my brother, Farshid. I am deeply appreciative of everything you have done for me.

Finally, I reserve my final and most special words for my beloved partner, *Elaheh*, who has supported me in every possible way, with infinite patience and love throughout this journey. This achievement would not have been possible without her by my side.

Reconstruction Automatique et Personnalisée de Modèle 3D de l'Articulation Tibiofibulaire Distale à Partir d'Images Radiographiques Biplanaires de la Cheville pour l'Évaluation de la Syndesmose

Pejman HASHEMIBAKHTIAR

RESUMÉ

L'articulation tibiofibulaire joue un rôle clé dans la stabilité de la cheville, et ses blessures, notamment les lésions de la syndesmose, peuvent entraîner des déficiences fonctionnelles majeures. Une évaluation précise est essentielle pour le diagnostic et la planification du traitement. Les techniques d'imagerie traditionnelles, telles que les radiographies conventionnelles, la tomodensitométrie (TDM) et l'imagerie par résonance magnétique (IRM), présentent des limites en termes d'exposition aux radiations, de coût et de prise en charge en position debout. Le système d'imagerie EOS, qui fournit des radiographies bi-planaires avec une faible dose de radiation en position debout, constitue une alternative prometteuse. Cependant, sa mise en œuvre pour la reconstruction 3D de l'articulation tibiofibulaire est complexe en raison de la projection bidimensionnelle, qui entraîne une perte d'information tridimensionnelle, de la superposition des structures et des faibles distances inter-osseuses.

Cette thèse propose une méthode innovante pour la reconstruction 3D personnalisée de l'articulation tibiofibulaire à partir de radiographies bi-planaires. L'approche combine la modélisation statistique de formes et d'intensité avec des techniques d'enregistrement 2D/3D basées sur l'apprentissage profond pour une reconstruction articulaire précise. Un modèle statistique de formes et d'intensité (SSIM) est développé à partir d'une base de données de scanners CT, garantissant une représentation anatomiquement cohérente. Le pipeline de reconstruction comprend l'extraction automatique de l'articulation à partir des radiographies à l'aide de réseaux neuronaux profonds de type auto-encodeur, suivie d'un cadre d'enregistrement 2D/3D basé sur l'apprentissage profond en une seule étape.

Les résultats expérimentaux confirment l'efficacité de la méthode proposée pour reconstruire des modèles 3D précis de l'articulation tibiofibulaire à partir de radiographies bi-planaires de la cheville. Les modèles reconstruits permettent le calcul automatique de paramètres cliniques pertinents pour l'évaluation de la syndesmose, offrant ainsi une alternative non invasive aux scanners CT. Cette recherche fait progresser le diagnostic orthopédique personnalisé en exploitant des techniques de reconstruction basées sur l'intelligence artificielle, améliorant la détection des blessures et les résultats des traitements.

Mots-clés : Articulation tibiofibulaire, cheville, reconstruction 3D, radiographies bi-planaires, apprentissage profond, modèle statistique de formes et d'intensité, lésion de la syndesmose, enregistrement 2D/3D, imagerie EOS

Automatic Personalized 3D model Reconstruction of the Distal Tibiofibular Joint from Biplanar Ankle X-Ray Images for Syndesmosis Assessment

Pejman HASHEMIBAKHTIAR

ABSTRACT

The tibiofibular joint plays a critical role in ankle stability, and its injuries, particularly syndesmosis injuries, can lead to significant functional impairments. Accurate assessment of such injuries is crucial for diagnosis and treatment planning. Traditional imaging techniques, including conventional radiographs, Computed Tomography (CT) scans, and Magnetic Resonance Imaging (MRI), each have limitations in terms of loss of 3D information, radiation exposure, cost, and the ability to capture weight-bearing conditions. The EOS Imaging system, which provides biplanar X-ray images with low radiation in a standing position, presents a promising alternative. However, its application in 3D reconstruction of the tibiofibular joint remains challenging due to factors such as two-dimensional projection, which leads to a loss of three-dimensional information and superimposition of structures, and the small inter-bone distances of the tibiofibular joint.

This thesis proposes a novel method for personalized 3D reconstruction of the tibiofibular joint from biplanar radiographs. The approach integrates statistical shape and intensity modeling and deep learning-based 2D/3D registration techniques to achieve accurate joint reconstruction. A statistical shape and intensity model (SSIM) of the tibiofibular joint is developed using a dataset of CT scans, ensuring an anatomically coherent representation. The reconstruction pipeline includes automatic extraction of the joint from radiographs using deep autoencoder-based neural networks, followed by a one-shot deep-learning-based 2D/3D registration framework.

Experimental results demonstrate the effectiveness of the proposed method for reconstructing anatomically accurate 3D models of the tibiofibular joint from ankle biplanar radiographs. The reconstructed models enable automatic calculation of clinically relevant syndesmosis parameters, offering a non-invasive alternative to CT scans for injury assessment. This research contributes to the advancement of personalized orthopedic diagnostics by leveraging AI-driven reconstruction techniques, ultimately improving injury detection and treatment outcomes.

Keywords: Tibiofibular joint, ankle, 3D reconstruction, biplanar radiographs, deep learning, statistical shape and intensity model, syndesmosis injury, 2D/3D registration, EOS imaging

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABREVIATIONS AND ACRONYMS

2D	Two-Dimensional
2D/3D	Two-Dimensional to Three-Dimensional
3D	Three-Dimensional
AI	Artificial Intelligence
AITFL	Anterior Inferior Tibiofibular Ligament
AP	Anteroposterior
CNN	Convolutional Neural Network
CT	Computed Tomography
Deg	Degrees
DF	Distal Fibula
DLT	Direct Linear Transform
DP	Descriptive Parameters
DRR	Digitally Reconstructed Radiograph
DTF	Distal Tibiofibular
FFD	Free-Form Deformation
GPA	Generalized Procrustes Analysis
ICP	Iterative Closest Point
IOL	Interosseous Ligament
JE-Net	Joint Extraction Network
LT	Lateral
MCS	Medial Clear Space
MDL	Minimum Description Length

Mm	Millimeter
MOV	Modes of Variation
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NSCC	Non-Stereo Corresponding Contours
NSCP	Non-Stereo-Corresponding Points
P2P	Point-to-Point
P2S	Point-to-Surface
PCA	Principal Component Analysis
PDM	Point Distribution Model
PITFL	Posterior Inferior Tibiofibular Ligament
PSNR	Peak Signal-to-Noise Ratio
ROI	Region of Interest
SCP	Stereo-Corresponding Points
SGD	Stochastic Gradient Descent
SSIM	Statistical Shape and Intensity Model
SSiM	Structural Similarity Index Measure
SSM	Statistical Shape Model
TFCS	Tibiofibular Clear Space
TFO	Tibiofibular Overlap
TP	Tibial Plafond
US	Ultrasound

LIST OF SYMBOLS

x, y, z	Coordinates of a vertex in 3D space
CT_i	The ith CT volume
CT_i^{seg}	The ith segmented CT volume
$ankle_i$	The ith segmented CT volume representing the ankle (bones and soft tissues)
$joint_i^{vol}$	The ith segmented CT volume representing the joint (tibia and fibula)
$tibia_i^{vol}$	The ith segmented CT volume representing the tibia
$fibula_i^{vol}$	The ith segmented CT volume representing the fibula
$tibia_i^{mesh}$	The ith tibia mesh generated from its volume
$fibula_i^{mesh}$	The ith fibula mesh generated from its volume
$tibia_{target}^{mesh}$	The target tibia mesh used for non-rigid registration of the other tibia meshes
$fibula_{target}^{mesh}$	The target fibula mesh used for non-rigid registration of the other fibula meshes
$tibia_i^{m-mesh}$	The ith tibia mesh which is matched to the target tibia mesh (results in the same topology)
$fibula_i^{m-mesh}$	The ith fibula mesh which is matched to the target fibula mesh (results in the same topology)
$tibia_i^V$	The ith tibia mesh vertices matrix
$fibula_i^V$	The ith fibula mesh vertices matrix
$tibia_i^F$	The ith tibia mesh faces matrix
$fibula_i^F$	The ith fibula mesh faces matrix
$joint_i^V$	The ith joint mesh (combined tibia and fibula) vertices matrix
$joint_i^F$	The ith joint mesh (combined tibia and fibula) faces matrix
$joint_i^{mesh}$	The ith joint mesh structure
$joint_i^{mesh^{aligned}}$	The ith joint mesh structure aligned with other meshes in an iterative process

$joint_i^{mesh^{EOS}}$	The ith joint mesh structure aligned with other meshes in an iterative process and aligned to the face-front positioning similar to the EOS cabin
$joint_{mean}^{mesh^{EOS}}$	The mean mesh of the joint meshes aligned with the EOS cabin face-front direction
Φ	Eigenvectors matrix
ψ_i	The ith eigenvector
Λ	Eigenvalues matrix
λ_i	The ith eigenvalue
N	Number of shapes in the dataset
C	Covariance matrix
K	Number of vertices in a mesh
b_i	The ith shape parameter
k_max	Maximum number of shape parameters chose for the SSIM
I^{simple}	The simple intensity value as the average for tibia and fibula volumes
rot_x, rot_y, rot_z	Rotation of the mesh around axes X, Y and Z, respectively
S	Scale coefficients
$joint_{Reconstructed}^{V}$	The reconstructed joint's vertices matrix
$joint_{Ground\_Truth}^{V}$	The ground truth for reconstructed joint's vertices matrix
$N(0, \lambda_k)$	Normal distribution
$joint_i^{mesh^{DB}}$	The ith generated mesh from the model to form the training database
$b_k^{i\ normalized}$	The ith normalized shape parameter
y_{pred}	Generated image using the denoising autoencoder
y_{true}	Ground truth for generated image using the denoising autoencoder
α	The weight parameter, controlling the combination of loss functions
$ankle_i^{EOS}$	The ith segmented CT volume representing the ankle (bones and soft tissues) aligned with the EOS cabin face-front direction
$joint_i^{vol^{EOS}}$	The ith segmented CT volume joint (tibia and fibula) aligned with the EOS cabin face-front direction

INTRODUCTION

The tibiofibular joint, a key component of the ankle, plays a crucial role in weight-bearing and stability. Injuries to this joint, particularly syndesmosis injuries, can significantly impair mobility and function (Paez, Lurie, Upasani, & Pennock, 2021). Proper diagnosis and treatment planning require accurate imaging techniques, yet current methods have limitations (Kellett, Lovell, Eriksen, & Sampson, 2018a). Conventional radiographs, while widely accessible, lack three-dimensional (3D) information, which causes low intraclass correlation values for some measurements (Dhont et al., 2023). Computed Tomography (CT) scans provide high-resolution 3D images but expose patients to significant radiation. Magnetic Resonance Imaging (MRI) offers detailed soft tissue visualization but is costly, has a long processing time and does not capture weight-bearing conditions, similar to CT scans. The weight-bearing position allows natural loading conditions to be applied to the ankle, facilitating a more accurate diagnosis that reflects the joint's behavior during everyday activities.

EOS Imaging, a low-radiation biplanar X-ray system, has emerged as a promising alternative for obtaining a three-dimensional (3D) model from joints and bones for orthopedic diagnostics. It retrieves two biplanar radiographs from the patient’s body while they are in a standing position. While the EOS Imaging software enables 3D model generation for the spine, hip, and knee, it has not yet been extended to the ankle joint, where significant challenges persist. These include structural superimposition, tightly connected anatomical components, and the inherent loss of depth information in 2D projections. As a result, reconstructing anatomically accurate 3D models of the ankle from biplanar radiographs requires advanced computational techniques.

Thus, this work aims to reconstruct a 3D model of the patient’s ankle joint, specifically the tibiofibular joint from biplanar radiographs, with the clinical objective of enabling the calculation of relevant ankle joint measurements.

Chapter 1 presents the clinical context motivating this thesis, followed by Chapter 2, which provides a comprehensive literature review of the 3D reconstruction approaches from biplanar radiographs and its related steps. Chapter 3 outlines the research problem and defines the objectives of the study. In Chapter 4, we detail the theoretical framework and methodology

employed for the 3D reconstruction of the tibiofibular joint from ankle radiographs, while Chapter 5 presents the experimental results. The thesis is completed with Chapter 6, which offers a discussion of the findings, and the overall conclusion and recommendations are presented in Chapter 7.

CHAPITRE 1

CLINICAL CONTEXTS AND 3D RECONSTRUCTION OF THE TIBIOFIBULAR JOINT FROM BIPLANAR X-RAY IMAGES

This chapter provides an overview of the ankle's structure, focusing on both its skeletal and soft tissue components. It also discusses how injuries to the syndesmosis can occur based on the anatomical structure and outlines the different degrees of such injuries. Following this, diagnostic techniques utilizing medical imaging are explored, with particular emphasis on the potentials for diagnosis syndesmosis injuries through Three-Dimensional (3D) reconstruction of the Tibiofibular joint structure.

1.1 Introduction

Ankle injuries are among the most common traumatic events affecting the lower limbs, often involving the lateral ligamentous complex, particularly the lower lateral ligaments (Lamer et al., 2018). In 18% of cases, the syndesmotic structures may also be affected (Lamer et al., 2018). The severity of instability in the syndesmosis depends on which ligaments are affected (Lamer et al., 2018). To fully understand the types of traumas that can impact the ankle ligaments, it is important to first examine the anatomy and biomechanics of the ankle in more detail.

1.2 Ankle joint anatomy

The ankle joint is composed of three bones: the distal tibia, distal fibula, and talus (Figure 1.1). Muscles, tendons, and ligaments are distributed along the surfaces of the bones, facilitating the complex movements essential for motion and maintaining balance. Three primary ligaments connect tibia and fibula consistently (Figure 1.2), forming the distal syndesmotic articulation (Lin, Gross, & Weinhold, 2006): the anterior inferior tibiofibular ligament (AITFL), the posterior inferior tibiofibular ligament (PITFL), and the interosseous ligament (IOL). These ligaments offer significant stability to the joint, limiting the external rotation of the fibula to 2°

relative to the tibia, while the ankle mortise only expands by approximately 1 mm as the ankle transitions from full plantar flexion to full dorsiflexion (Lin et al., 2006).

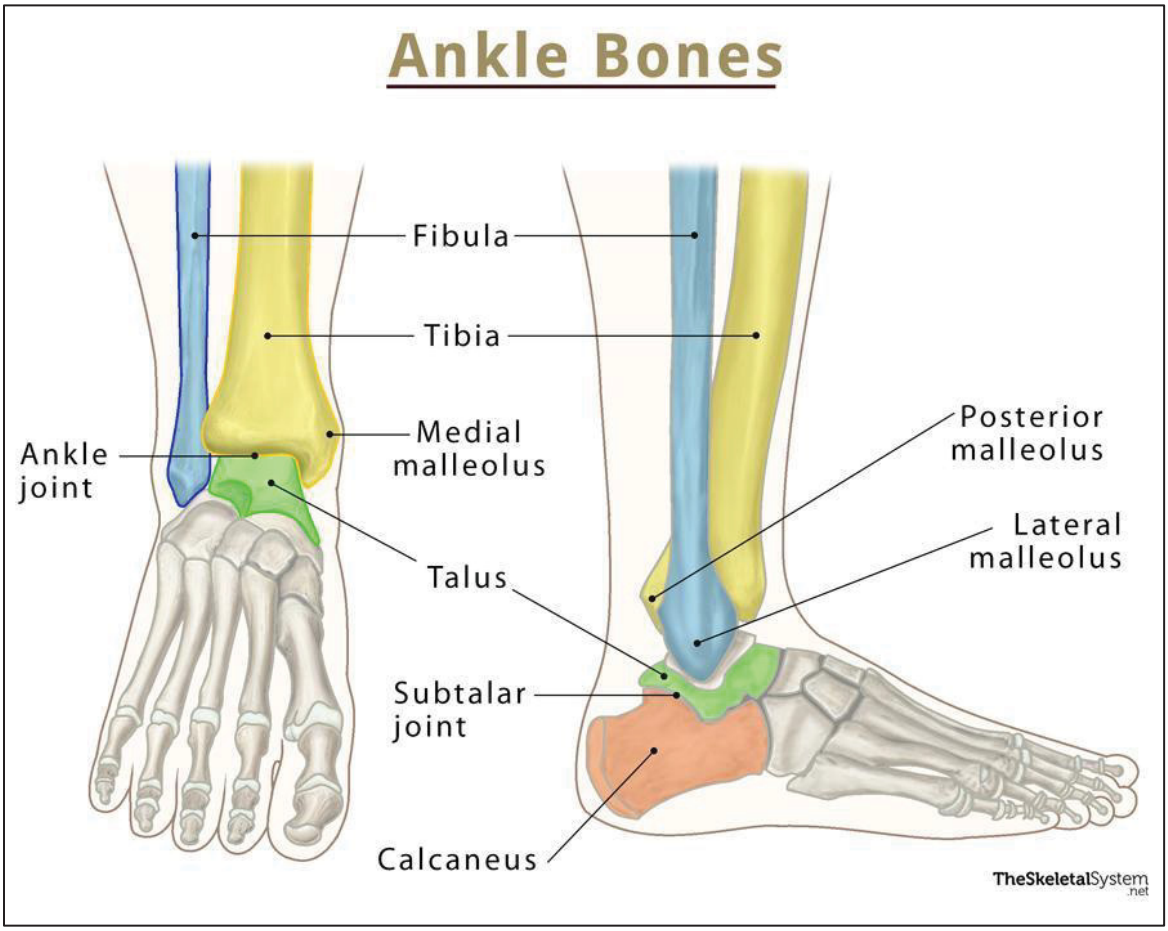

Figure 1.1 Ankle Bones
Taken from The Skeletal System (2022)

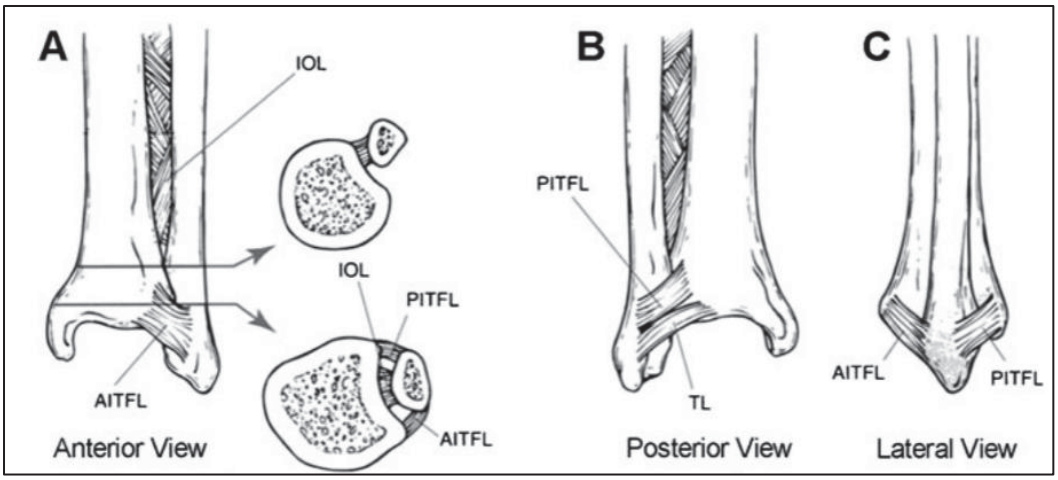

Figure 1.2 Major ligaments forming the Ankle joint
Taken from Hunt (2013)

1.3 Syndesmosis injury

Excessive rotation or force applied on the ankle may cause tears or ruptures to the ligaments, making the articulation unstable. These forces can lead to widening between the fibula and tibia at the ankle mortise, disrupting the ligaments (Hunt, 2013). Depending on the severity and spot of the injury, AITFL, PITFL, IOL or other ligaments could be affected (Figure 1.3).

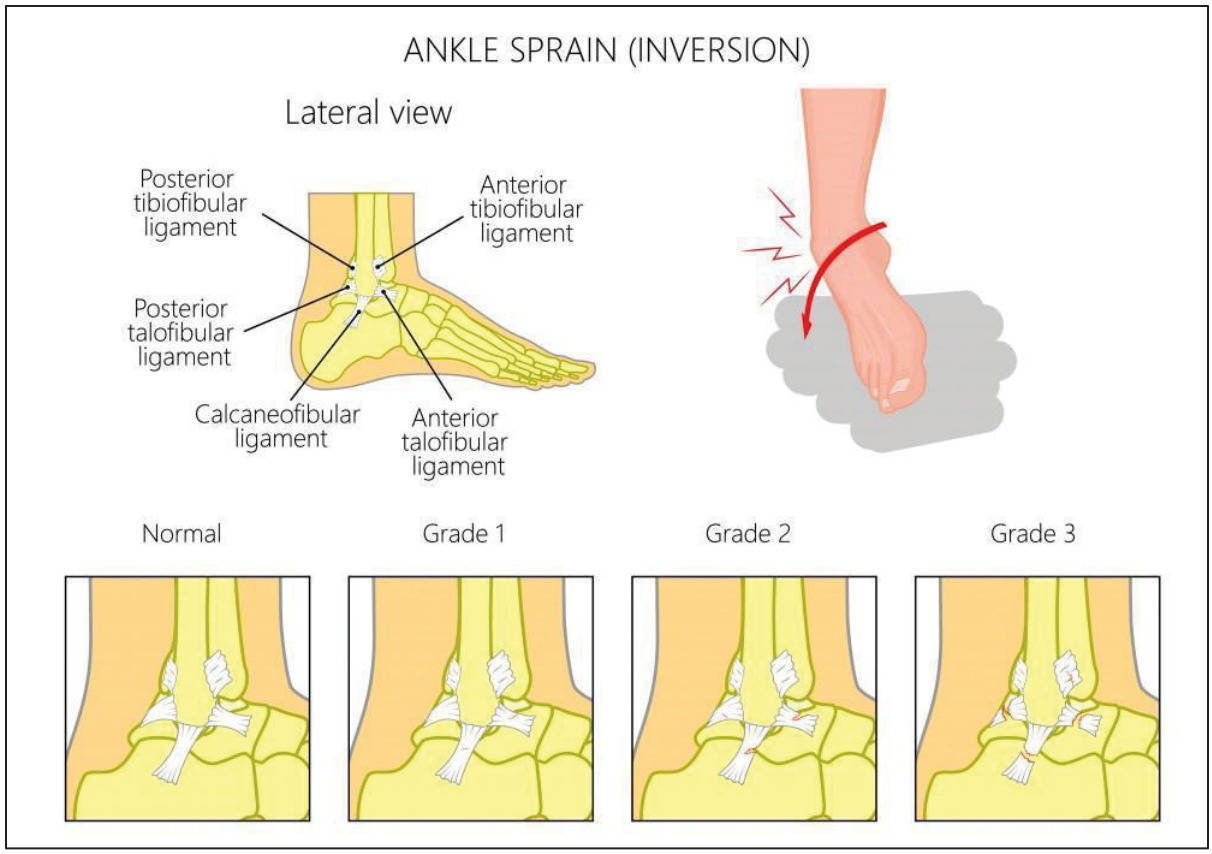

Figure 1.3 Ankle sprain and ligaments injuries
Taken from Eugene Stautberg III, MD (2025)

In the context of syndesmosis injury, it is widely accepted that there are three grades of injuries, ranging from grade 1 to grade 3, with grade 1 being the least severe and grade 3 the most severe (Hunt, 2013):

- Grade 1: Injuries are clinically mild and characterized by a stable syndesmotic joint. These injuries involve partial damage to the lateral ligaments.
- Grade 2: A complete disruption of the AITFL and/or the interosseous ligament (IOL) causes grade 2 of syndesmosis injury.
- Grade 3: A complete injury to lateral ligaments (AITFL, IOL, PITFL) leads to an unstable articulation.

1.4 Diagnosis of the syndesmosis injury using Medical Imaging

Various medical imaging techniques are available for the non-invasive diagnosis of injuries, each with its own set of characteristics, advantages, and limitations. An essential step in diagnosing syndesmotic injuries is selecting the most appropriate imaging method to accurately identify ligament damage or ankle instability (Kellett, Lovell, Eriksen, & Sampson, 2018b). Recent comprehensive literature reviews have compared different imaging modalities for diagnosing syndesmotic injuries (Kellett et al., 2018b; Krähenbühl et al., 2018). X-ray imaging with and without stress, Ultrasound, Computed Tomography (CT) scans, Nuclear Medicine scans, Magnetic Resonance Imaging (MRI) and Arthroscopy are the techniques used for diagnosing syndesmotic injuries (Kellett et al., 2018b). Among them, weight-bearing radiographs in Antero-Posterior (AP) and Mortis view, Ultrasound, CT scan and MRI are the most popular imaging techniques used for this purpose (van Dijk et al., 2016). In the following sections, we will explore these techniques and their approach for identifying syndesmotic injuries

1.4.1 Weight-Bearing Conventional Radiographs

In conventional X-ray imaging techniques, the patient's body is exposed to X-ray to capture the skeletal structure on the radiograph. On radiographs, soft tissues are not clearly visible, with the focus primarily on the bony structures. However, syndesmotic widening can indicate injuries to the ligaments (Figure 1.4). Therefore, observing the relationship between the distal Tibia and distal Fibula can provide indirect evidence of a syndesmosis injury, when the bones are separated more than a normal distance (Kellett et al., 2018b). A summary of the measurements used to assess the syndesmotic injury on conventional radiographs is presented in Figure 1.5 (Krähenbühl et al., 2018). Among these parameters, three are most commonly used for diagnosing syndesmotic injuries (Krähenbühl et al., 2018) :

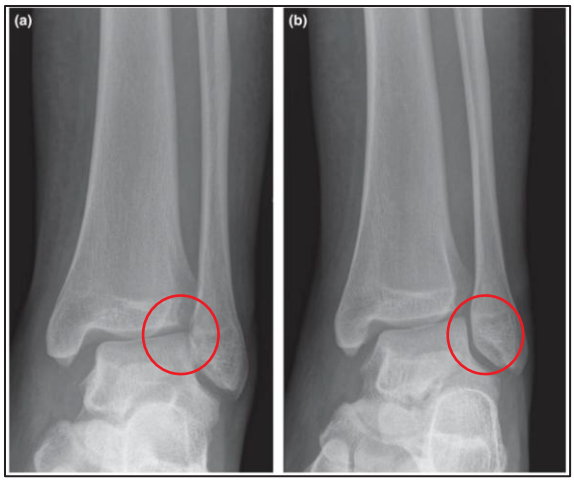

Figure 1.4 Ankle radiographs, AP (a) and Mortis (b) view demonstrating the syndesmotic widening and widening of medial clear space in red circles
Taken from Kellett et al. (2018b)

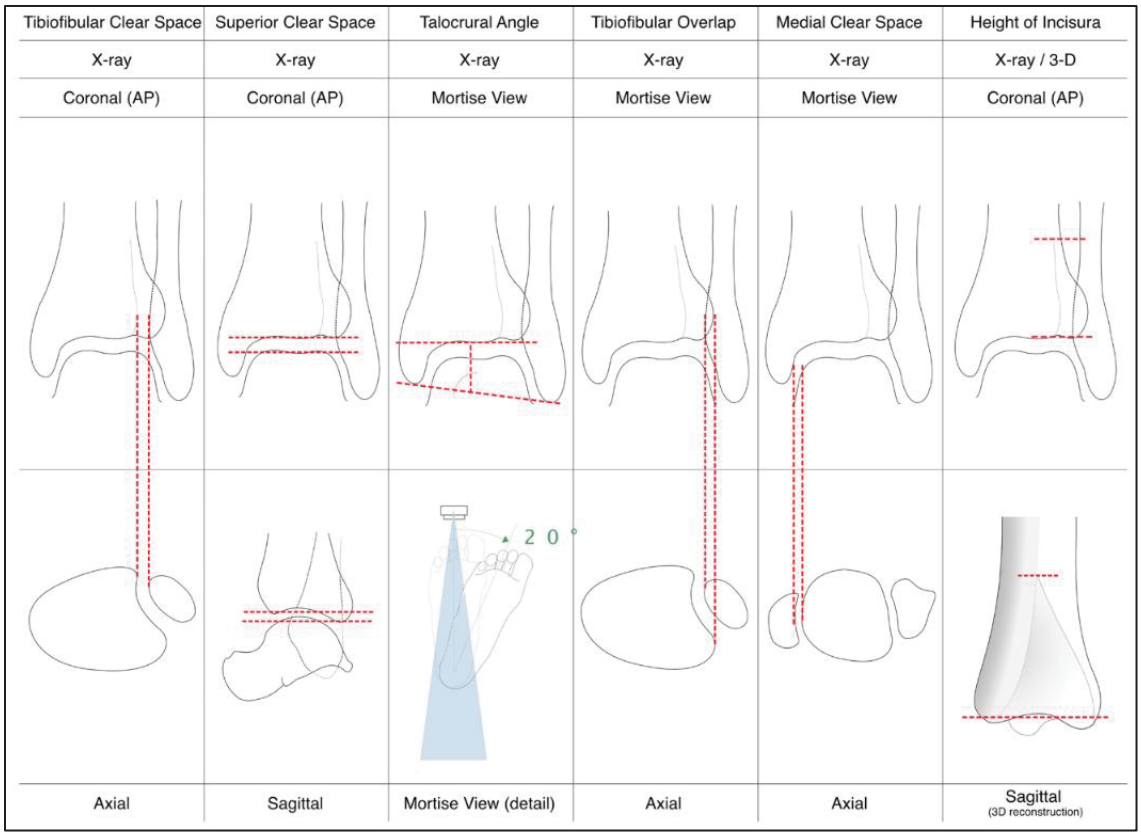

Figure 1.5 Summary of the defined measurements for syndesmosis injury assessment on radiographs
Taken from Krähenbühl et al. (2018)

- Tibio-Fibular Clear Space (TFCS): This metric measures the distance between the medial border of the fibula and the lateral border of the peroneal incisura of the tibia, measured from the AP view 1 cm above the tibial plafond (Kellett et al., 2018b). Based on various studies, the normal value for this measurement is reported to be typically less than 6 mm (Kellett et al., 2018b), with values ranging between 1.8 mm and 6.5 mm (Krähenbühl et al., 2018).
- Tibio-Fibular Overlap (TFO): This is the horizontal distance between the medial border of the fibula and the lateral border of the anterior tibial tubercle (Kellett et al., 2018b). It is also measured 1 cm above the tibial plafond. The normal value for this parameter is reported as at least 1 mm on the mortis view and 6 mm on the AP view (Kellett et al., 2018b). However, other studies report values ranging from 2.2 mm to 4.7 mm in the mortis view and 8 mm to 9.4 mm in the AP view (Krähenbühl et al., 2018)..
- Medial Clear Space (MCS): This refers to the widest distance between the medial borders of the malleolus, measured in the mortis view (Kellett et al., 2018b) and the normal value for this metric is greater than 4 mm (Kellett et al., 2018b). However, other studies report a range between 3 mm and 5.5 mm for this parameter (Krähenbühl et al., 2018).

Based on the statistics taken from different studies, the authors concluded that one may not solely rely on measurements taken from conventional radiographs for the assessment of the injury (Krähenbühl et al., 2018). The issue arises from the fact that in conventional radiographs, structures are superimposed during the projection process, which results in reduced contrast on the images (Michael, 2001). Thus, finer details may not be clearly visible, potentially leading to inaccuracies in assessing the relationship between the distal Tibia and distal Fibula. We may overcome this limitation using CT scans instead of conventional radiographs. CT scans were developed to provide depth information by reconstructing a 3D image of the patient's body using X-ray technology (Michael, 2001).

1.4.2 Computed Tomography scan

Computed Tomography (CT) Scan produces 2D images of trans-axial planes (slices) by rotating the X-ray source and detector around the patient's body (Figure 1.6), while they are in a lying position (Michael, 2001) . Combining the generated images produces one slice of CT which helps eliminate structural overlap. When this process is applied to the entire body, it results in a comprehensive 3D image of the patient's body (Michael, 2001).

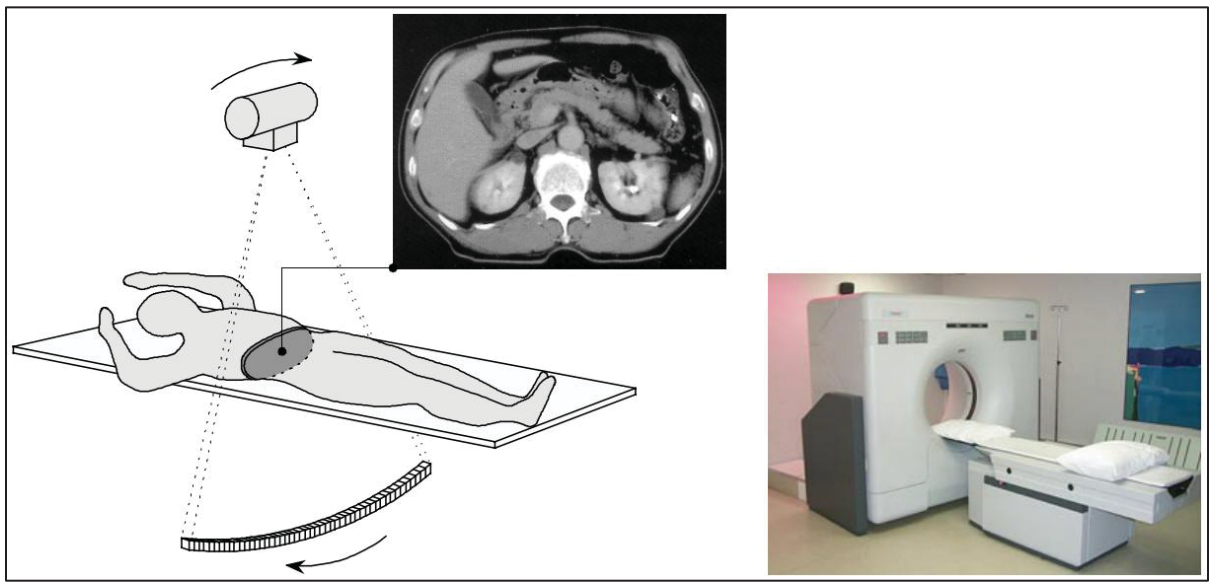

Figure 1.6 X-ray Computed Tomography (CT)
Taken from Michael (2001)

Since CT scans are based on X-ray imaging technology, the assessment of syndesmotic injury should still focus on the relationship between the bones, similar to the assessment strategy in radiographs. However, unlike X-ray images, CT scan slices offer high-contrast images with non-superimposed structures, providing clearer details for more accurate analysis (Michael, 2001). As a summary of recent studies, syndesmotic parameters are calculated on the CT slice located 1 cm above the Tibial Plafond (Krähenbühl et al., 2018). In the other study, a set of six measurements and two angles have been proposed to evaluate syndesmosis on specific CT slices (Nault, Hébert-Davies, Laflamme, & Leduc, 2013). Figure 1.7 illustrates these

measurements and angles, while Table 1.1 provides a description of the parameters. These measurements are taken from the CT slice located 9.45 mm proximal to the tibial plafond slice (15 slices in the used dataset), except for angle 2, which is measured on the Talar dome CT slice. All the measurements are made on non-pathological syndesmosis cases. The measurements range from 0.8 mm to 13 mm.

Among the parameters defined in Table 1.1, A, B and C are the main parameters to assess the syndesmosis injury (Nault et al., 2013). For the observed non-pathological cases, parameters A, B and C are ranging from 1.5 mm to 6.44 mm, 2.1 mm to 11.5 mm, and 1 mm to 5.8 mm respectively (Nault et al., 2013).

Although the metrics on the CT slice are measured more accurately than those on radiographs, CT scan exposes the patient's body to much more radiation than conventional radiograph as it takes many images to reconstruct the 3D volume. In addition, in a conventional CT scan, the patient is lying down (Figure 1.6) and therefore, weight-bearing load is not applied on the ankle which could affect the assessment process (Dhont et al., 2023).

Although CT is inherently a 3D imaging modality, the referenced study performed measurements on a single slice, thereby restricting the analysis to a 2D plane. This limitation in capturing the full three-dimensional anatomy may contribute to a lower intraclass correlation coefficient for certain syndesmosis measurements (Dhont et al., 2023), highlighting the potential for improved accuracy through true 3D assessments.

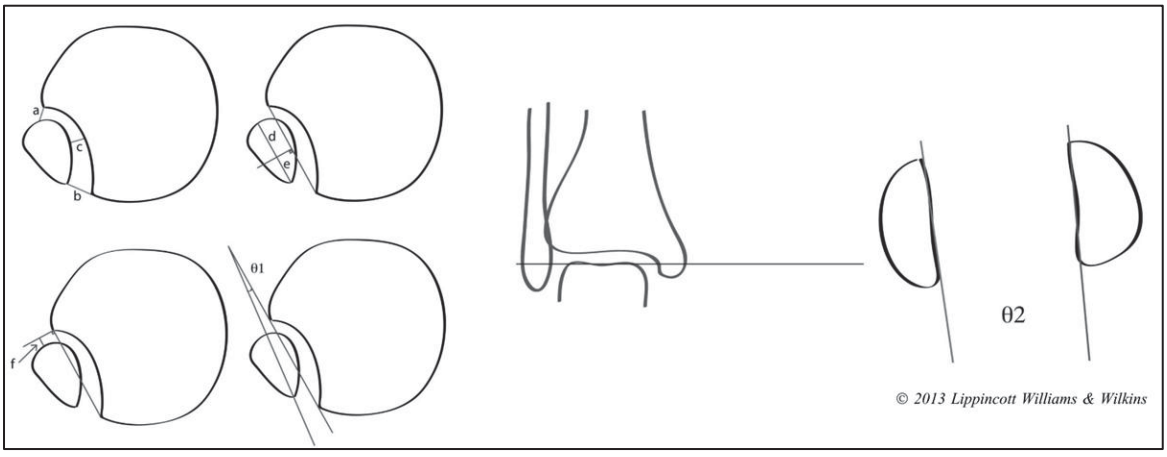

Figure 1.7 Syndesmosis clinical parameters 'a' to 'f' and 'angle 1' on an axial view, shown 9.45 mm above the Tibial Plafond. The 'angle 2' shown on axial view of the two malleoli at talar dome
Taken from Nault et al. (2013)

Table 1.1 Syndesmosis clinical parameters description (shown in Figure 1.7)
Taken from Nault et al. (2013)

Parameter	Description
a	Distance between the most anterior point of the incisura and the nearest most anterior point of the fibula
b	Distance between the most posterior point of the incisura and the nearest most posterior point of the fibula
c	Distance between the tibia and the fibula in the middle of the incisura
d	3 step measurements A line is drawn between the most anterior and most posterior point of the incisura A perpendicular line is drawn in the middle of the first line Distance between the anterior part of the fibula and the perpendicular
e	Distance between the posterior part of the fibula and the perpendicular in line with the anterior measure
f	The same perpendicular line is brought at the level of the most anterior point of the incisura Distance between that line and the most anterior point of the fibula
angle 1	Angle between a line drawn between the anterior and posterior point of the incisura and a line drawn in the fibula representing its orientation (internal rotation being a negative angle)
angle 2	This angle is measure at the level of the talar dome. The angle is between the talar side of the 2 malleoli

1.4.3 Stress X-ray

In clinical studies, stress is applied to the patient's ankle while monitoring his/her reaction and joint response, allowing for a real-time assessment of syndesmotic instability. Latent

syndesmosis may not be detectable on plain X-rays and may not significantly impact ankle functionality. However, surgical intervention is often required to repair the syndesmosis.
Diagnosing syndesmotic injuries through stress X-ray imaging requires highly experienced experts, as interpretation can be complex (LaMothe et al., 2018). However, the results often lack consistent reproducibility, making objective and standardized assessment challenging.

1.4.4 Ultrasound

Dynamic ultrasound imaging can be used to detect syndesmosis injuries by having the physician hold the ultrasound probe while the patient moves the relevant joint (Kellett et al., 2018a). In ultrasound imaging, the superficial structure of the anterior inferior tibiofibular ligament (AITFL) is typically visible. As a result, syndesmotic injuries involving this ligament can be detected using this approach.
This technique allows for real-time assessment of ligament integrity and joint stability. However, similar to the stress X-ray, the accuracy of this method is highly dependent on the expertise of the physician, as proper probe positioning and image interpretation require significant experience.

1.4.5 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) uses very strong magnetic fields (Figure 1.8), which are effective in visualizing soft tissues, such as ligaments, making it suitable for the direct assessment of syndesmotic injuries (Dhont et al., 2023). Injured ligaments often show changes in morphology, size, and signal intensity on MRI scans (Figure 1.9), providing valuable insights into the extent of the injury (Kellett et al., 2018b). A complete ligament tear on MRI scans can be interpreted by signs such as ligament discontinuity, a wavy appearance, irregular contours, or inadequate visualization of the ligament (Kellett et al., 2018b).

Figure 1.8 Magnetic Resonance Imaging (MRI) machine
Taken from Wikipedia (2025)

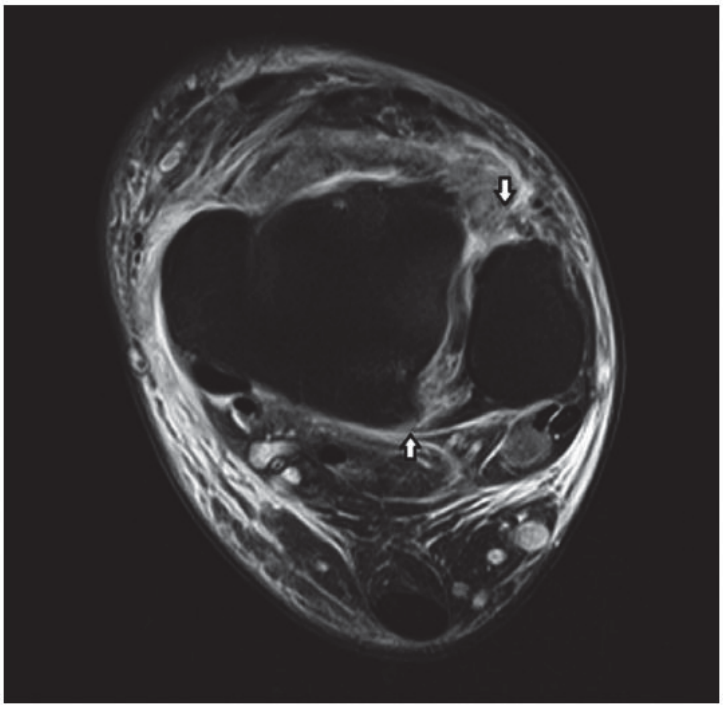

Figure 1.9 MRI at the syndesmotic level showing injured ligament
Taken from Krähenbühl et al. (2018).

Although MRI does not expose patients to radiation, it is not suitable for individuals with implanted pacemakers, drug infusion pumps, neurostimulators, or any other iron-based metal implants due to its use of a magnetic field. Additionally, like CT scans, patients must remain in a lying position during the procedure (Figure 1.8), which could potentially impact the assessment of the syndesmosis condition.

1.4.6 Comparing imaging techniques for assessment of syndesmotic injury

Comparing existing solutions for diagnosing syndesmotic injuries can provide valuable insights and help guide our conclusions. Table 1.2 provides a brief comparison of the most used techniques.

Table 1.2 Comparison of the imaging techniques for assessment of syndesmotic injuries

Technique	Radiation	Price	Standing position of the patient	Volume Imaging	Soft Tissue	Superimposition issue	Require an Expert to perform
X-ray	low	$	Yes	No	No	Yes	No
Stress X-ray	low	$	No	No	No	Yes	Yes
Ultrasound	None	$$	Yes	No	Yes	Yes	Yes
CT scan	Up to 200 Chest X-rays	$$$	No	Yes	No	No	No
MRI	None	$$$	No	Yes	Yes	No	No

In the following, we will discuss each factor and its importance in detecting syndesmosis injury.

- Radiation: X-ray imaging exposes the patient to radiation. While CT scan also involve radiation, it requires hundreds of images, resulting in exposure equivalent to up to 200 chest X-rays. MRI and ultrasound, on the other hand, do not present any radiation-related concerns.

- Price: X-rays are the simplest form among medical imaging techniques, making them more cost-effective compared to MRI and CT scans, which are generally more expensive.
- Standing position of the patient: Acquiring X-ray images while the patient is standing is crucial for detecting syndesmosis injuries. By positioning the ankle under weight-bearing conditions, the imaging better expresses real-life scenarios, allowing for a more accurate assessment of potential injuries. Similarly, dynamic ultrasound is conducted with the patient standing. Although stress X-ray imaging is typically performed while the patient is lying down, the external rotation and applied forces exert sufficient pressure on the ankle to facilitate the diagnosis of syndesmosis injuries. In contrast, MRI and conventional CT scans, as shown in Figure 1.6 and Figure 1.8, are typically taken while the patient is lying down, which causes the joints to be in a relaxed state. Many patients, however, experience discomfort when standing, as the joints are subjected to weight-bearing stress.
- Volume imaging: Volume imaging provides a more accurate representation than two-dimensional imaging, making injury detection easier. While CT scans and MRIs utilize volume imaging, X-ray based imaging techniques and ultrasound methods do not.
- Soft tissue: Soft tissues play a crucial role in identifying syndesmosis injuries. MRI and ultrasound effectively visualize soft tissues, whereas X-ray-based imaging techniques primarily capture bony structures. As a result, detecting injuries without considering the soft tissue requires alternative solutions, such as the ones outlined in sections 1.4.1 and 1.4.2.
- Superimposition issue: While volume imaging provides a comprehensive 3D view of the ankle joint, 2D imaging projects the structures onto a plane, leading to the superimposition of bones. This issue is illustrated in Figure 1.4, where the tibia and fibula are overlapped in a radiograph. Superimposition can obscure parts of the bones, making it difficult to accurately assess their alignment and relative positioning.

While current clinical practices depend on the imaging techniques mentioned earlier, each method has its own unique advantages and limitations. A promising alternative is 2D/3D reconstruction, which involves generating 3D model from one or more 2D radiographs. This

approach combines the benefits of low-radiation X-ray imaging with the assessment ability in 3D space.

1.5 Biplanar Radiographs by EOS Imaging system

The EOS™ X-ray machine (Figure 1.10) is capable of simultaneously capturing biplanar X-ray images through slot scanning of the entire body in an upright, physiological, load-bearing position, using ultra-low-radiation doses. It provides both AP and LT images that are spatially calibrated, offering more detailed information than a single radiograph (Illés & Somoskeöy, 2012).

Figure 1.10 EOS™ X-ray machine, capturing two AP and LT images simultaneously
Taken from Illés & Somoskeöy (2012)

The EOS Imaging system offers several advantages (Melhem, Assi, El Rachkidi, & Ghanem, 2016), including:

- Low-radiation dose: The system utilizes an ultrasensitive multi-wire proportional chamber, which requires only a minimal amount of X-ray exposure to generate clear images. This low-dose approach makes it a preferable option compared to CT imaging, particularly for conditions that require periodic imaging for ongoing analysis.
- Weight-bearing standing position: The EOS system captures images while the patient is standing, allowing joints, such as those in the spine and knee, to be imaged under normal weight-bearing loads. This contrasts with techniques like conventional CT scanning, where the patient must lie down, potentially altering the natural positioning and pressure on the joints.
- True-to-size imaging: The EOS system provides highly accurate representations of bone structures without distortion.

However, despite these advantages, there are certain limitations to the EOS system (Melhem et al., 2016):

- Occlusion and superimposition of bony structures: The projection of a single ray through the body can sometimes lead to occlusion, where certain parts of the bones are hidden. Additionally, when bones are superimposed, such as the femoral head and pelvis, the X-ray image may capture the overlapping bones as a single structure, making it difficult to distinguish the bones.
- Motion-related image distortion: If the patient is unable to stand still during the imaging process, the resulting images may exhibit distortions.
- Reduced contrast: The X-ray images generated by the EOS system tend to have less contrast and brightness compared to conventional digital radiography, which may affect the clarity of the images.
- Limited bone structure details: Since it relies on only two X-ray images, it represents less information about the structure of the bone than 3D imaging modalities, as it lacks the depth of information for reconstruction.

The EOS^{TM} software provides semi-automated 3D reconstruction for certain skeletal structures, including the spine, pelvis, and femur (Melhem et al., 2016), depicted in Figure

1.11. As a result, we may have the 3D model built from the radiographs, taking advantage of X-ray imaging while having access to 3D imaging results, with a comparable accuracy to CT scan (Melhem et al., 2016).

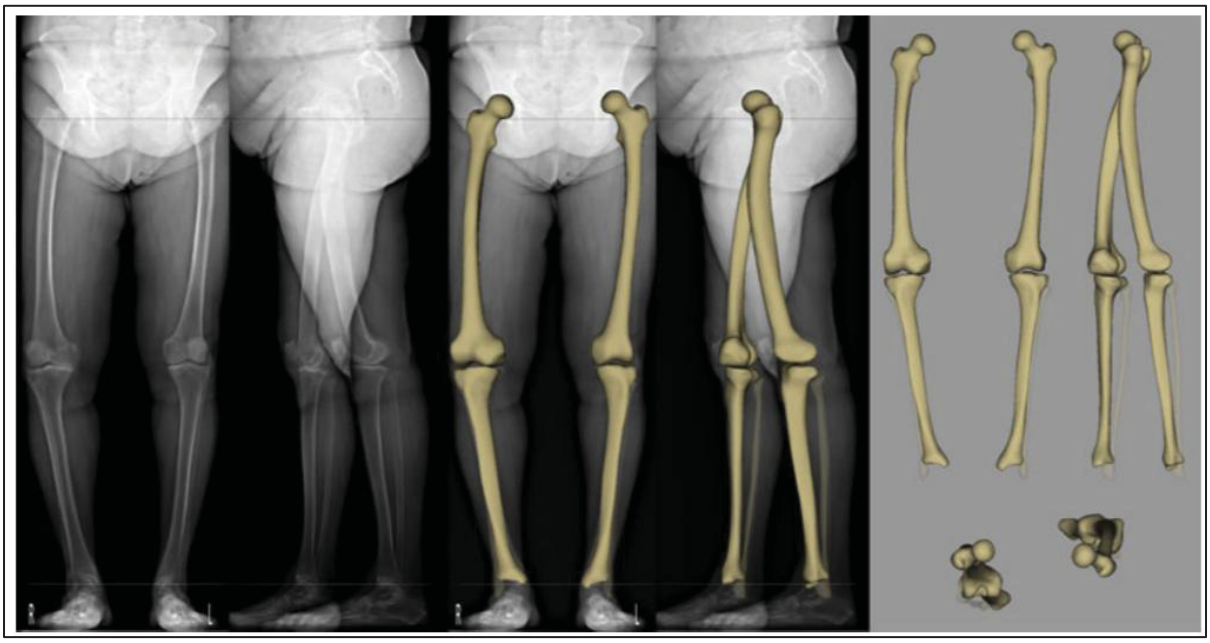

Figure 1.11 EOS™ 2D images and SterEOS™ 3D Reconstruction of the lower limb
Taken from Illés & Somoskeöy (2012)

1.5.1 3D Reconstruction of the Tibiofibular joint from biplanar radiographs

The reconstruction of the ankle using EOS images is not supported by SterEOS™. However, such reconstruction could still be valuable in evaluating syndesmosis injuries by analyzing the relevant clinical parameters in 3D as the 3D model can present the most definitive information (Ebinger, Goetz, Dolan, & Phisitkul, 2013). To measure these parameters, only a 3D reconstruction of the tibiofibular joint (Figure 1.12) is necessary, and the talus can be excluded from the model as explained in Section 1.4.2 and Figure 1.7. The process of creating a 3D personalized model of the Tibiofibular joint from 2D X-ray images presents significant challenges, since the structures are aligned closely in 3D space where distances can be as small

as 1 mm (Nault et al., 2013) and thus, there is considerable overlap of the structures, resulting from the projection process.

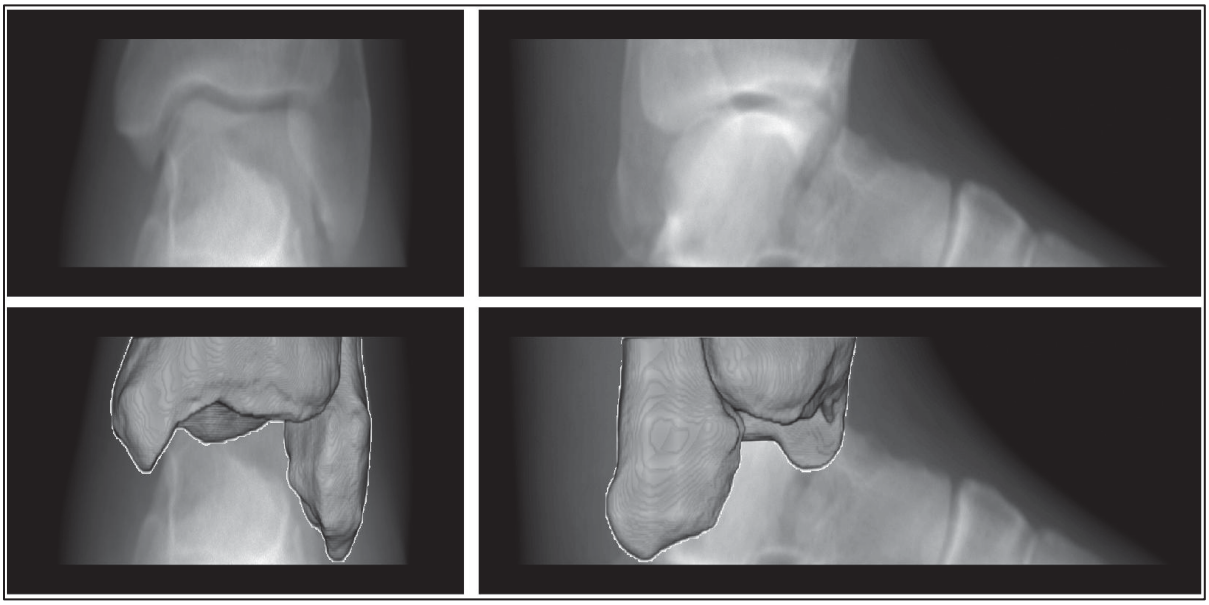

Figure 1.12 Ankle biplanar radiographs, having the tibiofibular joint highlighted

1.6 Conclusion

Accurate pre-surgery diagnosis is essential for effective syndesmotic injury treatment. While MRI provides insight into soft tissues, it is costly, it does not place the patient in a weight-bearing position, and not universally accessible. Among techniques that do not visualize soft tissues, CT scan offers 3D bone information but is conventionally taken in a non-weight-bearing position and exposes patients to significant radiation. Furthermore, current studies often compute these parameters on individual CT slices, limiting the analysis to two-dimensional measurements. The most critical factor in diagnosing syndesmotic injuries is imaging the patient while standing, putting the ankle under stress, and assessing the clinical parameters in 3D.

Currently, CT scans and MRIs are commonly used to diagnose syndesmosis injuries. While CT scans offer greater sensitivity in detecting small syndesmotic injuries compared to standard radiographs, MRI remains the gold standard for diagnosing ligamentous injuries due to its

ability to differentiate between sprains and tears. Finding an alternative imaging modality with similar diagnostic capabilities would be ideal.

EOS Imaging offers several advantages over CT and MRI, as highlighted earlier. Due to these benefits, we would opt for EOS Imaging over other methods. However, addressing the challenges posed by its limitations is necessary.

Although distances in the ankle can be calculated using CT scans, either in 2D on a slice or on the volume in 3D, the same approaches cannot be applied to EOS radiographs. The primary difference lies in the type of information provided by these two techniques. CT scans yield 3D information, while EOS Imaging provides two orthogonal 2D X-ray images. To detect injuries by analyzing the bones' relationship which is explained in Section 1.4.2, a 3D reconstruction of the tibiofibular joint from the EOS images is required. This approach would allow injury detection by assessing the clinical measurements in 3D. In this manner, we may leverage the advantages of EOS while overcoming its challenges, enabling us to identify potential syndesmotic injuries.

The challenges in 3D reconstruction of the Tibiofibular joint using EOS biplanar images are summarized as follows:

- Multi-object structure representation and the relationships
- Small inter-distances between bones
- Complex 3D structures of the bones which may not be representable by geometrical shapes such as spheres, cylinders, etc.
- Superimposition and visibility issues of structures due to projection process in generating radiographs (as we see superposed tibia, fibula and talus in Figure 1.12).

In the next chapter, we will conduct a literature review on various solutions for reconstructing 3D models from biplanar X-ray images, the information and methods used in this process, supporting techniques and clinical parameters estimation on the reconstructed models. We will explore the steps involved, address challenges, and the limitations in literature.

CHAPITRE 2

LITERATURE REVIEW

This chapter reviews the literature related to 3D reconstruction from radiographic or X-ray images, with a focus on techniques used to reconstruct anatomical structures and calculate clinical parameters from the resulting 3D models. The discussion encompasses the methodologies and the information employed in the reconstruction process, along with their advantages, limitations, and relevance to the 3D reconstruction of the tibiofibular joint from biplanar radiographs. In addition, we examine supporting methods that enhance the reconstruction pipeline, such as image translation techniques. Finally, the metrics commonly used to evaluate the accuracy and effectiveness of 3D reconstruction approaches are discussed.

2.1 Introduction

Plain X-ray images are generated by directing X-ray beams through a part of the body. While X-rays provide less detailed information than volume-based modalities such as CT or MRI, they are less expensive, faster to acquire, and expose patients to markedly lower radiation doses than CT scans (Goswami & Kr., 2015). Consequently, recent research increasingly leverages one or two calibrated, low-dose radiographs to reconstruct patient-specific 3D bone geometry as a cost-effective substitute for CT-based modelling (Melhem et al., 2016; Nguyen, Benameur, Mignotte, & Lavoie, 2023). However, in planar radiography, each detector pixel records the line integral of the tissue's linear attenuation coefficient along a single X-ray path, so every voxel intersected by that ray is collapsed into one value (Kak & Slaney, 2001). As this projection process superimposes anatomical layers that lie at different depths, the image contains only two-dimensional shadows with no intrinsic depth cues (Bushberg, Seibert, Leidholdt, & Boone, 2011). As a result, a single radiograph cannot unambiguously recover three-dimensional anatomy; reliable 3D reconstruction typically requires at least two calibrated views or strong anatomical priors (Kyung, Jo, Choo, Lee, & Choi, 2023).

The methods for 3D reconstruction from X-rays can be categorized based on the prior knowledge utilized during the reconstruction process and the methods used to incorporate this information into the process (Hosseinian & Arefi, 2015). Some approaches rely solely on the X-ray images as input, while others incorporate mathematical, geometrical, statistical, or parametric models containing information about the bones, which are then used by the 2D/3D reconstruction algorithms (Goswami & Kr., 2015). By utilizing these models, accurate 3D reconstructions can be achieved from the X-ray images (Hosseinian & Arefi, 2015). Methods that use such models require prior knowledge of the shape of interest, such as the structure of bones, relationship between bones in the joint, or other anatomical features.

2.2 3D Reconstruction from X-ray images

The process of reconstructing a 3D model from 2D X-ray images aims to generate a three-dimensional representation of an anatomical structure based on available X-ray projections. This is achieved by identifying the spatial correspondence between features in the 2D images and in 3D space (Hosseinian & Arefi, 2015). The technique used to align and map 2D image data to a 3D model is known as the 2D/3D reconstruction process (Unberath et al., 2021). Figure 2.1 illustrates examples of 3D reconstructions of the pelvis (Lamecker, Wenckebach, & Hege, 2006) and spine obtained using one and two X-ray images (Zhang et al., 2013), respectively.

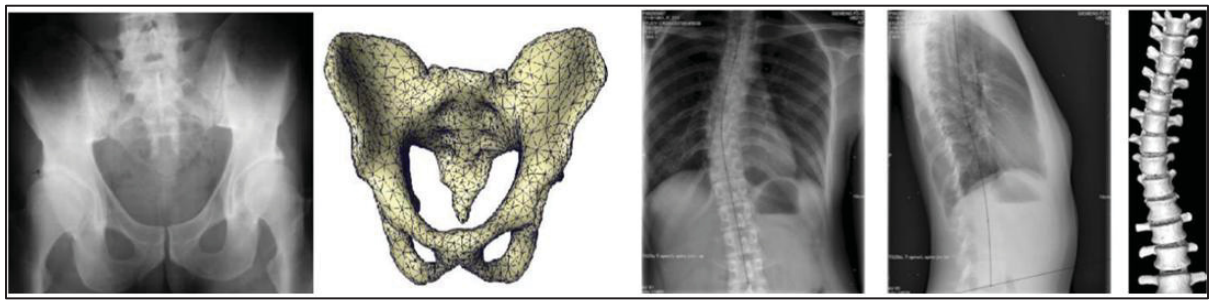

Figure 2.1 3D reconstruction of pelvis from one X-ray image and Spine from two biplanar X-ray images
Taken from Lamecker et al. (2006) and J. Zhang et al. (2013)

2D/3D reconstruction techniques can be categorized according to several criteria, including the applications, the prior knowledge incorporated into the process through models, the type of data used for aligning models with images, the specific algorithms employed to match 3D structures onto 2D views, and the level of automation (Hosseinian & Arefi, 2015; Unberath et al., 2021). The following sections provide a detailed discussion of each of these aspects.

2.2.1 Applications

3D reconstruction from radiographs has proven effective across several body regions, such as spine with an average reconstruction error of 3 to 5 mm (B Aubert, Cresson, Guise, & Vazquez, 2022; Hanaoka et al., 2011; Kadoury, Cheriet, Kadoury, Cheriet, & Labelle, 2009; Moura, Boisvert, Barbosa, Labelle, & Tavares, 2011; Moura, Boisvert, Barbosa, & Tavares, 2009; Pomero, Mitton, Laporte, De Guise, & Skalli, 2004a), pelvis, hip, and femur with an average error of 3 to 4 mm (Bah et al., 2015; Baka et al., 2011, 2010; Dong, Ballester, & Zheng, 2007; Dong & Zheng, 2009; Fleute & Lavallée, 1999; Lamecker et al., 2006; Mahfouz, Badawi, Fatah, Kuhn, & Merkl, 2006; Nolte, Xie, & Bull, 2023; Zheng et al., 2009), knee with average error of 2 to 3 mm (Hwang, Lee, & Shin, 2024; Kasten et al., 2020; Tchinde Fotsin, Vazquez, Cresson, & De Guise, 2019).

To the best of our knowledge, the problem of reconstructing an accurate 3D model of the ankle, specifically the distal tibiofibular joint, from 2D X-ray images has not been thoroughly investigated in the literature. The most closely related works to ours have performed 3D reconstruction from biplanar X-rays of the proximal tibia (Arn Roth et al., 2024), and tibia and fibula (Pan et al., 2023) without focusing on the distal region. Their application is centered on knee arthroplasty and preoperative osteotomy, which primarily involves the proximal sections of these bones. In contrast, our study is among the first to specifically address the technical and anatomical challenges associated with 3D reconstruction of the distal tibiofibular joint. One key challenge lies in the narrow spacing between the tibia and fibula in the distal region, which has been reported to range from approximately 1 mm to 16 mm (Nault et al., 2013).

Therefore, any reconstruction methodology must preserve this anatomical constraint to ensure the structural accuracy of the resulting model.

2.2.2 Shape and Appearance Models Used in 3D Reconstruction

As previously discussed, incorporating 3D models into the reconstruction process results in accurate 3D reconstructions from the X-ray images (Hosseinian & Arefi, 2015). Models are commonly employed to address the inherent limitations of radiographic imaging, particularly the loss of depth information resulting from 2D projection (Heimann & Meinzer, 2009). Given that a single-view or even a biplanar radiograph does not provide enough information to reconstruct a unique and complete 3D anatomical structure, such models act as prior constraints. They reduce the ambiguity in the reconstruction process by narrowing the solution space and guiding it toward anatomically realistic shapes (Heimann & Meinzer, 2009).
3D models used for reconstruction can take various forms, including generic template models, Statistical Shape Models (SSMs), models that incorporate both shape and intensity information such as Statistical Shape and Intensity Models (SSIMs), and parametric models. When dealing with joint structures, specialized models like coupled or articulated models are often required to account for the relative positioning or movement between connected bones. The following sections provide an overview of each of these model types.

2.2.2.1 Generic Models

Generic models, also referred to as anatomical templates, provide a mathematical representation of an object's geometric surface in 3D. These models are typically aligned and morphed to fit patient-specific radiographs.
Generic models are defined as a set of n vertices $v_i \in R^3, i = 1..n$, forming a point cloud which represents the surface of the shape. Alternatively, a triangular mesh-based approach can be employed, which not only specifies vertex positions but also includes connectivity information by defining m triangular faces, where each face consists of three connected

vertices forming a triangle. Equations (2.1) and (2.2) are the definition of the vertices and faces, respectively.

$$V = [v_1, v_2, \dots, v_n] = [x_1, y_1, z_1; x_2, y_2, z_2; \dots; x_n, y_n, z_n] \quad (2.1)$$

$$F = [f_1, f_2, \dots, f_m] = [v_1, v_2, v_3; v_2, v_4, v_5; \dots] \quad (2.2)$$

Figure 2.2 shows a triangular mesh representation of a distal femur generic model. The vertices are then used to deform by an optimization algorithm such as kriging and the use of control points to align the model to 2D X-ray images (Laporte, Skalli, de Guise, Lavaste, & Mitton, 2003).

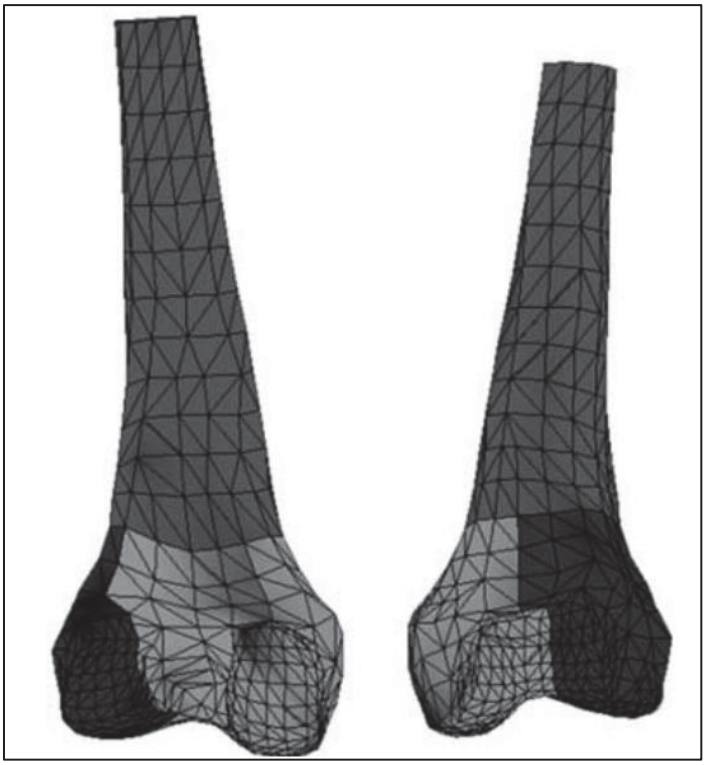

Figure 2.2 distal femur mesh representation
Taken from Laporte et al. (2003)

Generic models are usually constructed based on a single CT scan instance (Chaibi et al., 2012). The process begins with segmentation to isolate the target bone (Heimann & Meinzer,

2009), followed by a surface reconstruction step using meshing techniques. The final output is typically a 3D representation of the bone in the form of a point cloud or a triangular mesh.

These models facilitate the reconstruction process by providing a predefined structure that can be efficiently aligned or deformed to match patient-specific radiographic views (Mitton et al., 2006). This reduces the dimensionality of the problem, provides anatomical constraints, and helps overcome ambiguities inherent to limited 2D projections. They are applied to distal femur (Laporte et al., 2003), proximal femur (Galibarov, Prendergast, & Lennon, 2010), ribcage (Koehler, Wischgoll, & Golshani, 2019), complete femur (Arnold, Blemker, & Delp, 2001) and tibia and fibula (Pan et al., 2023). Some generic models are generated for foot and ankle (Ramlee, Kadir, & Harun, 2013), but they are not incorporated in the 3D reconstruction process from radiographs.

As mentioned, segmentation is an essential step in generating a model from scratch. CT scan segmentation has progressed from classical image processing techniques to powerful deep learning-based solutions. Initial methods relied on thresholding, edge detection, and region-growing algorithms, which worked well for high-contrast structures but struggled with complex or noisy data (Pham, Xu, & Prince, 2000). Machine learning methods, including random forests and support vector machines, introduced supervised feature-based segmentation with improved robustness. The advent of deep learning revolutionized the field, with convolutional neural networks (CNNs) such as U-Net becoming the standard for automatic segmentation tasks due to their ability to learn hierarchical features directly from images (Ronneberger, Fischer, & Brox, 2015). More recent advancements include 3D CNNs for volumetric segmentation and attention mechanisms to improve focus on relevant anatomical regions (Çiçek, Abdulkadir, Lienkamp, Brox, & Ronneberger, 2016; Rayed et al., 2024). These methods significantly outperform earlier techniques, though generalization across scanners and pathologies remains an ongoing challenge. As the reconstruction process accuracy is heavily dependent on the segmentation result, it is crucial to have an accurate result from segmentation process.

Once segmentation masks are generated from CT scans, the next step is meshing process which is converting voxel-based binary masks into continuous surface representations. One of the most widely used techniques for this is the Marching Cubes algorithm (Lorensen & Cline,

1987), which constructs high-resolution triangular meshes by evaluating the local configuration of voxel values and interpolating the surface within each cube. This method is efficient and produces topologically accurate surfaces, making it a foundational tool in medical imaging and surgical planning. In addition to Marching Cubes, more advanced meshing techniques include Poisson surface reconstruction and level-set-based methods, which offer smoother and more watertight surfaces, particularly useful for complex or noisy segmentations (Kazhdan, Bolitho, & Hoppe, 2006). Recently, neural implicit surface representations such as DeepSDF have been explored for learning smooth surface models directly from data(Park, Florence, Straub, Newcombe, & Lovegrove, 2019). However, classical meshing remains dominant in clinical pipelines due to its reliability, speed, and broad implementation in medical software.

In the context of ankle modeling, a generic model of ankle and foot has been built, but they failed to capture the local details of the structures in their approach (Ramlee et al., 2013). Their model encompasses all bones from the knee to the foot.

Generating the generic models is straightforward and easy (Heimann & Meinzer, 2009). However, because these models are based on a single shape, they cannot effectively capture shape variations (Timothy F. Cootes, 1977) and as a result, they have limited generalization performance (Luthi, Gerig, Jud, & Vetter, 2018). In other word, the reliance on a single reference shape in generic models results in poor adaptability to patient-specific anatomy, leading to increased reconstruction error and limited ability to resolve ambiguities inherent in radiographic projections (Ambellan, Lamecker, von Tycowicz, & Zachow, 2019; Luthi et al., 2018).

2.2.2.2 Statistical Shape Models

To address the limitations of shape representation inherent in generic models, Statistical Shape Models (SSMs) were introduced (T. F. Cootes, Taylor, Cooper, & Graham, 1995). These models are constructed from a set of N instances, where each shape is represented by a set of corresponding points distributed over the surface or by triangular meshes (Heimann & Meinzer, 2009). The object is then characterized by its mean shape and the statistical variations

observed across the dataset. This multi-instance framework enables the model to capture anatomical variability within a population, providing a more generalizable and data-driven representation of shape. By applying statistical analysis, typically using Principal Component Analysis (PCA) and on the aligned dataset, the model encodes both the average shape and the main modes of deformation (T. F. Cootes et al., 1995). An example is shown in Figure 2.3, where aligned pelvic shapes are overlaid after removing translation, rotation, and scale.

To construct an SSM, a dataset of N aligned 3D meshes is used, where each shape instance is brought into correspondence and statistically analyzed to extract a mean shape and principal deformation modes (Ambellan et al., 2019; Heimann & Meinzer, 2009). One of the most challenging aspects of constructing an SSM is achieving accurate point-to-point correspondences across all shapes in the training dataset. Errors in this step can significantly degrade the quality of the resulting model and its ability to represent anatomical variability (Ambellan et al., 2019; Davies, Cootes, & Taylor, 2001; Heimann & Meinzer, 2009). Matching techniques are based on surface deformation techniques (Choi, Lam, & Lui, 2015; Vikas & Bhallamudi, 2014), such as mapping shapes onto spherical models and match them in the latter space (Brechbuhler, Gerig, & Kubler, 1995; Hashemibakhtiar, Cresson, Guise, & Vázquez, 2023; Kirschner & Wesarg, 2010; Lee & Kazhdan, 2019; M. Styner et al., 2006), using multiscale matching (Eisenberger, Lähner, & Cremers, 2019), deform shapes through non rigid registration (A. F. Frangi, Rueckert, Schnabel, & Niessen, 2003; Alejandro F. Frangi, Rueckert, Schnabel, & Niessen, 2001; Heitz, Rohlfing, & Maurer, Jr., 2005; Rueckert, Frangi, & Schnabel, 2001; J. Wang & Shi, 2017; Wu, Li, Lu, Kim, & Ogunbona, 2017), and establishing and optimizing correspondence through achieving the best compactness, generalization, and specificity of the resulting SSM (J. H. Chen & Shapiro, 2009; J. H. Chen, Zheng, & Shapiro, 2010; Davies et al., 2001; Davies, Twining, Cootes, & Taylor, 2010; Heimann, Wolf, Williams, & Meinzer, 2005). Deep learning approaches are also incorporated for finding the correspondences through the dataset (Agrawal, Whitaker, & Elhabian, 2017; Groueix, Fisher, Kim, Russell, & Aubry, 2018).

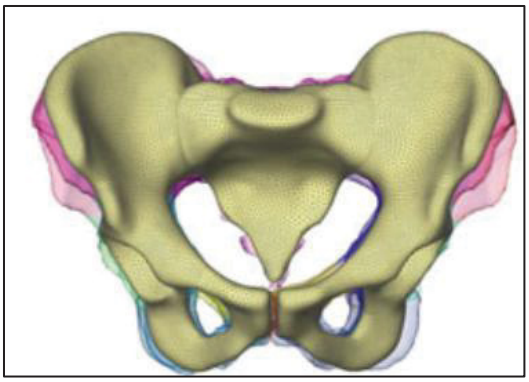

Figure 2.3 Overlay of several pelvis shapes which is used for constructing a SSM
Taken from Lamecker & Zachow (2016)

A general view of generating an SSM is depicted in Figure 2.4. It begins with segmenting the bone of interest from a series of CT scans (Skadłubowicz, Król, Wróbel, Hefti, & Krieg, 2009). A meshing technique is then applied to the segmented masks denoted as $CT_1^{seg}, CT_2^{seg}, \dots, CT_N^{seg}$ to create surface meshes $S_1, S_2, \dots, S_N$ (Hashemibakhtiar et al., 2024). These steps can be skipped if a dataset of N 3D meshes is accessible.

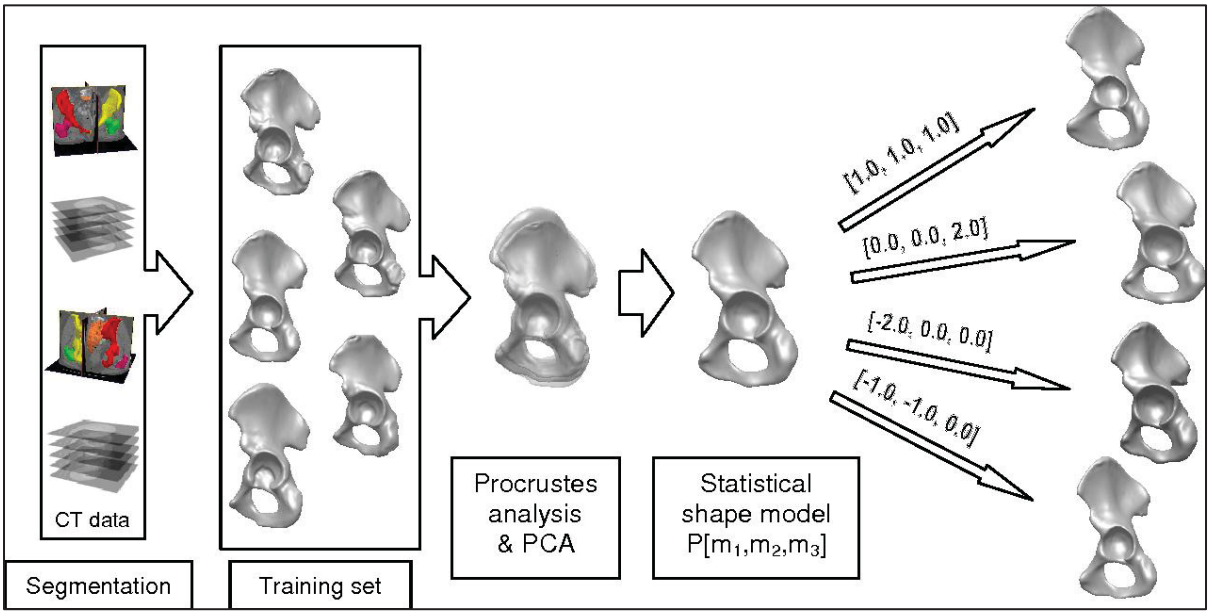

Figure 2.4 Steps to build a Statistical Shape Model
Taken from Skadłubowicz, Król, Wróbel, Hefti, & Krieg (2009)

As it is essential to have all meshes represented with the same topology (Heimann & Meinzer, 2009), a matching process, either automatic, semi-automatic, or manual, is employed to ensure a one-to-one anatomical correspondence across the dataset (Davies et al., 2001), resulting in $S_1^{st}, S_2^{st}, \dots, S_N^{st}$ which all represent same number of vertices and faces. Once aligned, the meshes undergo rigid transformation alignment using methods such as Generalized Procrustes Analysis (GPA) (Gower, 1975), which efficiently removes variations in rotation, translation, and scale in an iterative process. This normalization step is essential to ensure that the computed modes of variation reflect genuine anatomical differences rather than pose differences (Heimann & Meinzer, 2009). The transformation to remove scale, rotation, and translation is given by Equation (2.3) where R is the rotational matrix, s is the scaling factor t is the translation parameter. GPA iteratively aligns shapes and updates the mean configuration until convergence (Gower, 1975). This alignment step prepares the dataset, $S_1^{rm}, S_2^{rm}, \dots, S_N^{rm}$ for statistical analysis, commonly performed using Principal Component Analysis (PCA) (PCA) (Jollife & Cadima, 2016).

$$S_i^{rm} = sRS_i^{st} + t\ , i = 1..N \tag{2.3}$$

Having all the shapes aligned, we can compute the mean shape $\bar{S}$ using Equation (2.4).

$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i^{rm} \tag{2.4}$$

Given the matrix of the training shapes $Y = [S_1^{rm}, S_2^{rm}, \dots, S_N^{rm}] \in R^{3K\times N}$ where each shape consists of K vertices, the covariance matrix C is computed using Equation (2.5)

$$C = \frac{1}{N-1} YY^T \in R^{3K\times 3K} \tag{2.5}$$

The eigenvectors $\Phi = [\psi_1; \psi_2; \dots; \psi_{N-1}]$ and eigenvalues $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{N-1}]$ are determined using Singular Value Decomposition (SVD), as explained in Equation (2.6).

$$C\Phi = \Lambda\Phi \tag{2.6}$$

The principal Modes of Variation (MoV) are selected by analyzing the eigenvalues of the covariance matrix and computing the cumulative explained variance (T. F. Cootes et al., 1995). A subset of modes is typically retained to capture a predefined proportion of the total shape variability, as illustrated in Equation (2.7). The number of retained principal components M is typically selected such that they capture between 90% and 99% of the total variance in the training shapes, depending on the application-specific, balance between model compactness and expressiveness (T. F. Cootes et al., 1995; Heimann & Meinzer, 2009; Luthi et al., 2018).

$$Var(M) = \frac{\sum_{j=1}^{M} \lambda_j}{\sum_{k=1}^{N-1} \lambda_k} \tag{2.7}$$

Finally, having the mean shape $\bar{S}$ and the eigenvectors matrix Φ, the SSM is defined using Equation (2.8) where $b_k \in R^M$ follows a normal distribution bounded by $\pm 3\sqrt{\lambda_k}$ (T. F. Cootes et al., 1995).

$$S(b_1, b_2, \dots, b_{N-1}) = \bar{S} + \sum_{k=1}^{N-1} b_k . \psi_k \tag{2.8}$$

This process yields a compact yet informative representation of the anatomical structure, capturing its natural variations.

In the context of 3D reconstruction using SSM, the goal is to estimate the optimal shape parameters b such that the projection of the generated 3D shape best aligns with the observed 2D contours or silhouettes in the radiographs (Shiode et al., 2021). This parametric constraint drastically reduces the ill-pose problem of the 2D to 3D matching by limiting the solution space to statistically valid shapes.

Establishing accurate correspondence between 3D meshes is a critical preprocessing step for SSM (Sahillioğlu, 2020), after segmentation and meshing process. shape reconstruction, and

analysis. To build an SSM, correspondences must be defined across all training shapes(T. F. Cootes et al., 1995), and poor matching, such as aligning anatomically dissimilar regions, can introduce significant modeling errors and overestimated variability (Davies, Twining, Cootes, Waterton, & Taylor, 2002). The challenge intensifies with 3D, non-rigid, and non-isometric deformations, where both local and global shape features must be considered (Sahillioğlu, 2020). Two main solution categories are commonly identified (Hashemibakhtiar et al., 2023). Similarity-based methods rely on matching geometric descriptors invariant to transformations (Gehre, Bronstein, Kobbelt, & Solomon, 2018; Ling, Qinsong, Shengjun, & Xinru, 2021; Melzi, Ovsjanikov, Roffo, Cristani, & Castellani, 2018; Nogneng & Ovsjanikov, 2017; Ovsjanikov, Ben-Chen, Solomon, Butscher, & Guibas, 2012; Ren, Poulenard, Wonka, & Ovsjanikov, 2018; Vestner, Litman, Rodolà, Bronstein, & Cremers, 2017). Registration-based approaches, in contrast, deform one shape to another via a continuous transformation or through a shared parameter space (Cosmo et al., 2019; Dyke, Lai, Rosin, & Tam, 2019; Eisenberger et al., 2019; Lee & Kazhdan, 2019; Melzi et al., 2019). It has been elaborated that the quality of the shapes correspondences is also affected by the non-isometric deformation degree of the shapes (Hashemibakhtiar et al., 2023; X. Huang, Yang, Vouga, & Huang, 2020), which is mostly low in bony structures cases. Smooth Shell has represented the least matching error for shapes with low degrees of non-isometric deformation (Hashemibakhtiar et al., 2023; X. Huang et al., 2020).

Statistical Shape Models (SSMs) offer substantial advantages over generic models in the context of anatomical 3D reconstruction. Unlike generic models, which are typically constructed from a single anatomical instance and therefore fail to capture inter-subject variability, SSMs are built from a population of aligned shapes. This allows them to statistically represent both the mean anatomical structure and its plausible deformations (T. F. Cootes et al., 1995). As a result, SSMs are better equipped to generalize across different patients and accommodate natural anatomical variability, making them particularly effective in scenarios where imaging data, such as biplanar radiographs, lack depth information. By constraining the reconstruction process to a learned, low-dimensional shape space, SSMs reduce ambiguity and improve anatomical plausibility, thereby enabling more accurate and robust 3D reconstructions (T. F. Cootes et al., 1995; Heimann & Meinzer, 2009).

In addition to improving accuracy, SSM-based reconstruction methods facilitate greater automation by incorporating prior knowledge about the shape of interest, thereby reducing the need for extensive user intervention (Davies et al., 2001). However, the construction of an effective SSM requires a sufficiently large and diverse dataset, including both normal and pathological variations of the anatomical structure (Hosseinian & Arefi, 2015). The quality of the resulting model is highly dependent on the accuracy of shape alignment and the establishment of reliable point correspondences across the dataset (Heimann & Meinzer, 2009). The generalizability of the model increases with the quantity and diversity of training samples, while an optimal SSM is characterized by three critical properties: compactness, generalization, and specificity (Davies et al., 2001). Due to its ability to encapsulate statistical knowledge about anatomical shape and variability, an SSM provides a powerful prior that can be directly integrated into the 3D reconstruction pipeline to guide and regularize the estimation process (Ambellan et al., 2019).

2.2.2.3 Intensity Models

Statistical Shape and Intensity Models (SSIMs) extend the capabilities of Statistical Shape Models (SSMs) by simultaneously modeling both geometric variability and the associated intensity patterns of anatomical structures (C. J. F. Reyneke et al., 2019). This combined representation allows for a more complete characterization of anatomical regions, incorporating not only the shape but also density characteristics captured by imaging modalities such as CT (C. J. F. Reyneke et al., 2019). As a result, SSIMs have found increasing utility in applications such as 3D reconstruction, segmentation, image registration, and data augmentation (Ambellan et al., 2019; C. Reyneke, Thusini, Douglas, Vetter, & Mutsvangwa, 2018). Figure 2.5 shows a SSIM of pelvis, representing the bone’s density distribution as intensity information.

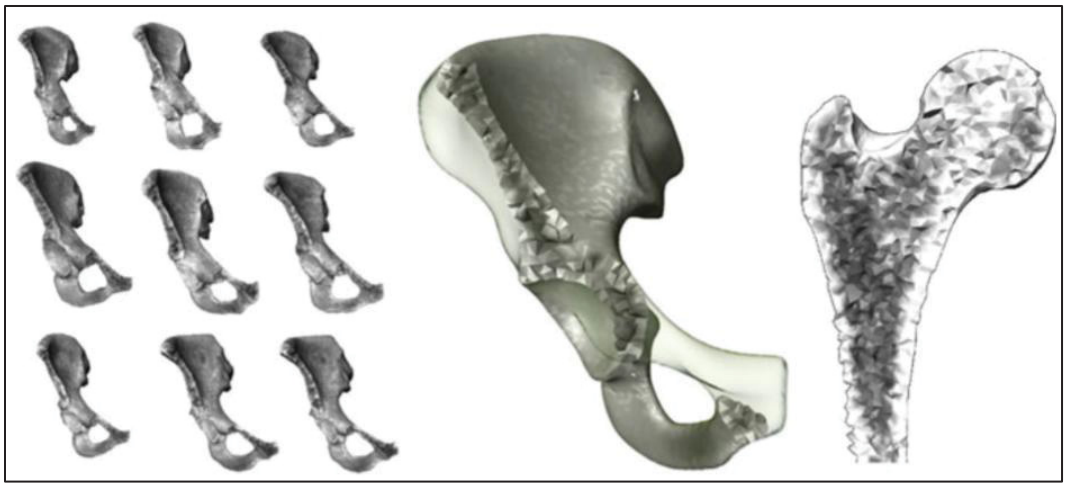

Figure 2.5 Statistical Shape and Intensity Model of pelvis
Taken from Lamecker & Zachow (2016)

SSIM construction can follow multiple strategies depending on how intensity values are associated with the anatomical representation (C. Reyneke et al., 2018). These strategies can be broadly categorized as follows:

- Voxel-Based Modeling: In voxel-based SSIMs, all training images are registered to a common anatomical frame. Intensity values are then stored voxel-wise, allowing the use of dimensionality reduction techniques such as Principal Component Analysis (PCA) to jointly model shape and image appearance (Figure 2.6. A). This approach is suitable for modeling detailed internal structures but can be memory-intensive and sensitive to registration errors (Hanaoka et al., 2011; C. Reyneke et al., 2018). Averaging the voxel intensities to simplify the information is applied to tackle the intensive processing issue. The bony structures could be split into cortical and trabecular bones (Figure 2.6.C) in which each area has the average constant voxel value of that region (Julien, Fotsin, Présentée, & Technologie, 2021; Sarkalkan, Weinans, & Zadpoor, 2014).
- Mesh-based Modeling: This method maps intensity values to the nodes of a triangular or tetrahedral mesh (Figure 2.6. B) constructed from the segmented anatomical surface or volume (Bah et al., 2015; C. Reyneke et al., 2018; Tchinde Fotsin et al., 2019; Zheng, Guoyan; Yu, 2017). Once all meshes are brought into correspondence, statistical

analysis is applied to both shape (vertex coordinates) and intensity (values assigned to the mesh nodes). This technique provides a more compact representation and is particularly effective for modeling structures like bones with clear internal density gradients (Aldieri, Terzini, Audenino, Bignardi, & Morbiducci, 2020).

- Particle-Based Modeling: In particle-based SSIMs, intensity values are assigned to particles distributed throughout the object volume (Oh & Koo, 2024; Xu & Elhabian, 2024). Each particle encodes both its spatial location and the corresponding local intensity, enabling the modeling of fine-grained internal variations. Such models are particularly useful for tasks involving image synthesis and non-rigid registration.

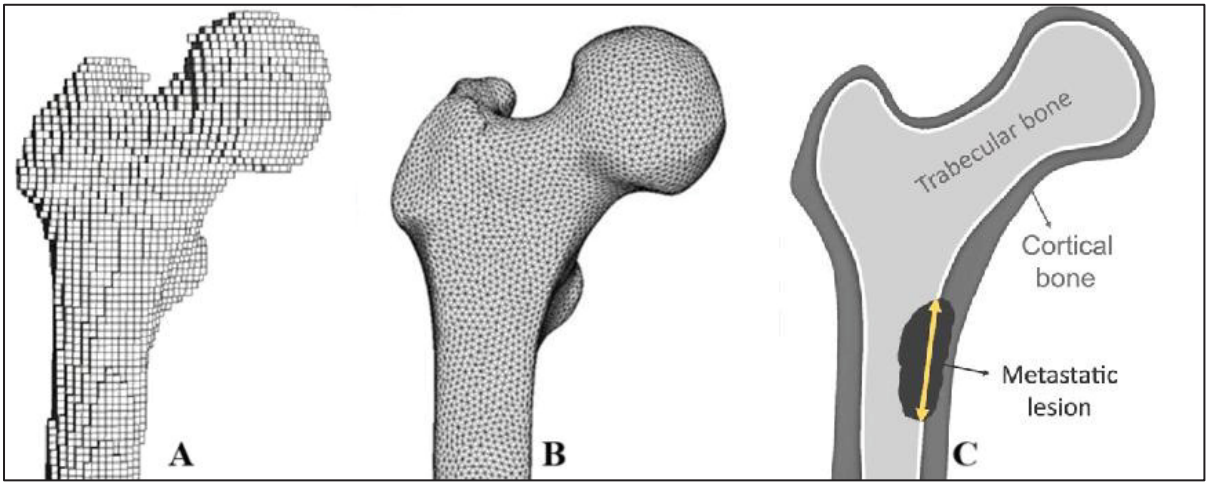

Figure 2.6 voxel-representation (A) and tetrahedral mesh representation (B) of the proximal femur. Splitting the trabecular and cortical bones of the proximal femur (C)
Taken from Sas, Tanck, Sermon, & van Lenthe (2020)

When extending a SSM to incorporate image intensity, two primary modeling strategies can be employed: independent and combined construction (C. J. F. Reyneke et al., 2019). These approaches differ in how the shape and intensity components are statistically coupled. Before integrating shape and intensity information, each of the N training shapes must represent their intensity data as a vector of P corresponding samples (e.g., tetrahedra, voxels, etc.), as shown in Equation (2.9).

$$I_j = \left(i_1^j, i_2^j, \dots, i_P^j\right), j = 1..N \tag{2.9}$$

In the independent modeling strategy, shape and intensity are treated as separate random variables, and PCA is applied independently to each. This approach enables more modular models in which shape and appearance variations can be analyzed separately, resulting in more accurate, patient-specific reconstruction (C. J. F. Reyneke et al., 2019). However, this is generally more time-consuming to implement rather than the combined approach. (C. J. F. Reyneke et al., 2019). A similar process to the SSM (as explained in Section 2.2.2.2) is applied on the intensity vectors and finally the independent SSIM can be presented with Equation (2.10) where I is the average vectors of intensities, Q represents the number of principal components of the intensity model I, while u_q and e_q denote the qth eigenvalue and eigenvector, respectively.

$$I = \bar{I} + \sum_{q=1}^{Q} w_q \sqrt{u_q} e_q \qquad w_q \sim N(0,1) \tag{2.10}$$

Although the independent modeling approach represents more interpretable intensity models, it may fail to capture the underlying correlations between anatomical shape and intensity variations. In contrast, the combined modeling approach concatenates the shape and intensity vectors from each training sample into a single feature vector, upon which a joint PCA is performed (C. J. F. Reyneke et al., 2019). This formulation captures shared variance between geometry and intensity and is particularly beneficial in tasks such as 2D/3D reconstruction, where both components are jointly informative. Combined SSIMs are generally more efficient to implement, as they involve fewer parameters and require less memory compared to independent models. This streamlined structure also reduces the computational time needed to estimate patient-specific parameters during reconstruction or analysis(C. J. F. Reyneke et al., 2019). PCA is applied on the Shape vectors and Intensity vectors separately, and after they are combined, another PCA applies on it to have the combined SSIM model (C. J. F. Reyneke et al., 2019). Equation (2.11) shows the combining formula, in which E_m^T and E_q^T are transposed versions of eigenvectors matrices of shape and intensity, respectively. W_m reconciles the shape

and intensity model parameters, as they are different in magnitudes (C. J. F. Reyneke et al., 2019).

$$w_{combined} = \begin{pmatrix} W_m w_m \\ w_q \end{pmatrix} = \begin{pmatrix} W_m E_m^T (S - \bar{S}) \\ E_q^T (I - \bar{I}) \end{pmatrix} \tag{2.11}$$

Choosing between voxel-based, mesh-based, and particle-based approaches for constructing SSIMs depends on the target application, the nature of the imaging data, and the required balance between anatomical detail and computational efficiency. Voxel-based SSIMs represent both shape and intensity using aligned 3D grids, making them well-suited for applications requiring high-resolution modeling of internal structure characteristics, such as bone density from CT scans (C. J. F. Reyneke et al., 2019). These models are advantageous when realistic appearance modeling is essential, such as intensity-based segmentation. However, they come at the cost of increased computational complexity and require precise voxel-level registration across training samples unless their simplified approaches are considered. In contrast, mesh-based SSIMs map intensity values to the nodes of surface or volumetric meshes, offering a more compact representation of anatomical structures. This method is particularly useful in statistical analysis of organ morphology, where region-specific average intensity is acceptable (Bah et al., 2015). While mesh-based models reduce data dimensionality, they may not fully capture fine-grained internal intensity gradients. Particle-based approaches provide a flexible alternative by automatically distributing corresponding particles across anatomical surfaces or volumes, each potentially carrying both positional and local intensity information (Schmid, Assassi, & Chênes, 2023). Although particle-based SSIMs are computationally lighter and more adaptable to shape variation, their resolution is generally lower, and modeling intensity reliably may require sampling strategies such as profile extraction or patch-based encoding. Ultimately, the decision among these approaches involves a trade-off between anatomical realism, computational cost, and the specific requirements of the reconstruction or analysis task.

In addition, the choice between these strategies depends on the target application: independent models may be more interpretable, whereas combined models offer richer representational

capacity by encoding the co-variation of anatomical shape and intensity (C. J. F. Reyneke et al., 2019).

Utilizing SSIM into 3D reconstruction process from radiographs help generation of realistic Digitally Reconstructed Radiographs (DRRs). DRRs are synthetic 2D projections generated from a 3D anatomical model to simulate X-ray images (Oh & Koo, 2024; C. J. F. Reyneke et al., 2019). In SSIM-based pipelines, both the shape and intensity of the model contribute to DRR generation, facilitating the creation of DRRs that more closely resemble real radiographs. This capability is especially important for 2D/3D reconstruction, where these DRRs are used to match and optimize alignment with actual patient X-rays. The result is a more robust and anatomically plausible estimation of the underlying 3D shape (C. J. F. Reyneke et al., 2019). 'Voxel Projector' is a simulation application developed by EOS Imaging ™, which simulates cabin projections (AP and LT images at the same time) from a volume, e.g., a CT-scan (EOS imaging, n.d.). In addition, SSIMs enable the creation of anatomically plausible synthetic datasets for training and validating machine learning models in low-data scenarios (Schmid et al., 2023).

As with the process of building an SSM, establishing correspondences for the SSIM is the most challenging part. Moreover, the accuracy of the segmentation process becomes even more critical here, as errors can significantly impact the collection of intensity values. Additionally, while incorporating more information improves the model, handling the larger amount of data can be time intensive. However, this trade-off results in the creation of more accurate models, especially for generating DRRs.

2.2.2.4 Multi-object and Articulated Models

Modeling anatomical structures comprising multiple interconnected components, such as joints, necessitates approaches that capture both individual object variations and their interrelationships. Multi-object and articulated models have been developed to address this complexity by jointly modeling the shapes and spatial configurations of multiple anatomical entities (Balestra, Schumann, Heverhagen, Nolte, & Zheng, 2014; Tchinde Fotsin et al., 2019).

These models may be based on SSM, SSIM or parametric models. Figure 2.7 shows a coupled model of multiple bones in foot.

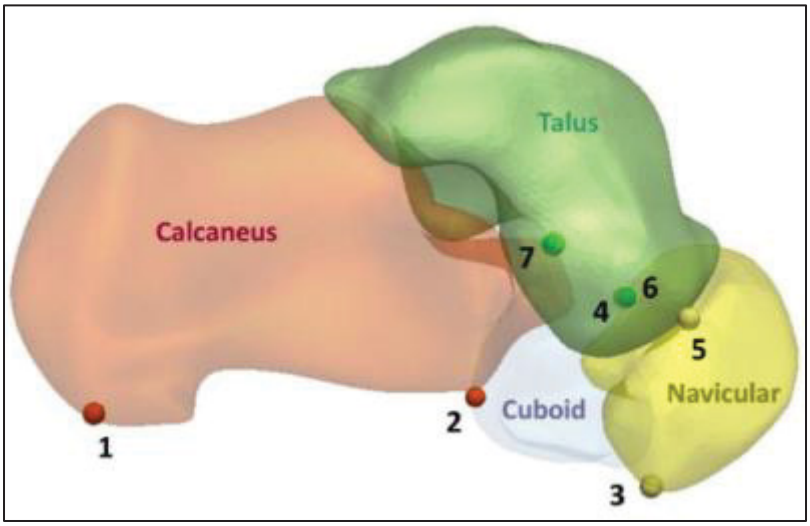

Figure 2.7 Coupled Shape Model of the foot
Taken from Brehler et al. (2019)

One foundational approach involves constructing articulated SSMs that incorporate explicit joint constraints to model the relative motions between connected bones such as developing an articulated SSM for 2D-3D reconstruction of the hip joint from biplanar radiographs, integrating rotational parameters to model the ball-and-socket joint behavior (Balestra et al., 2014) . Similarly, knee joint is presented by an articulated SSM, capturing both statistical shape variations and physiological joint movements (Bindernagel et al., 2011; Tchinde Fotsin et al., 2019). Spine modeling has been performed by modeling each vertebra and connecting them through an arithmetic spine model, with a registration process applied to adjacent vertebrae (Boisvert, Cheriet, Pennec, Labelle, & Ayache, 2008; Seitel, Rasoulian, Rohling, & Abolmaesumi, 2015). Figure 2.8 depicts the knee joint definition using articulated modeling (Bindernagel et al., 2011).

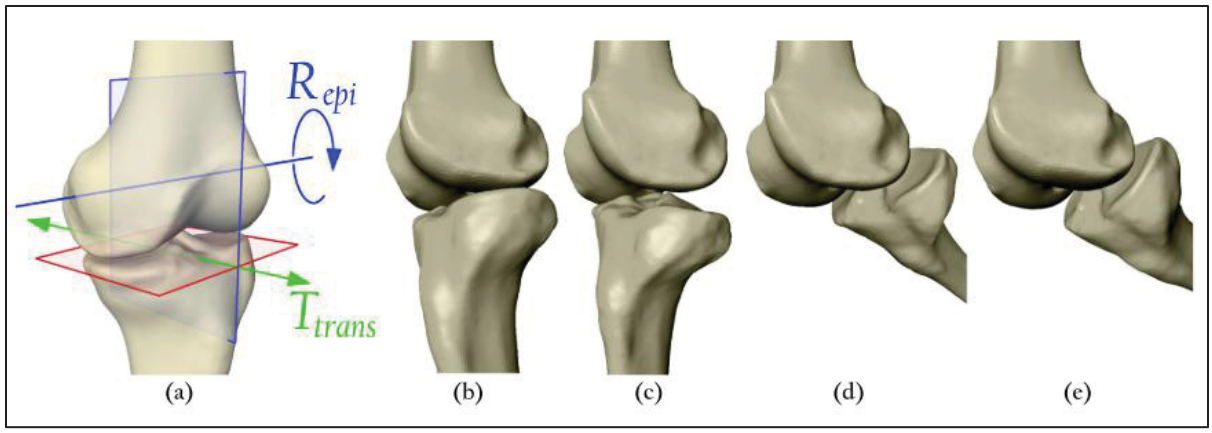

Figure 2.8 Knee joint modeled by involving the kinematics (rotation) in different states
Taken from Bindernagel et al. (2011)

In other approached, modeling individual shapes in a hierarchical structure and defining their relative positions in a separate model is applied (Cerrolaza, Villanueva, & Cabeza, 2012). The global model is created at low resolution, while each object is modeled at high resolution to capture shape variations more accurately (Figure 2.9). This approach, known as multiresolution analysis, is particularly effective for shapes that exhibit high local variability and low global variability (Cerrolaza et al., 2012; Wilms, Handels, & Ehrhardt, 2017).

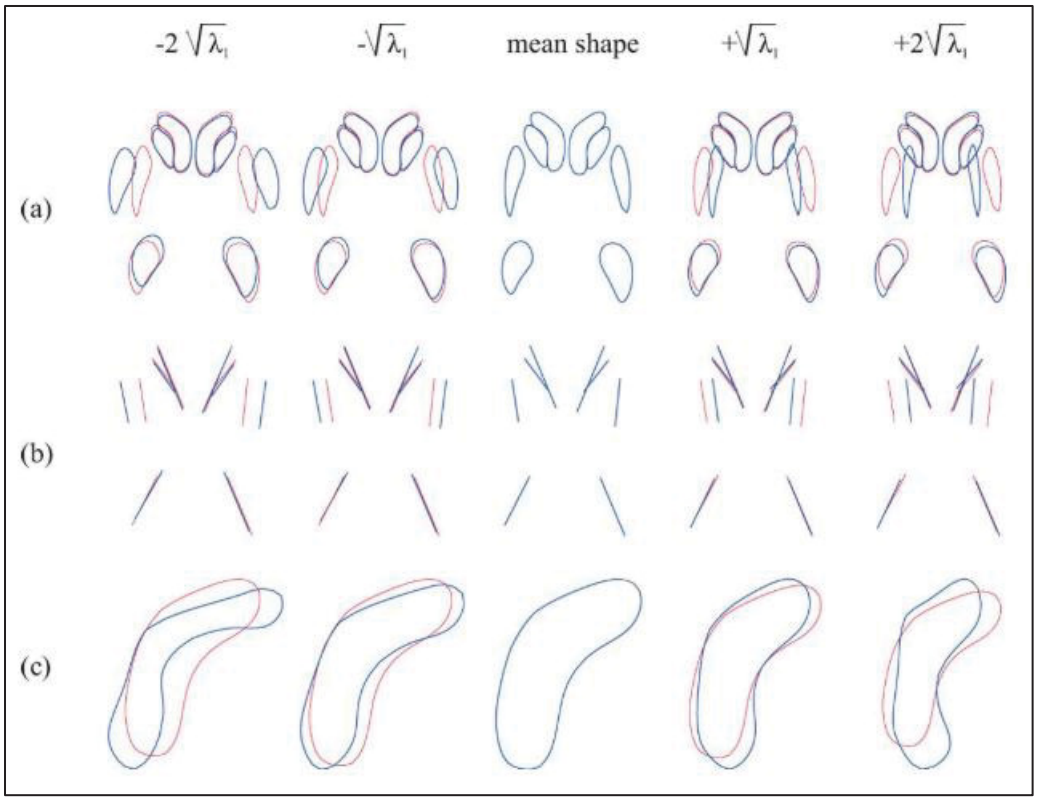

Figure 2.9 Brain MRI shapes in the multi-resolution SSM with all objects modeled at high resolution (a) while the resolution decreases for the other level (b) and finally modeling pieces at high resolution (c)
Taken from Cerrolaza et al. (2012)

Another approach involves concatenating the principal components of all bones and analyzing them as a unified model (Gollmer, Simon, Bischof, Barkhausen, & Buzug, 2012; Montillo, Song, Liu, & Miller, 2013). By concatenating the objects, the relationships between the shapes are also encoded within the model (Wilms et al., 2017). In a mesh-based dataset consisting of L objects, with N cases, the multi-object meshes can be represented as shown in Equation (2.12). Here, V_i^L and F_i^L denote the vertex and face matrices, respectively, for the Lth object of the ith shape in the dataset:

$$\begin{aligned} S_i^{mo} &= \{V_i^{mo}, F_i^{mo}\}, i = 1..N \\ V_i^{mo} &= [V_i^1; V_i^2, \dots, V_i^L], i = 1..N \\ F_i^{mo} &= [F_i^1; F_i^2, \dots, F_i^L], i = 1..N \end{aligned} \tag{2.12}$$

This representation ensures that each object within a shape is explicitly defined by its geometric structure, facilitating further processing in statistical shape modeling and multi-object analysis. Having $V_i^{mo}, i = 1..N$ presented with the same topology, a GPA process (Equation (2.3)) can align all shapes. Finally, PCA is applied on the multi-object meshes to have the mean shape and variations. Equation (2.13) depicts the multi-object PCA, where $\overline{S^{mo}}$ represents the mean shape of the multi-object model, and Φ^{mo} is the multi-object eigenvectors matrix. Multi-object shape parameters $b_k^{mo} \in R^M$ follows a normal distribution bounded by $\pm 3\sqrt{\lambda_k^{mo}}$ (Balestra et al., 2014; Bindernagel et al., 2011; T. F. Cootes et al., 1995; Tchinde Fotsin et al., 2019).

$$\overline{S^{mo}} = \frac{1}{N}\sum_{i=1}^{N} S_i^{mo}$$
$$S^{mo}(b_1^{mo}, b_2^{mo}, \dots, b_{N-1}^{mo}) = \overline{S^{mo}} + \sum_{k=1}^{N-1} b_k^{mo}.\psi_k^{mo} \quad (2.13)$$

In multi-object models, it is crucial to define a criterion to ensure that adjacent bones do not interfere with each other (Zheng, Guoyan; Yu, 2017). This can be achieved by placing a separator between the structures, such as drawing a line between the tibia and fibula (Zheng, Guoyan; Yu, 2017) or by using specific points as landmarks (Brehler et al., 2019) as illustrated in Figure 2.7. Encoding the relationships within the model also helps prevent the bones from interfering with one another.

Articulated SSMs explicitly encode joint mechanics by incorporating rotational or translational degrees of freedom between bones. These models are highly effective for anatomies like the hip or knee, where the articulation follows well-defined biomechanical constraints (Balestra et al., 2014; Bindernagel et al., 2011; Tchinde Fotsin et al., 2019). However, implementing such models requires detailed geometric and arithmetic definitions of the joint, including the articulation axis and motion constraints, which may not be available or reliably measurable in all anatomical regions.

In the context of ankle modeling, a study generated a coupled model of the foot and ankle, excluding the tibia and fibula, and focusing solely on the talus (Brehler et al., 2019). They utilized manual localization of landmarks to avoid overlap between the bones, but they did not model the relationships between the bones. Other researchers have employed planes (e.g., in knee joints) or imposed limitations on joint movement based on the joint's geometric structure (e.g., in the femur and pelvis). These methods, however, are not applicable to the ankle, as the joint's shape sometimes obscures visibility in certain areas, and the bones are tightly connected. Consequently, modeling the relationship between the bones is a challenging task.

Multi-resolution modeling is particularly effective for capturing fine-grained local variability, especially in anatomies with large-scale spatial organization, such as the brain or abdomen. However, such models are computationally demanding to implement, requiring a layered registration pipeline and separate training stages for local and global components (Cerrolaza et al., 2018).

In contrast, concatenation-based coupled models offer a balanced compromise between anatomical realism and implementation feasibility. PCA enables the model to learn not only the individual variations but also the co-variation patterns across objects. Unlike articulated models, this method does not require explicit joint mechanics, and unlike multi-resolution models, it simplifies training by using a single statistical process for all objects. It also inherently maintains inter-object relationships, reducing the likelihood of bone overlap or unrealistic deformations when properly aligned and regularized (Brehler et al., 2019).

2.2.2.5 Parametric Models

In parametric modeling, anatomic and clinical parameters are considered as information (Hosseinian & Arefi, 2015). Instead of using all geometrical points as proposed in SSM, some anatomical Descriptive Parameters (DP) from the anatomy of interest are extracted and used in the reconstruction process (Hosseinian & Arefi, 2015). Figure 2.10 shows a parametric model of femur, represented by its main features: femoral neck shaft angle and three spheres fitting the femoral head and the posterior condyles (Chaibi et al., 2012). Other parameters such as length of the diaphysis, the trochanters, the femoral neck and the trochlea are also added to

the model to complete the simplified parametric model (Chaibi et al., 2012) . The parametric models may be used to build statistical shape model. In a recent work, statistical spine model is built using a global shape of the spine and local geometrical representation of vertebra using simplified parametric model (B. Aubert et al., 2019). Figure 2.11 shows the spine statistical model. All parameters are concatenated and statistically processed to have the final model (B. Aubert et al., 2019).

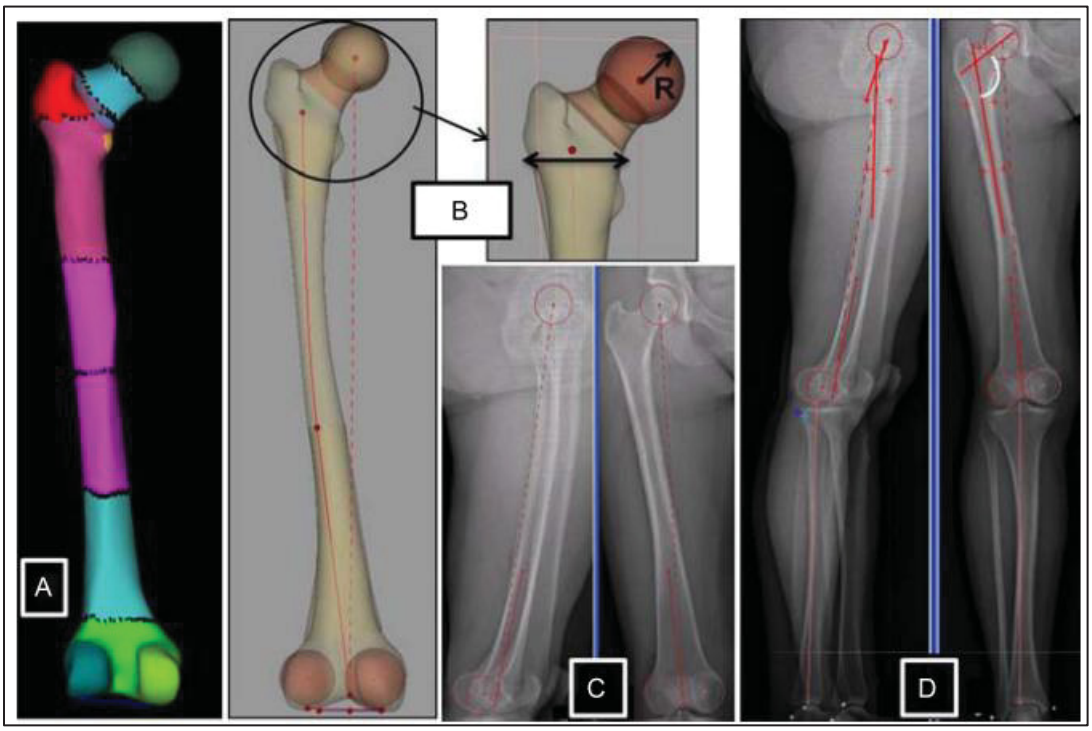

Figure 2.10 A parametric model of femur (B)
Taken from Chaibi et al. (2012)

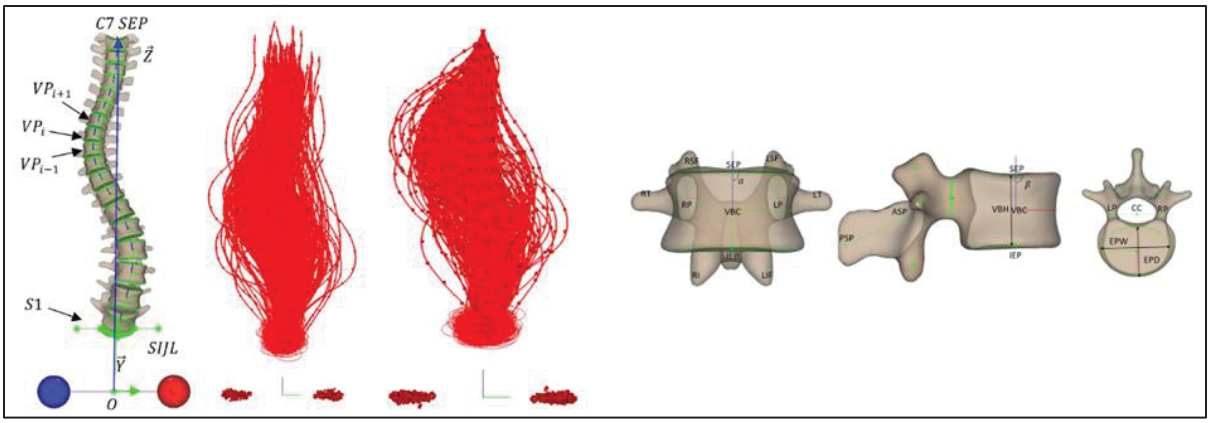

Figure 2.11 Global statistical spine model and the local simplified parametric vertebrae model
Taken from Aubert et al. (2019)

These methods offer a fast and reliable initialization for the reconstruction process as they model the shape with fewer parameters than SSMs (Hosseinian & Arefi, 2015). Various geometric primitives, such as points, lines, circles, spheres, and cylinders, are employed to represent the bone shapes. Similar to the creation of an SSM, a database of parameters is generated from the training data, and the model is then constructed based on this information. The parametric models have been used to reconstruct the spine and vertebra (Cresson et al., 2009; Humbert, De Guise, Aubert, Godbout, & Skalli, 2009), proximal femur (Baudoin, Skalli, de Guise, & Mitton, 2008), lower limb (Chaibi et al., 2012).

The advantages of parametric models include their ability to incorporate prior anatomical knowledge through other models such as SSMs, leading to more accurate and robust reconstructions. However, parametric models' applicability becomes limited in anatomical regions with complex or ambiguous joint geometry where the structures are not representable by primitives such as spheres, cylinders, etc.

2.2.3 Model-based 2D/3D Reconstruction

With prior 3D information about the structure of interest encoded into models, the model can be deformed such that its characteristics align with those observed in 2D radiographs (P. Markelj, Tomaževič, Likar, & Pernuš, 2012). The model-based reconstruction framework

offers a principled way to constrain the ill-posed nature of 2D-to-3D reconstruction by ensuring anatomical plausibility through the learned shape priors. This is applied through estimation of the model and pose parameters such that its projections best match the 2D images (Dworzak et al., 2010). This process yields a 3D representation consistent with the information contained in the radiographs. Two key aspects of this process are the type of information extracted from the model and radiographs for establishing correspondence, and the method employed to perform the matching between them (P. Markelj et al., 2012). Another important factor is the degree of freedom of the geometric transformation, which can be either rigid or non-rigid. Rigid registration involves three translation and three rotation parameters, and it is defined as a transformation where no spatial distortions are introduced (Bakic et al., 2006; Xin Chen et al., 2007; Guoyan, 2007; Knaan & Joskowicz, 2003; Livyatan, Yaniv, & Joskowicz, 2007 2003; Weese et al., 1997; Zöllei, Grimson, Norbash, & Wells, 2001). In contrast, non-rigid transformations occur when structures undergo deformation during the transformation process. Developing personalized models of bones is inherently a non-rigid registration task, as bone structures vary among different patients (Candemir et al., 2014; Cresson, Branchaud, Chav, Godbout, & de Guise, 2010; Zheng, 2011). In the context of model deformation for 3D reconstruction, achieving accurate patient-specific models typically requires non-rigid registration. However, many studies have relied on rigid transformations due to the increased complexity and computational demands associated with non-rigid methods (P. Markelj et al., 2012).

In following sections, first we explain different methods that employ the matching process and then, explore different information extracted from the images and the models such as points, contours, etc.

2.2.3.1 Reconstruction methods

The 2D/3D registration process aims to align a 3D model and one or more 2D images, into a common coordinate frame (P. Markelj et al., 2012). Figure 2.12 depicts a geometrical setup of 2D/3D registration process, which elaborates finding transformations T_j for all j images that brings the 3D model X_A^{3D} in the coordinate system S_w, to its projection in 2D, $X_{B(j)}^{2D}$ and $X_{B(j+1)}^{2D}$,

which are presented in coordinate systems S_j and S_{j+1}, respectively. It is conventional that S_j and S_{j+1} are orthogonal planes.

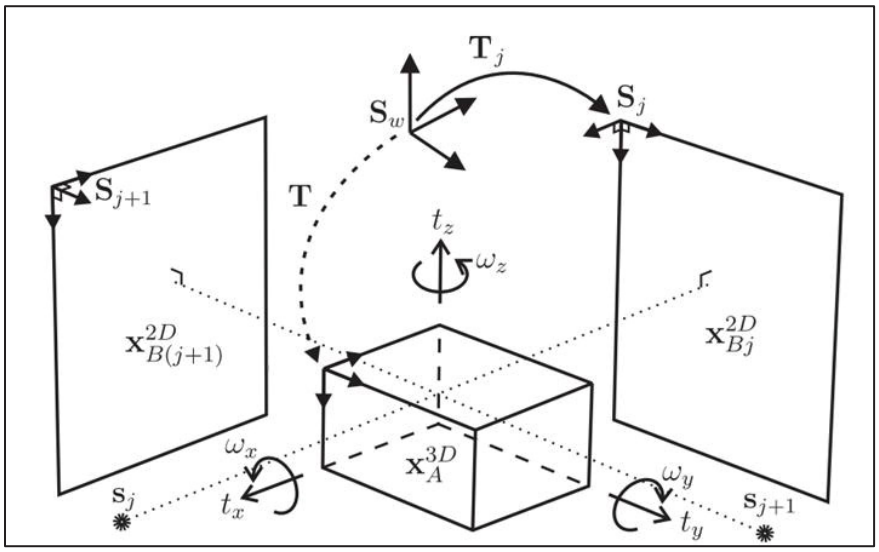

Figure 2.12 Registration of a 3D image into two 2D images
Taken and Modified from Markelj et al. (2012)

The alignment process can be performed using two main strategies: either by projecting the 3D data into a 2D coordinate space and conducting registration in 2D or by backprojecting the 2D data into 3D space, allowing registration to take place between two 3D structures. Figure 2.13 and Figure 2.14 illustrate both approaches for registering a pre-interventional 3D image $A(x_A^{3D})$ with N 2D intra-interventional images $B_i(x_{B^i}^{2D})$, where $i = 1..N$. Equation (2.16) depicts finding the optimal transformation.

$$\dot{T} = argmax_T \sum_{j=1}^{N} CF_j^{2D} = argmax_T \sum_{j=1}^{N} CF^{2D}(P_j(A^T(x_A^{3D})), B_j\left(x_{B_j}^{2D}\right)) \tag{2.14}$$

In the 2D-based registration approach (Figure 2.13), the goal is to determine a transformation T that aligns the 2D projections of the transformed 3D image $A^T(x_A^{3D})$, denoted as $P_i(A^T(x_A^{3D}))$ with the 2D images $B_i(x_{B^i}^{2D})$ based on a set of 2D criterion functions CF_i^{2D} for $i = 1..N$. Alternatively, in the 3D-based registration approach (Figure 2.14), the intra-

interventional 2D images $B_i(x_{B^i}^{2D})$ are back-projected into 3D space, forming $\beta_i(B_i(x_{B^i}^{2D}))$. The registration is then performed in 3D, ensuring that the transformed version of the pre-interventional image $A^T(x_A^{3D})$ aligns with all back-projected images using 3D criterion functions CF_i^{3D}. Equation (2.17) depicts finding the optimal transformation. Additionally, Figure 2.15 provides a visual representation of both the projection and back-projection processes.

$$\dot{T} = argmax_T \sum_{j=1}^{N} CF_j^{3D} = argmax_T \sum_{j=1}^{N} CF^{3D}(A^T(x_A^{3D}), \mathrm{B}_j(B_j\left(x_{B_j}^{2D}\right))) \tag{2.15}$$

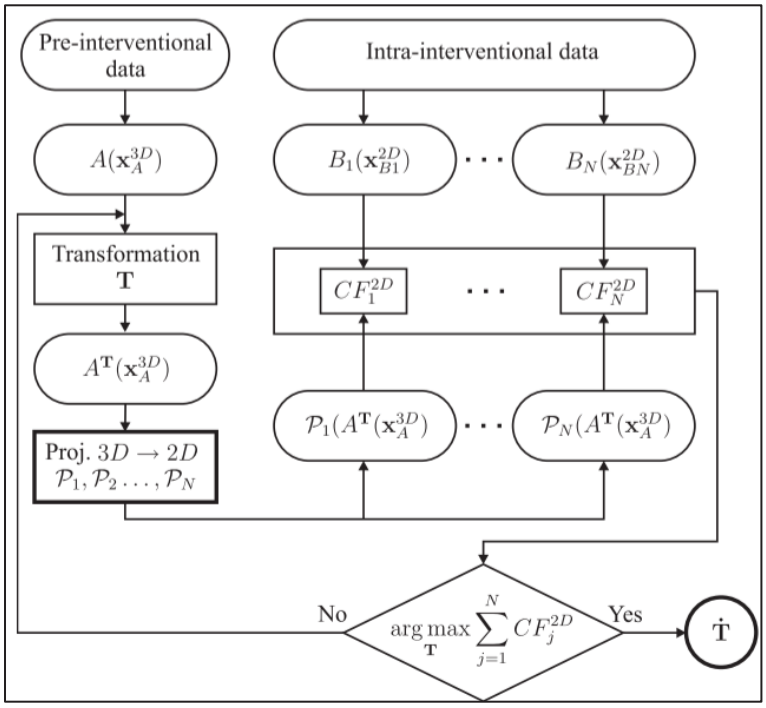

Figure 2.13 2D/3D registration process based on projection strategy
Taken from Markelj et al. (2012)

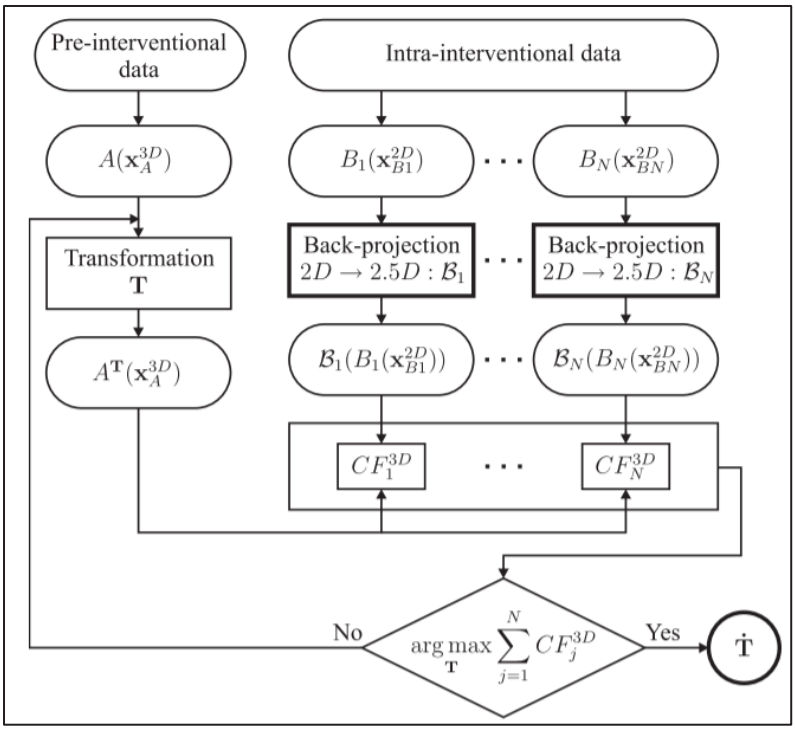

Figure 2.14 2D/3D registration process based on back-projection strategy
Taken from Markelj et al. (2012)

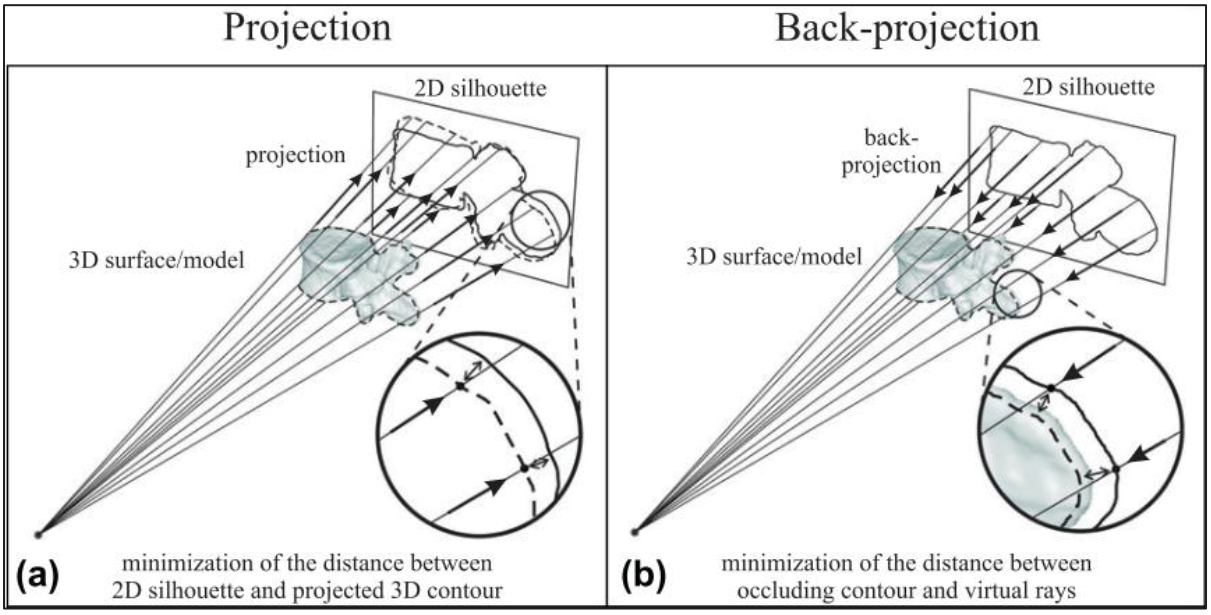

Figure 2.15 Projection and back-projection of a vertebra for registration process

Taken from P. Markelj et al. (2012)

The 2D/3D reconstruction methods are traditionally relied on registration techniques that iteratively deform and align the 3D model to one or multiple images in 2D (P. Markelj et al., 2012). These registration-based methods solve an optimization problem to find the best spatial and shape parameters of the 3D model, such as a Statistical Shape Model, to minimize the discrepancy between the model's and the actual X-ray images (Chaibi et al., 2012; J. Chen et al., 2025; Heimann & Meinzer, 2009). Alternatively, the parameters of the transformation T can be calculated which align the model to the radiographs. These parameters can either be computed directly or explored within a predefined search space, with the objective of minimizing the distance between image features, such as edges and contours, by optimizing an energy function (P. Markelj et al., 2012). These optimization searches can be conducted globally or locally to identify the optimal solution.

Several optimization techniques are commonly used for image registration. Gradient descent or ascent is used to optimize the cost function (Aouadi & Sarry, 2008; Bakic et al., 2006; Primož Markelj, Tomaževič, Pernuš, & Likar, 2008; Mertzanidou, Hipwell, Tanner, & Hawkes, 2010; Weese et al., 1997; Zöllei et al., 2001). These methods are simple to implement and effective for smooth and convex cost functions, but they com with the cost of possibility of being trapped in the local minima (Qiu, Hammernik, Qin, Chen, & Rueckert, 2021). In addition, their performance degrades with noisy data or poor initialization (Mertzanidou et al., 2014). Iterative downhill simplex algorithms (Knaan & Joskowicz, 2003; Livyatan et al., 2003) does not require gradient information which makes the implementation feasible in some cases that the gradient information may not be feasible to be calculated. On the other hand, they might not be efficient for high dimensional problems and may converge to non-optimal point. Levenberg-Best-neighbor search optimization (X. Chen et al., 2007; X. Chen, Varley, Shark, Shentall, & Kirby, 2008) may be trapped in local minima. The deformation algorithms thin-plate splines (TPS) (Harmouche, Cheriet, Labelle, & Dansereau, 2010), iterative closest point (ICP) (Cresson et al., 2009; Pan et al., 2023), moving least squares deformation (Cresson et al., 2010), and Free-Form Deformation (FFD) algorithms (Fang, Wang, Chen, Jian, & He, 2016) which are based on optimization strategy are very dependant on the initialization and may be stuck in local minima with poor initialization. Evolutionary algorithms are also used to search the best model parameters such as Gradient Descent Evolution optimization

(Lamecker et al., 2006) and Genetic Algorithm (Mahfouz et al., 2006; Tchinde Fotsin et al., 2019).

A common characteristic among the optimization-based approaches is that it is a critical prerequisite to determine an optimal initial model state which could lead the search to the global optimal point (P. Markelj et al., 2012). A poor initialization may mislead the optimization process, resulting in suboptimal reconstruction outcomes. In addition, iterative optimization could be time consuming for some cost functions (Goswami & Kr., 2015; P. Markelj et al., 2012).

In recent years, deep learning has enabled registration-free approaches for 2D/3D reconstruction, where a 3D shape or image is directly inferred from one or more 2D radiographs without the need for iterative alignment or optimization. These methods represent a shift from traditional model-to-image registration toward end-to-end learning-based frameworks that reconstruct 3D anatomy with high speed and minimal user intervention. The methods may range from receiving two radiograph and reconstruct the 3D CT-like volume image (Balakrishnan, Zhao, Sabuncu, Guttag, & Dalca, 2018; Y. Chen et al., 2024; Fuentes-Jimenez, Casillas-Perez, Pizarro, Collins, & Bartoli, 2018; Kasten et al., 2020; Jian Wang & Zhang, 2020; Zhao, Lau, Luo, Chang, & Xu, 2019, Arn Roth et al., 2024). Figure 2.16 shows the general scheme of these methods.

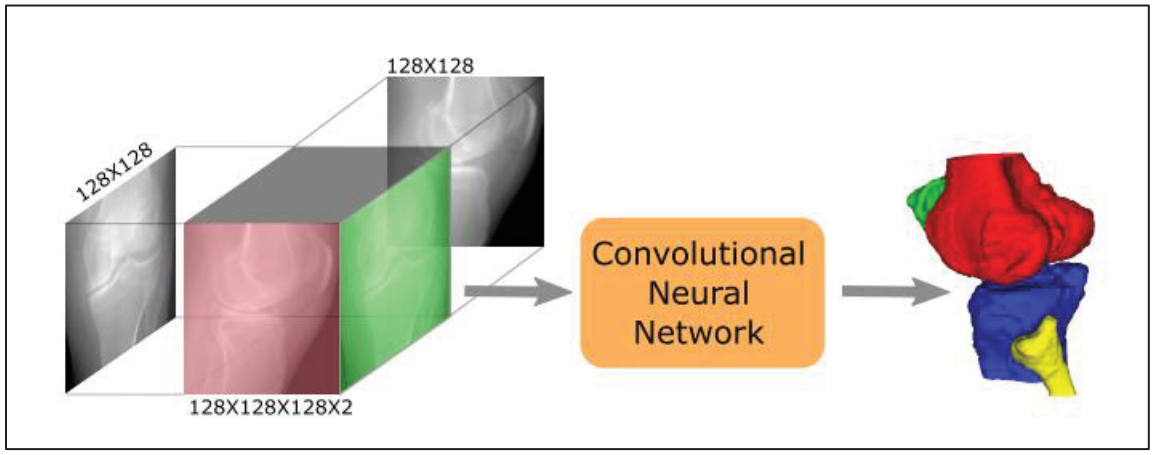

Figure 2.16 The general scheme of the registration-free reconstruction from biplanar radiographs
Taken from Kasten et al. (2020)

These methods leverage deep learning architectures, such as Convolutional Neural Networks (CNNs), to directly estimate the 3D shape or its underlying parameters from 2D radiographs without iterative alignment (Bayat et al., 2020). These approaches could learn a correspondence between 2D inputs and 3D structures (Arn Roth et al., 2024; Kasten et al., 2020), enabling real-time reconstruction without explicit registration or finding the pose, shape, or both parameters from the radiographs, using regressor models (Bayat et al., 2020; Hashemibakhtiar et al., 2024; D. X. Huang et al., 2022; Liu et al., 2025; Miao et al., 2016; WEI, Haishan, Alpers, Rak, & Hansen, 2021). The transformation parameters are revealed via receiving 2D images and through regression process. In other work dense deformation field is revealed with a deep regressor architecture to deform an atlas image to apply the reconstruction (van Houtte, Gao, Sijbers, & Zheng, 2021). Such direct regression or translation approaches bypass the need for handcrafted alignment objectives, making them particularly attractive for anatomies with complex geometry or low-contrast image conditions.

Hybrid methods also emerged to combine the traditional iterative optimization with deep learning approaches for guiding the reconstruction process, such as automatically localizing landmarks required for deformation on the radiographs (B. Aubert et al., 2019).

These registration-free methods offer several advantages. They are highly efficient and suitable for real-time applications, particularly in clinical workflows where computation time is a limiting factor (B. Aubert et al., 2019; Miao et al., 2016). They also demonstrate robustness to noise and variability in image acquisition, and in some cases, outperform classical registration pipelines in terms of reconstruction fidelity (Bayat et al., 2020).

However, these approaches are not without limitations. They typically require large, annotated datasets for training, which may not always be available for specific anatomies or pathologies (Fuentes-Jimenez et al., 2018; Mescheder, Oechsle, Niemeyer, Nowozin, & Geiger, 2018). Obtaining high-quality paired 2D-3D training data is difficult due to patient privacy concerns and limitations in imaging techniques. Some studies address this by generating synthetic data, but such data may not generalize well to real-world scenarios (R. J. Chen, Lu, Chen, Williamson, & Mahmood, 2021). Furthermore, models trained on a specific dataset may struggle to adapt to new imaging conditions or scanner types (Ghafoorian et al., 2017). Domain adaptation and transfer learning techniques are often required to improve generalization and

robustness (Guan & Liu, 2022). Additionally, high-resolution 3D volumes demand substantial memory and computational resources. To address this, strategies like network pruning, quantization, and multi-scale learning can be employed to optimize efficiency. Lastly, traditional registration techniques yield interpretable optimization results, whereas deep learning models often function as "black boxes".

In deep learning-based 2D/3D reconstruction, tuning of training parameters is crucial for model performance and generalization. Key considerations include dropout rate, learning rate, batch size, early stopping, hyperparameter optimization, output normalization, and data augmentation strategies.

- Optimization Algorithms: Training deep neural networks requires the use of efficient optimization algorithms to minimize loss functions in high-dimensional, often non-convex parameter spaces. Among the most widely used methods are Stochastic Gradient Descent (SGD) (Duchi, Hazan, & Singer, 2011), Adagrad (Duchi et al., 2011), and Adaptive Moment Estimation (Adam) (Kingma & Ba, 2014). Each of these methods has its pros and cons which are detailed in the literature (Liashchynskyi & Liashchynskyi, 2019).
- Dropout Rate: Dropout is a regularization technique that mitigates overfitting by randomly deactivating neurons during training. Optimizing the dropout rate is essential; excessively high rates can lead to underfitting, while very low rates may not sufficiently prevent overfitting. Studies have demonstrated that appropriate dropout rates enhance generalization in medical image processing tasks (Phukpattaranont & Rojchanaumpawan, 2025).
- Learning Rate: The learning rate determines the step size during weight updates, and the batch size defines the number of samples processed before updating the model. A high learning rate can cause the model to converge too quickly to a suboptimal solution, while a low rate may result in a long training process. Learning rate decay is a widely adopted strategy in training deep neural networks to improve convergence and model generalization. It involves gradually reducing the learning rate over time during training, allowing the optimizer to take large steps initially for rapid learning, and smaller steps later for fine-tuning around local minima. Among various decay

strategies, step decay, where the learning rate is reduced by a fixed factor after a predefined number of epochs (e.g., halved every 10 epochs), is commonly used due to its simplicity and effectiveness. Other approaches include exponential decay, polynomial decay, and adaptive methods, where the learning rate adjusts based on validation performance. Decaying the learning rate helps prevent overshooting and can lead to better convergence, especially in non-convex optimization problems typical in deep learning.

- Early Stopping: Early stopping is a strategy to prevent overfitting by halting training when the model's performance on a validation set stopped improving. This technique ensures that the model maintains optimal generalization capabilities.
- Batch size: It refers to the number of training samples processed simultaneously before the model's parameters are updated. It determines how many samples are used to estimate the gradient of the loss function during each iteration. A smaller batch size (e.g., 1 to 32) leads to more frequent updates with higher variance in the gradient estimates, which can help the model escape local minima but may result in noisier convergence. In contrast, a larger batch size (e.g., 128 or more) provides more stable and accurate gradient estimates but requires more memory and may lead to slower generalization or convergence to sharper minima. The choice of batch size is a trade-off between computational efficiency, training stability, and generalization performance.
- Hyperparameter Optimization: Selecting optimal hyperparameters is vital for model performance. Grid search (Liashchynskyi & Liashchynskyi, 2019) is a systematic method that exhaustively searches through a specified subset of hyperparameters. While computationally intensive, it can identify effective hyperparameter combinations that enhance model accuracy.
- Output Normalization in Regression Networks: Normalizing output values in regression tasks ensures that the model's predictions are on a comparable scale, facilitating more stable and efficient learning (Khakhar & Buckman, 2022). Normalization techniques, such as min-max scaling or z-score normalization, are

commonly employed to achieve this. Each parameter is scaled within its scope individually.

- Data Augmentation: Data augmentation enhances model robustness by increasing the diversity of the training dataset. It can be simply done by rotation, horizontal and vertical flips, resizing, and translation (Kumar, Asiamah, Jolaoso, & Esiowu, 2025). Recent advancements have introduced adversarial data augmentation techniques to produce challenging training samples, further enhancing model performance.

In addition to hyperparameter selection, the architecture of the neural network plays a critical role in performance. Two of the most commonly used Convolutional Neural Network (CNN) architectures for regression tasks are ResNet-50 (He, Zhang, Ren, & Sun, 2015) and VGG-16 (Simonyan & Zisserman, 2014). Although originally developed for classification, these architectures can be adapted for regression by modifying the output layers and activation functions (Lathuilière, Mesejo, Alameda-Pineda, & Horaud, 2018). They have been successfully applied to tasks such as anatomical landmark localization (Mehra, R, Anbarasi L, Ravi, & Al Mazroa, 2024), pose parameter estimation (Kendall, Grimes, & Cipolla, 2015), and alignment assessment (Tack, Preim, & Zachow, 2021).

When dealing with multi-channel input such as biplanar radiographs, one strategy involves processing the AP and LT images through separate channels and subsequently concatenating their features at the dense layer level (Benjamin Aubert, Vazquez, Cresson, Parent, & De Guise, 2016). Alternatively, concatenating the images at the input level and processing them through shared convolutional layers allows the network to learn common representations, which can exploit mutual anatomical information present in both views.

2.2.3.2 Information used in the model-based Reconstruction process

The information that is used in the registration process may be features, intensity, or gradient information (P. Markelj et al., 2012). These features are extracted directly from the image, such as points and contours, pixel intensities, edges information, or a combination of all these factors (Hosseinian & Arefi, 2015). While points represent the simplest features used in registration, as seen in methods like Iterative Closest Point (ICP) (Bakic et al., 2006; Candemir

et al., 2014; Fang et al., 2016; Kadoury et al., 2009), contours provide richer information (Arnold et al., 2001; X. Chen et al., 2007, 2008), which can be used to minimize silhouette distances in both 2D and 3D. On the other hand, intensity-based approaches utilize pixel values and methods such as Maximum Intensity Projection to generate Digitally Reconstructed Radiographs (DRR) (Aouadi & Sarry, 2008; Guoyan, 2007; Knaan & Joskowicz, 2003; Mertzanidou et al., 2010; Shalaby, Farag, Ross, & Hockenbury, 2012; Weese et al., 1997; Zheng, 2008). However, the computational complexity of generating DRR for every iteration in the registration process is a notable limitation of these methods. Gradient-based methods also use processed intensity information (Primož Markelj et al., 2008; Tomazevic et al., 2002). The following information is used in the reconstruction process:

- Stereo Corresponding Points and Non-Stereo Corresponding Points: In Stereo Corresponding Points (SCP), points, the low-level primitives, are identified and matched across multiple images during the registration process (Hosseinian & Arefi, 2015). In this approach, an expert identifies at least 6 corresponding points across two or more X-ray images (Hosseinian & Arefi, 2015). These points can then be backprojected into 3D and used in the reconstruction process. A drawback for methods based on points is not being suitable for reconstructing bony structures with continuous shapes due to the absence of distinct anatomical landmark (Hosseinian & Arefi, 2015). Figure 2.17 shows SCP on a vertebra. Non-Stereo Corresponding Points (NSCP) improve the Challenges in SCP methods (Hosseinian & Arefi, 2015). These points can be identified on one of the images only which helps the reconstruction process where the corresponding points are not visible or identifiable in all images. NSCP is defined as a point located along the line connecting the X-ray source to the projection of the point in a single view (Hosseinian & Arefi, 2015). method enables 3D reconstruction by incorporating additional landmarks, thereby increasing the accuracy of the process (Mitton et al., 2000; Mitulescu, Semaan, De Guise, Leborgne, & Adamsbaum, 2001; Mitulescu, Skalli, Mitton, & De Guise, 2002). Figure 2.18 shows NSCPs on AP and LT views of a vertebra.

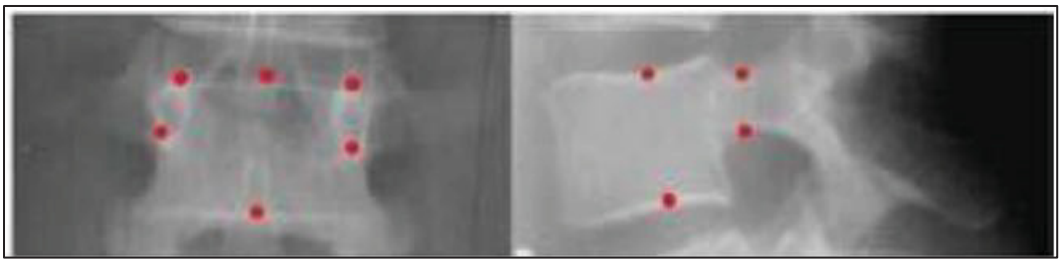

Figure 2.17 SCPs shown on AP and LT views of a vertebra
Taken from Hosseinian & Arefi (2015)

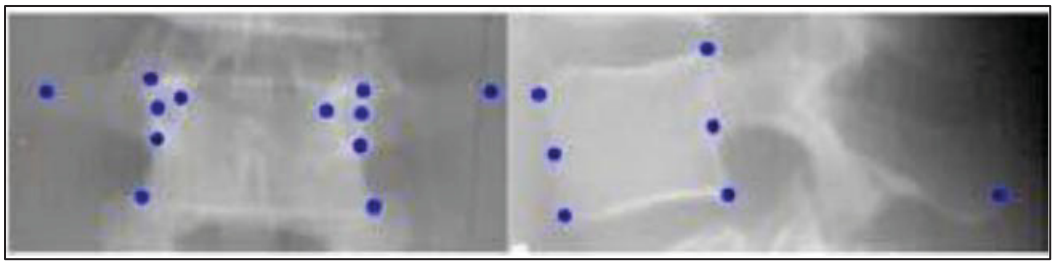

Figure 2.18 NSCPs shown on AP and LT views of a vertebra
Taken from (Hosseinian & Arefi, 2015)

- Contours: As mentioned, point-based methods are unsuitable for bony structures with continuous shapes, such as the knee joint, due to the absence of specific landmarks (Laporte et al., 2003). Contour-based methods were subsequently introduced to enhance the 3D reconstruction process from stereo radiographs using Non-Stereo Corresponding Contours (NSCC) (Laporte et al., 2003). These NSCC are identified on 2D radiographs and matched to the silhouette of the model projection through an iterative optimization process (Baka et al., 2011). This process results in higher accuracy in comparison to point-based methods (Laporte et al., 2003). Figure 2.19 shows examples of NSCC on femur and pelvis. Different deformable models are integrated with NSCC (Dong et al., 2007; Dong & Zheng, 2009; Fleute & Lavallée, 1999; Galibarov et al., 2010; Laporte et al., 2003; Wei, Wang, & Chen, 2009; Zheng, 2006; Zheng et al., 2009; Zheng & Nolte, 2006). In some works, NSCC is combined with point-based features to increase the reconstruction accuracy (Akkoul, Hafiane, Lespessailles, & Jennane, 2013; Mitton et al., 2006).

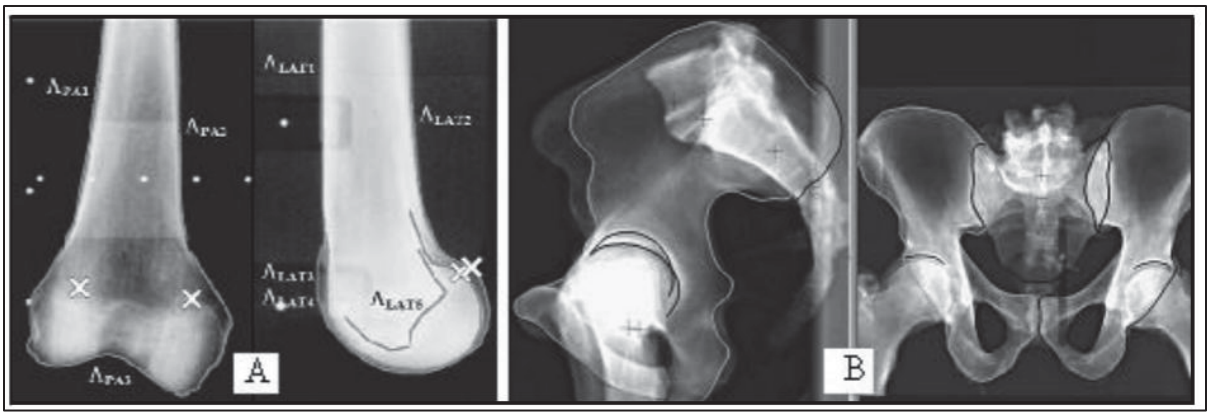

Figure 2.19 Anatomical contours on (A) distal femur and (B) pelvis
Taken from Laporte et al. (2003) and Mitton et al. (2006)

In contour-based methods, higher-level geometric primitives are utilized compared to point-based methods, thereby avoiding the challenge of finding corresponding points. While contour-based methods offer acceptable accuracy, they still require a significant amount of time for 3D reconstruction, e.g., 15 to 35 minutes for lower limb (Hosseinian & Arefi, 2015). Additionally, contours may introduce errors when segmentation or edge extraction is not performed accurately (Hosseinian & Arefi, 2015).

- Intensity: Intensity-based 2D/3D registration methods rely on directly comparing the pixel intensities of a 3D anatomical model's synthetic projections, known as Digitally Reconstructed Radiographs (DRRs), to those of the acquired 2D X-ray images (Zheng, Guoyan; Yu, 2017). DRRs are generated by simulating X-ray attenuation through the volumetric model. rays are cast from a virtual source through the CT volume or mesh-based representation, and line integrals of intensity (e.g., Hounsfield units) are computed to form a projection image that mimics actual radiographs (Levine & Levine, 2020; C. J. F. Reyneke et al., 2019). Once DRRs are produced at a given pose, an optimization algorithm deforms or reorients the 3D model to minimize a similarity metric between the DRR and the real X-ray. Common intensity-based metrics include normalized cross-correlation, mutual information, and gradient correlation, which measure the degree of alignment in image intensities or their gradients (Heimann & Meinzer, 2009; P. Markelj et al., 2012). By iteratively adjusting the model's shape and pose to maximize such similarity, the reconstruction process converges on a 3D

configuration whose projections best match the observed radiographs, ensuring an anatomically plausible fit grounded in the actual image data.

Although, DRRs are used extensively in the model-based 2D/3D reconstruction process, the generated DRRs are different in characteristics than a real radiograph (B. Aubert, Cresson, De Guise, & Vazquez, 2023). Thus, Image translation methods are presented to make these images as similar as possible (B. Aubert et al., 2023). Other works are done to alter the radiographs to have the desired image content, such as removing bony structures from lung X-ray (Kalisz & Marczyk, 2021) using an autoencoder which can be used to alter pixel intensities.

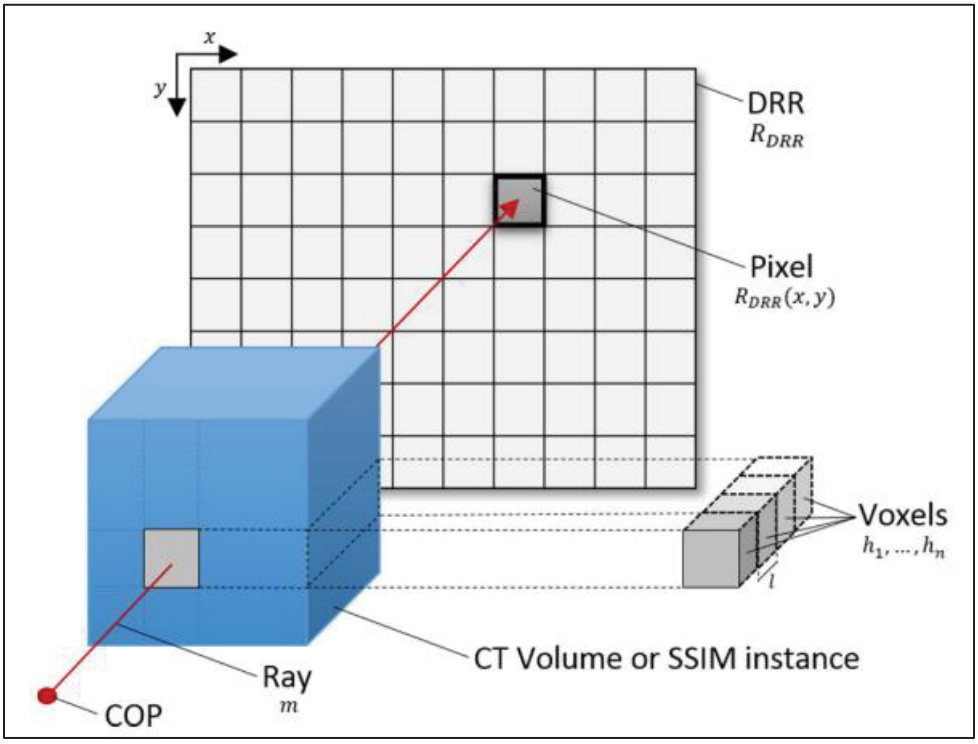

Figure 2.20 Generation of DRR using ray-casting
Taken from C. J. F. Reyneke et al. (2019)

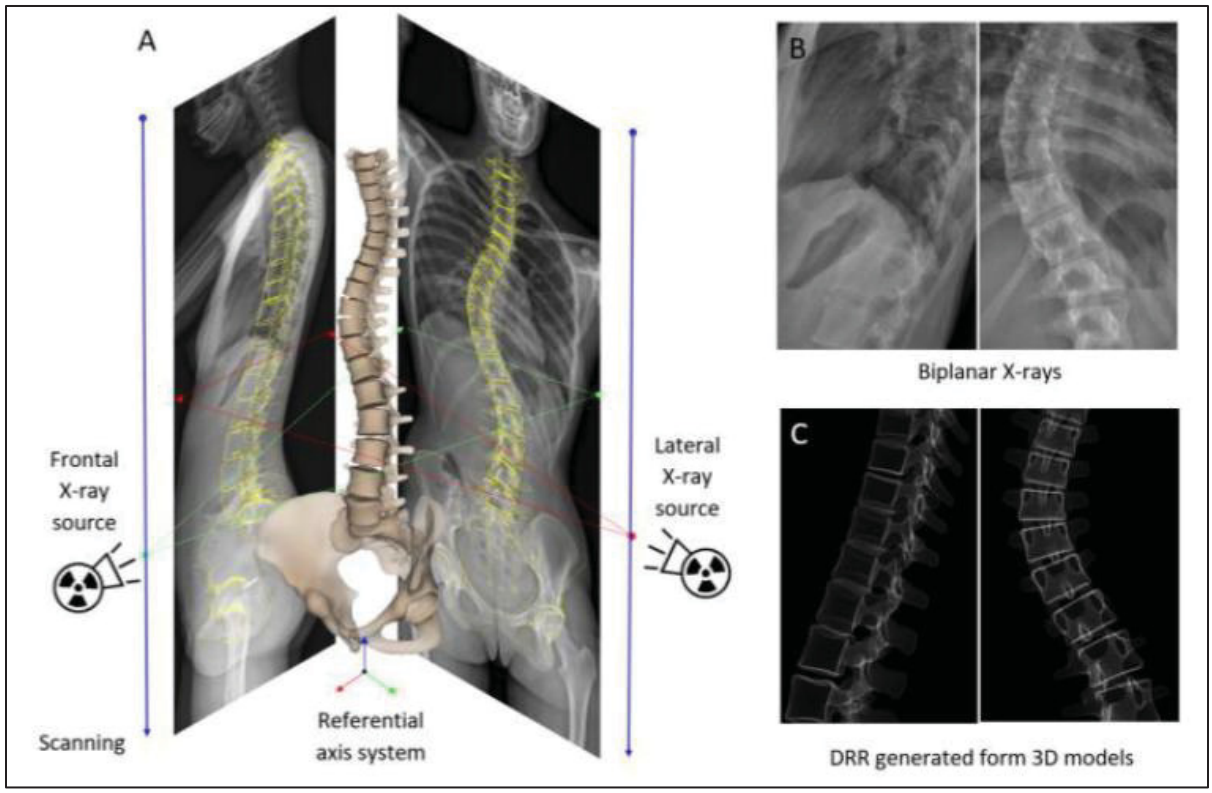

Figure 2.21 Biplanar X-rays from the spine vs. the DRR generated from 3D model
Taken from Aubert et al. (2023)

2.2.4 Level of automation

The reconstruction approaches can be fully automated, semi-automated, or manually executed (Unberath et al., 2021). Manual reconstruction methods typically require specialists to manually place dozens, or even hundreds of anatomical landmarks or to trace the full bone contours on each radiograph which a is labor-intensive and time-consuming process (Subburaj, Ravi, & Agarwal, 2009). For instance, techniques that rely on identifying anatomical landmarks typically require 2 to 4 hours to process the entire spine, which limits their practicality for routine clinical use (Pomero, Mitton, Laporte, De Guise, & Skalli, 2004b).
Semi-automated approaches significantly reduce this workload. These methods usually rely on a small number of user-defined reference points, such as the centre of superior and inferior endplates of vertebrae (Moura et al., 2011) or center of the femoral head (Chaibi et al., 2012). Once these points are provided, the reconstruction can be done much quicker than manual

process, e.g., 1.5 minutes for vertebrae (Moura et al., 2009) by refining a SSM through intensity-based or contour-based registration.

Finally, fully automated approaches methods go even further by removing the need for user input altogether. These approached can directly estimate the 3D bone image (Kasten et al., 2020), or the parameters of a predefined shape model (Nguyen et al., 2023) from one or more input radiographs. The resulting reconstructions are typically generated in under a second, while maintaining consistent accuracy (Kasten et al., 2020).

Given the time-consuming nature of manual techniques and their strong dependence on the operator's expertise, recent research and commercial efforts are increasingly shifting toward semi- and fully automated methods to improve efficiency, reliability, and clinical usability.

2.3 Calculation of clinical parameters on the reconstructed 3D model

After 3D reconstruction (whether statistical shape model–based, intensity-based, or learning-based), most studies follow a geometric analysis workflow to calculate clinical parameters like lengths, angles, and distances. Different approaches are applied to retrieve the clinical measurements:

- Automatic calculation: In this approach, the parameters are mentioned in the parametric model (B. Aubert et al., 2019; Boutillon, Salhi, Burdin, & Borotikar, 2022) or the parameters are calculated automatically on the reconstructed model based on geometrical definitions (Cerveri et al., 2010; Chaibi et al., 2012; Liu et al., 2025).
- Manual calculation: In this approach after the reconstruction, manual annotations are used to define landmarks and calculate clinical parameters
- Transfer landmarks in Statistical Shape Model deformation: As each point on the shape is assumed to correspond to the same anatomical location on each training example (T. F. Cootes et al., 1995), the model captures the main modes of variation of the object (Timothy F. Cootes, 1977) and as the deformation to the shape is also applied according to the learned modes of variation, the selected landmark on the mean shape of the model represents the same anatomical landmark after deformation is applied on the shape.

This can be used to assign landmarks on the mean shape and automatically calculate clinical measurements using the same assigned landmarks.

2.4 Validation Methods and Metrics

In the domain of 2D/3D reconstruction from biplanar images, rigorous validation metrics and methods are essential to assess the accuracy and reliability of the reconstructed models. These evaluations encompass various aspects, including the quality of the Statistical Shape Model used for the reconstruction, the performance of learning-based registration networks, the accuracy of geometric reconstructions, and the accuracy of the calculated clinical measurements. In following sections, we explored the appropriate metrics and methods.

2.4.1 Statistical Shape Model

To assess the Statistical Shape Model (SSM), four key metrics are introduced, which can also be used to indirectly evaluate the quality of the correspondences across the dataset (M. A. Styner et al., 2003):

- Compactness: Compactness measures how efficiently the statistical shape model (SSM) captures shape variations using the least number of parameters. A more compact model retains the most significant shape variations while minimizing redundancy, making it more efficient for practical application (Audenaert et al., 2019; Goparaju et al., 2021). Equation (2.16) calculates the compactness of model using M Modes of Variation (MoV) where the λ_m is the m-th eigenvalue of the model. The higher compactness values, one need lesser variable to reconstruct the shapes, considering the population variance.

$$Compactness(M) = \frac{1}{\sum \lambda} \sum_{m=1}^{M} \lambda_m \tag{2.16}$$

- Generalization: Generalization assesses the model's capability to accurately represent unseen shapes that were not part of the training set. A well-generalized SSM should reconstruct new instances with minimal error, demonstrating its robustness in adapting

to novel anatomical variations (Audenaert et al., 2019; Goparaju et al., 2021). The metric is the average reconstruction error over all the performed K tests, which may be calculated using Equation (2.17). S_i^{rm} is the excluded shape and $S_i^{rm'}(M)$ is the reconstructed shape using M MoV. A smaller error in the model's generalization ability indicates a more robust and well-constructed shape model.

$$Generalization(M) = \frac{1}{K}\sum_{i=1}^{K}\left\|S_i^{rm} - S_i^{rm'}(M)\right\|^2 \tag{2.17}$$

- Specificity: Specificity quantifies how realistic the shapes generated by the model are, ensuring they resemble actual anatomical structures. It is evaluated by synthesizing new shapes from the model and comparing them to real examples; a low specificity error indicates that the generated shapes remain within the natural anatomical variations (Audenaert et al., 2019; Goparaju et al., 2021). The approximation error is defined as the distance between a generated shape instance $S_i^{rm}(M)$ and its closest matching sample in the training dataset, denoted as $S_i^{rm'}$. The specificity metric is then calculated as the average approximation error across all N generated shape instances as shown in Equation (2.18)

$$Specificity(M) = \frac{1}{N}\sum_{i=1}^{N}\left\|S_i^{rm}(M) - S_i^{rm'}\right\|^2 \tag{2.18}$$

- Accuracy: Accuracy measures how well the statistical shape model reconstructs shapes from its training data. It is often evaluated by comparing reconstructed shapes to their original counterparts and computing the deviation between them. High accuracy indicates that the model effectively captures shape details without significant loss of fidelity (Audenaert et al., 2019; A. Patil, Kulkarni, Xie, Bull, & Jones, 2023). Equation (2.19) is used to calculate model's accuracy.

$$Accuracy(M) = \frac{1}{N}\sum_{i=1}^{N}\left\|S_i^{rm'}(M) - S_i^{rm}\right\|^2 \tag{2.19}$$

Validation of the SSIM has been explored in a few articles in the literature (C. J. F. Reyneke et al., 2019).

2.4.2 Reconstruction

Validating the performance of 2D/3D registration techniques is critical to ensure the accuracy, robustness, and clinical applicability of the methods. For assessing the 3D geometric reconstruction accuracy in comparison to the ground truth 3D model, we may use Point-to-Point or Point-to-Surface metrics.

Point-to-Point error (P2P) error measures the Euclidean distance between corresponding points on the reconstructed model P and the ground truth Q which are represented by vertices, p_i and q_i where $i = 1..N$. This metric is useful when there is a known correspondence between vertices on both models. Equation (2.20) represents this error.

$$P2P(P,Q) = \frac{1}{N}\sum_{i=1}^{N}\|p_i - q_i\|^2 \tag{2.20}$$

As the model Q is also representing some error from the surface S which it is made from, the gold standard metric for the reconstruction process is Point-to-Surface (P2S) error which measures the error between the reconstruction model P and the surface S which could be achieved from segmentation mask of the organ of interest (Besl & McKay, 1992). Considering S as a very dense mesh with M vertices where $N \ll M$, the P2S error can be calculated using Equation (2.21)

$$P2S(P,S) = \frac{1}{N}\sum_{i=1}^{N} min_{s\in S}\|p_i - s\|^2 \tag{2.21}$$

In the context of regression while optimizing the model parameters we may use the Mean Squared Error (MSE) or Mean Absolute Error (MAE) between the output of the regressor Y and the ground truth $\hat{Y}$ where they have N parameters. MAE is presented in Equation (2.22). In addition, the Hausdorff distance quantifies the greatest point-to-point distance between two

sets of points (e.g., the reconstructed 3D surface and the CT scan surface). A lower Hausdorff distance indicates a better match between the two surfaces (Taha & Hanbury, 2015).

$$MAE(Y, \hat{Y}) = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{2.22}$$

In the context of image translation and image reconstruction, we may use the Peak Signal-to-Noise Ratio (PSNR) to assess the quality of the output of the translation process and the ground truth. Considering Y as the output image of the method, $\hat{Y}$ the desired output, and I the maximum intensity value in the image, the Mean Squared Error (MSE) and the PSNR may be calculated using Equations (2.23) and (2.24).

$$MSE(Y, \hat{Y}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2.23}$$

$$PSNR(Y, \hat{Y}) = 10 \times \log_{10}\left(\frac{MAX_I^2}{MSE(Y, \hat{Y})}\right) \tag{2.24}$$

The Structural Similarity Index (SSiM) is a perceptual metric that quantifies image quality degradation caused in image reconstruction (Z. Wang, Bovik, Sheikh, & Simoncelli, 2004). Unlike simpler metrics as MSE or PSNR, SSiM evaluates perceived visual quality by incorporating luminance, contrast, and structural information. Given two images x and y, SSiM may be calculated using Equation (2.25) where μ_x and μ_y are means of x and y, σ_x and σ_y are the variance of the images, σ_{xy} is the covariance between the two images, and C_1 and C_2 are stabilizing constants which are calculated by Equation (2.26) where L is the dynamic range of the pixel values, $K_1 \approx 0.01$ and $K_2 \approx 0.03$.

$$SSiM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{2.25}$$

$$C_1 = (K_1 L)^2, C_2 = (K_2 L)^2 \tag{2.26}$$

2.4.3 Clinical Measurement

Validating clinical measurements derived from 2D/3D reconstruction models is crucial to ensure their reliability and applicability in clinical settings. The calculated measurements are compared against the ground truth which are measured by one or more experts. If more than one expert has participated in the experiment, the mean of their observations is used as the reference (Benjamin Aubert, Vazquez, Cresson, Parent, & De Guise, 2016).

The mean bias error (MBE) between observations of a specific parameter O_i and their reference values R_i in a dataset of N reconstructed shapes where $i = 1..N$ can be calculated as in Equation (2.27).

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(O_i - R_i) \tag{2.27}$$

The standard deviation of the error also may be calculated using Equation (2.28) where $e_i = O_i - R_i$ and $\bar{e}$ is the mean of errors e_i.

$$SDE = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(e_i - \bar{e})^2} \tag{2.28}$$

In addition, Bland-Altman plots are widely used in medical imaging research to assess the agreement between two measurement techniques (Martin Bland & Altman, 1986). These plots are typically employed to compare the clinical measurements (e.g., bone lengths, joint angles, distances between landmarks) obtained from the reconstructed 3D models against a reference

standard, often measurements extracted directly from CT scans or manually annotated ground-truth data.

The Bland-Altman analysis visualizes the difference between the two methods (y-axis) against their mean (x-axis) for each measurement. This allows researchers to observe any systematic bias and to assess the limits of agreement, typically defined as the mean difference ±1.96 times the standard deviation of the differences. A narrow range within these limits, centered around zero, indicates strong agreement and supports the clinical reliability of the reconstruction pipeline.

Considering measurement sets M_1 and M_2 with size of n from two different methods, following are the formula for plotting the Bland-Altman. Equations (2.29) depicts mean of the measurements, difference of the measurements, mean difference (Bias) and Standard Deviation of differences, respectively. The upper and lower Limits of Agreements (LoA) are defined as $\bar{d} + 1.96 \times SD_d$ and $\bar{d} - 1.96 \times SD_d$. Figure 2.22 shows an example of Bland-Altman plot for Cobb angle calculation agreement between the proposed method and the reference (B. Aubert et al., 2019).

$$
\begin{aligned}
Mean_i &= \frac{M_{1i} + M_{2i}}{2} \\
Diff_i &= M_{1i} - M_{2i} \\
\bar{d} &= \frac{1}{n}\sum_{i=1}^{n}(M_{1i} - M_{2i}) \\
SD_d &= \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}((M_{1i} - M_{2i}) - \bar{d})^2}
\end{aligned}
\tag{2.29}
$$

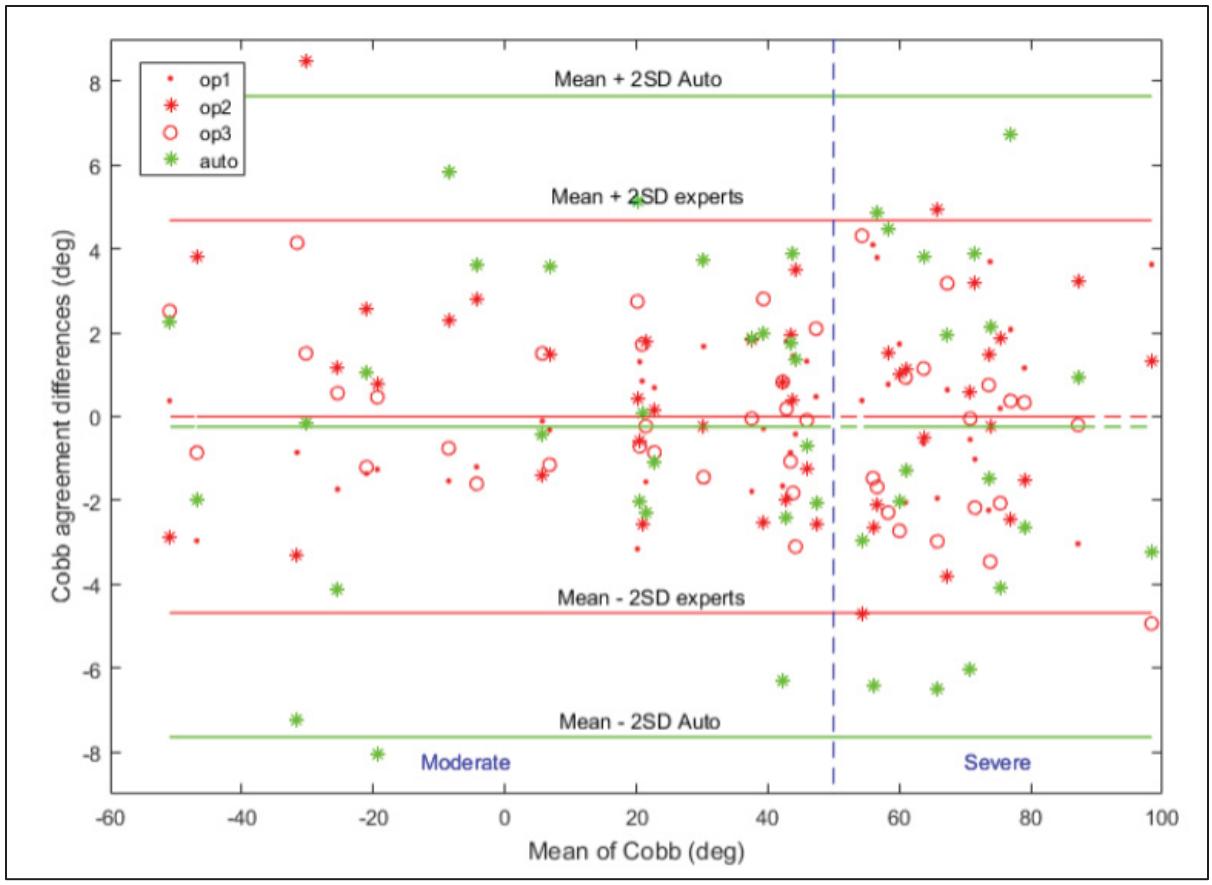

Figure 2.22 Bland-Altman plot example
Taken from B. Aubert et al. (2019)

2.5 Conclusion

This chapter reviewed existing approaches for 2D/3D reconstruction from biplanar radiographs, with a specific focus on applications targeting the ankle joint. Accurate reconstruction is critical, as the tibiofibular joint features inter-bone distances ranging from 1 mm to 16 mm, necessitating a reconstruction precision of 1 mm or better to ensure clinically valid outcomes.

The 3D reconstruction process typically involves three core components: the model used as prior anatomical knowledge, the reconstruction methodology applied to align the model with radiographic data, and the type of information leveraged to guide the reconstruction.

Generic models, while simple to implement, do not account for anatomical shape variability. This lack of representational fidelity increases the search space during optimization, particularly problematic in regions like the ankle, where closely connected structures and occlusions challenge both feature extraction and intensity-based matching. In contrast, Statistical Shape Models (SSMs) incorporate inter-subject shape variability, enabling more anatomically plausible reconstructions and reducing the optimization complexity by constraining the solution space. However, effective SSM construction requires a large dataset of segmented anatomies with point-wise correspondence across samples.

An extension of this approach, the SSIM, integrates both shape variation and intensity, resulting in more realistic Digitally Reconstructed Radiographs (DRRs). This combination enhances the fidelity of the reconstruction, especially in regions with poor geometric features. Nevertheless, the inclusion of voxel- or pixel-based intensity information demands pre-alignment of the shapes for statistical analysis and may significantly increase computational complexity. Employing coarser intensity representations can mitigate this overhead, albeit at the expense of reduced DRR realism.

For anatomical structures composed of multiple interacting bones, such as the tibiofibular joint, articulated or coupled models should also be considered. In practice, concatenation-based models offer a more feasible solution in anatomically ambiguous regions, such as the ankle syndesmosis, where joint movement is minimal. These models simplify implementation while maintaining sufficient anatomical relevance for reconstruction tasks.

Parametric models, which rely on simplified geometric representations, are commonly used in joints with well-defined shapes, such as the femoral head, which approximates a sphere and can be parameterized with a few rotational degrees of freedom. However, the bones of the tibiofibular joint possess irregular and non-convex geometries that are not easily captured using simple parametric forms.

Registration-based methods align a prior model to radiographic projections by iteratively minimizing the mismatch between DRRs and actual radiographs. These methods are computationally expensive and sensitive to initialization due to the non-convex nature of the optimization landscape, often leading to suboptimal convergence in the presence of local minima.

Conversely, registration-free, learning-based approaches offer improved computational efficiency and robustness. These methods learn the mapping between input radiographs and model or pose parameters directly, often incorporating high-level feature extraction into their neural architectures. This enables accurate reconstruction even under occlusion or limited visibility. However, such approaches require access to large, annotated datasets, which are currently unavailable for the ankle, posing a major limitation for practical deployment.

Incorporating intensity information via SSIMs provides rich anatomical detail and contributes to the generation of realistic DRRs. In contrast, feature-based methods, such as those relying on landmarks or contours, depend on an additional preprocessing step, either through expert annotation or automated feature extraction. Point-based methods are particularly limited when anatomical landmarks are difficult to identify in 2D radiographs due to structural overlap. These methods also struggle to reconstruct continuous anatomical surfaces. Contour-based methods may offer improved reconstruction of bone boundaries but rely heavily on accurate edge detection, which is also compromised in the presence of superimposed anatomical structures.

Calculating clinical parameters such as inter-bone distances and joint angles on reconstructed 3D models necessitates reliable localization of anatomical landmarks. While manual annotation remains common practice, it introduces variability and reduces reproducibility. In contrast, when using SSMs, the deformation process preserves landmark correspondence, allowing for automatic and consistent measurement extraction following reconstruction.

In summary, the combination of Statistical Shape and Intensity Models with registration-free, learning-based reconstruction techniques presents a promising framework for automated 3D reconstruction of the ankle joint from biplanar radiographs. Despite current limitations, including data availability and computational complexity, this approach holds significant potential to enhance accuracy, efficiency, and clinical applicability in orthopaedic imaging.

CHAPITRE 3

OBJECTIVES OF THE STUDY AND CONTRIBUTIONS OF THE THESIS

3.1 Introduction

In the previous chapter, we reviewed the literature on 2D/3D reconstruction from X-ray images, focusing on the challenges involved, models, information and the methods proposed to address them. Numerous approaches have been explored to tackle the problem, each with their own advantages and drawbacks. We also reviewed the applications, with particular emphasis on those related to the foot and ankle. In this chapter, we present the problematics and objectives of the study and the methods we employed to achieve these objectives. Our goal is to bridge the gaps and overcome the challenges highlighted in the literature, while incorporating the best practices from previous studies. Finally, the contributions of this thesis will be discussed.

3.2 Problem statement

Despite the growing use of 3D reconstruction techniques in orthopedic imaging, the ankle joint, specifically the tibiofibular syndesmosis, remains underexplored compared to more commonly studied joints such as the hip, spine, and knee. Furthermore, current syndesmosis clinical measurements still rely heavily on two-dimensional (2D) imaging modalities, such as CT slices or radiographs, which fail to capture the complex 3D spatial relationships between tightly connected bones in the ankle. This limitation hinders accurate assessment of critical clinical parameters and may compromise diagnostic and treatment decisions. Furthermore, existing 3D reconstruction methods are time-intensive due to optimization-based registration or are not yet adapted to the anatomical structure of the ankle in learning-based approaches. In addition, accuracy of the reconstruction process must comply with the tight space in the ankle which is around 1 mm wide. There is a clear need for a low-dose, automated, and anatomically precise 3D reconstruction framework that can operate on biplanar radiographs while the patient is

standing, addressing the challenges of bone overlap, tightly connected bones, articulation ambiguity, and limited visibility in ankle imaging.

3.3 Objectives of the study

This study focuses on the 3D reconstruction of the tibiofibular joint from biplanar radiographs of the ankle, with the goal of calculating the syndesmosis clinical parameters. The primary challenge in this work is that structures in ankle radiographs often overlap, making image features difficult to process in some cases. For the tibiofibular joint, we only need to represent the tibia and fibula bones; however, these bones overlap with each other and with other structures, such as the talus. Figure 3.1 shows these superposed areas.

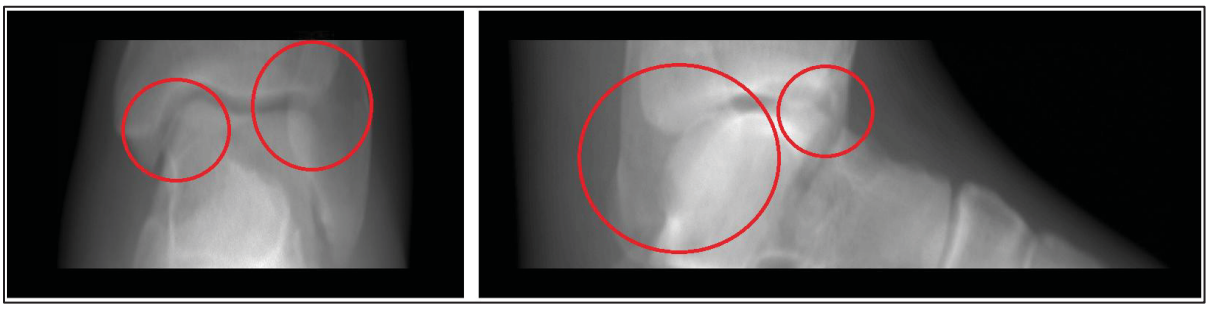

Figure 3.1 AP and LT radiographs of ankle indicating superposed tibia, fibula and talus

Additionally, soft tissues in the radiographs could affect the bones' visibility, complicating the reconstruction process. The 3D reconstruction must be highly accurate (e.g., within 1 mm), as the distances between the tibia and fibula in the tibiofibular joint can be as small as 1 mm. As the bones are tightly connected, the reconstruction process should consider the relation between the bones and prevent them from intersecting. The landmarks used for calculating clinical parameters must also be precisely localized, and ideally, this localization should be automated to facilitate the automatic calculation of clinical parameters. Furthermore, the non-rigid 2D/3D registration process must function in an environment where image features are often barely visible, which could potentially mislead the algorithm.

In conclusion, the following are the objectives of this study:

- Develop a pipeline which receives ankle biplanar radiographs (AP and LT) and reconstructs the tibiofibular joint with an average surface reconstruction error with 1 mm or less.
- Automatically calculate the syndesmosis clinical parameters on the built model in 3D

In the following section, we will discuss the methods that will help us achieve these objectives.

3.4 Overview of the Methods

To address the outlined objectives and overcome the associated challenges, we integrated key strengths from existing reconstruction methodologies identified in the literature. Our approach combines SSIM with registration-free learning-based techniques to enable accurate 3D reconstruction of the ankle joint, suitable for assessing syndesmosis-related clinical parameters. The SSIM incorporates prior anatomical and intensity information, which proves particularly advantageous in scenarios where bony structures are obscured due to superimposition. To enhance efficiency and robustness, we employed regressor networks capable of directly estimating model parameters by learning from the features in DRRs, thereby producing 3D models whose projections closely align with the input DRRs. The synthetic training dataset required for this learning process was generated using the SSIM. Once the 3D model is reconstructed, clinically relevant parameters are automatically computed using predefined anatomical landmarks.

The following section details the methodological steps undertaken to implement and validate this approach.

3.4.1 Building a 3D Statistical Shape and Intensity Model of the Tibiofibular Joint

A SSIM of the tibiofibular joint must be developed to capture both the mean shape and the variations in shape and intensity of the bones. This model plays a key role in guiding the 2D/3D registration network, especially in the case of ankle in which bones are superposed in many regions of the radiographs. The use of SSMs allows capturing the bone shape variations and relation of the bones, enhancing the generalization of the algorithm. Joint modeling is also

incorporated into the SSIM to account for the relation between the bones. To create this model, a database of ankle CT scans is processed (segmentation, alignment, etc.). The SSIM integrates both shape and intensity information, which is used to generate multiple DRRs that guide the registration process.

3.4.2 Registering the 3D model of the Tibiofibular Joint to biplanar radiographs

With the 3D model of the joint, we develop a 2D/3D registration approach that learns the shape and pose parameters of the model directly from the DRRs, generated by the SSIM. This approach avoids the iterative registration, avoiding the local minima problem. We aim to use a regressor network to predict these parameters in a single shot. Although training such a network requires a large amount of data, this is not an issue as we can generate as many instances as needed from the model. Once the shape and pose parameters of the joint are obtained, we can reconstruct the 3D surface of the bones.

3.4.3 Isolation of the Tibiofibular Joint skeletal structure in biplanar radiographs

The other step is to isolate the tibiofibular joint from other structures, such as the talus and soft tissues, in the radiographs, as there are many unwanted structures in the images. The presence of other structures makes the 3D reconstruction process difficult. This allows the reconstruction network to focus solely on the registration task. Additionally, since the SSIM serves as prior information for the reconstruction process and includes intensity information of the bones, we should align the intensity information of the isolated joint in the radiographs with the intensity information represented in the SSIM. For example, if the intensity information in the SSIM is simplified using tetrahedral modeling, we will ensure that the radiographs match this simplified representation.

3.4.4 Calculation of clinical parameters on the Personalized Reconstructed Joint for diagnosis of Syndesmosis injury

After a pair of orthogonal radiographs are processed to isolate the joint and reveal the shape and pose parameters of the bone surfaces, the reconstructed 3D joint can be used to assess the syndesmosis clinical parameters. This is achieved using predefined landmarks on the model, eliminating the need to reidentify the landmarks. Therefore, the parameters can be calculated automatically as soon as the reconstruction is completed.

3.5 Contributions of the Thesis

The contributions of this thesis are as follows:

- Development of a tibiofibular SSIM to generate a large training dataset for the learning process, addressing the challenge of limited datasets for training such networks. Additionally, the model aids the registration process, particularly in situations involving occlusion, by providing a structured representation of the joint's shape and intensity.
- Development of a network that processes orthogonal DRRs of the joint’s projections and extracts the shape and pose parameters of the SSIM, facilitating the 3D reconstruction of the joint.
- Deployment of a method to isolate the tibiofibular joint by removing soft tissues and other bony structures, while adjusting the DRR intensity values to closely resemble the projections of the SSIM of the joint.
- Establishment of a pipeline that integrates all steps to process biplanar X-ray images of the ankle and reconstruct a 3D model of the tibiofibular joint surfaces in a single shot.
- Automatic calculation of syndesmosis clinical parameters from biplanar radiographs of the ankle.

3.6 Conclusion

In this chapter, we have outlined the challenges, the problematic and objectives of the study, and the contributions of this thesis. The next chapter will provide a detailed description of the methodology used to achieve these objectives.

CHAPITRE 4

PROPOSED METHODOLOGY

4.1 Introduction

In this chapter, we elaborate on the methodology to address the challenge of building a personalized 3D model of the tibiofibular joint from ankle biplanar images and calculating the syndesmosis clinical parameter automatically from the built 3D model. This is achieved by incorporating a SSIM of the tibiofibular joint and a deep regressor network to perform a one-shot 2D/3D registration process. Figure 4.1 shows the proposed pipeline, starting from the ankle radiograph and resulting in the 3D personalized model of the tibiofibular joint. The following steps are proposed:

- A: Localizing the ankle in the radiographs to focus on the Region of Interest
- B: Isolating and intensity-modifying the tibiofibular joint in the AP and LT radiographs, which facilitates the 2D/3D registration process by the regressor network in step C.
- C: Extracting the shape and pose parameters of the SSIM from the concatenated AP-LT images of the tibiofibular joint to generate a 3D personalized model of the joint.

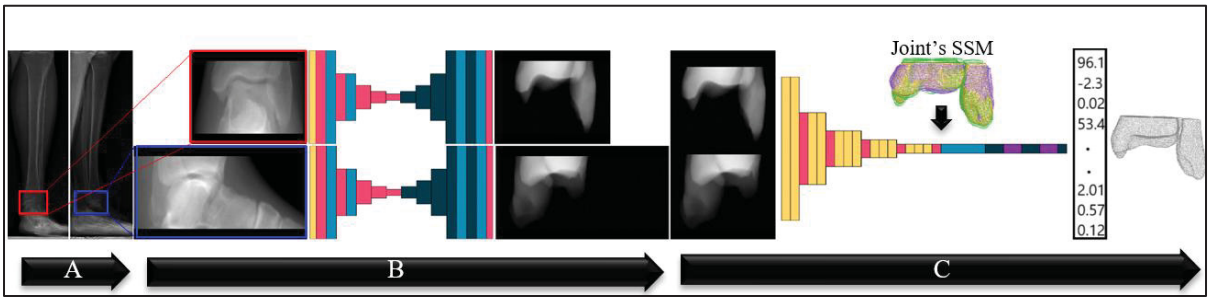

Figure 4.1 The pipeline of the proposed method

After 3D reconstruction, the syndesmosis clinical parameters can be calculated on the reconstructed joint surface. Figure 4.2 shows the parameters 'a', 'b' and 'c' on the 3D model. In the following sections, we elaborate on each step in detail.

To train the networks in the pipeline and guide the registration process, we need a dataset which will be discussed in the next section.

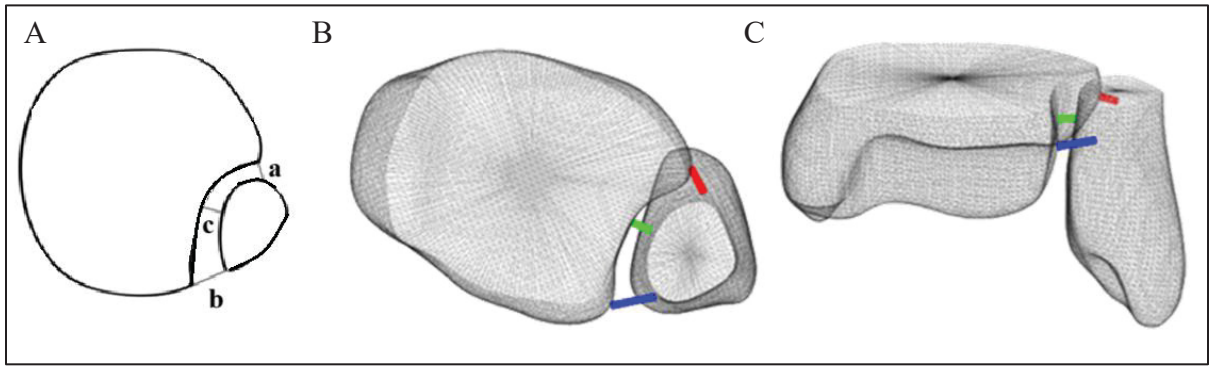

Figure 4.2 parameters 'a', 'b' and 'c', calculated on the CT scan slice (A) and on the personalized 3D model (B and C) in red, blue and green respectively

4.2 Dataset

Steps B and C in the pipeline (Figure 4.1) are learning-based approaches, requiring access to a CT scan database that includes the ankle structure. These CT scans are also incorporated into Step A to simulate the selection of Region of Interest in the evaluation process. A dataset of 60 ankle CT scans was collected following the required ethical approvals from Sacré-Cœur Hospital, ÉTS, and the CHUM Research Center in Montréal, Canada. These scans were specifically acquired to ensure the inclusion of slices necessary for calculating syndesmosis clinical parameters.

4.2.1 Cleaning and Preprocessing of CT Scan Dataset

The preparation of this dataset is a crucial step in the methodology, as the CT scan data serves as the foundation for all subsequent processes. One key observation is that the superior-most slice in each scan does not correspond to the same anatomical landmark across all cases. Figure 4.3 illustrates sagittal views of the lateral projection from three different cases, highlighting the variation in scan coverage between cases.

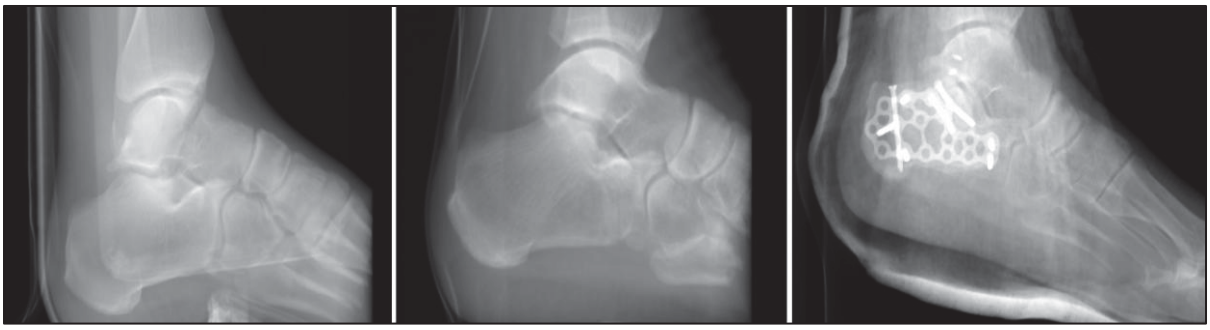

Figure 4.3 Sagittal view of three samples from the ankle dataset

To ensure consistency in the anatomical structures included in the dataset, particularly to maintain a uniform bone scope for model building process, the following selection criteria were applied:

- The Tibial plafond slice was manually identified in all 60 cases.
- As previously mentioned, the slice necessary for calculating the syndesmosis clinical parameters is located 15 slices above the Tibial plafond (Nault et al., 2013). Therefore, we verified that each case in the dataset included this specific slice.
- Subsequently, we selected cases that included an additional 15 slices above that specific slice, bringing the total number of slices above the Tibial plafond to 30. This 15-slice "safety margin" ensures that the syndesmosis calculation slice remains within the scope, even when adjustments, such as cropping the top of the model or other image processing operations, are performed.

Applying these criteria led to the removal of seven cases, reducing the dataset to 53 cases. These cases were then randomly split into two groups:

- 50 cases were allocated for training the networks in Steps B and C.
- 3 cases were reserved as unseen test cases for evaluating the pipeline's performance.

Since segmented bony structures and the foot are essential for both training the networks in Step B and generating the joint's Statistical Shape and Intensity Model in Step C, the next section details the segmentation process, followed by the other processing steps.

4.2.2 Multi-class Segmentation of the CT Scan Dataset

Once the dataset was cleaned, a semi-automatic multi-class segmentation method was applied to each CT scan case $CT_i, i = 1..53$ to segment the background, the joint (comprising the tibia and fibula), and the foot. Figure 4.4 presents a sample CT scan slice (A), along with the segmented tibia and fibula (B), and the foot segmentation (C).

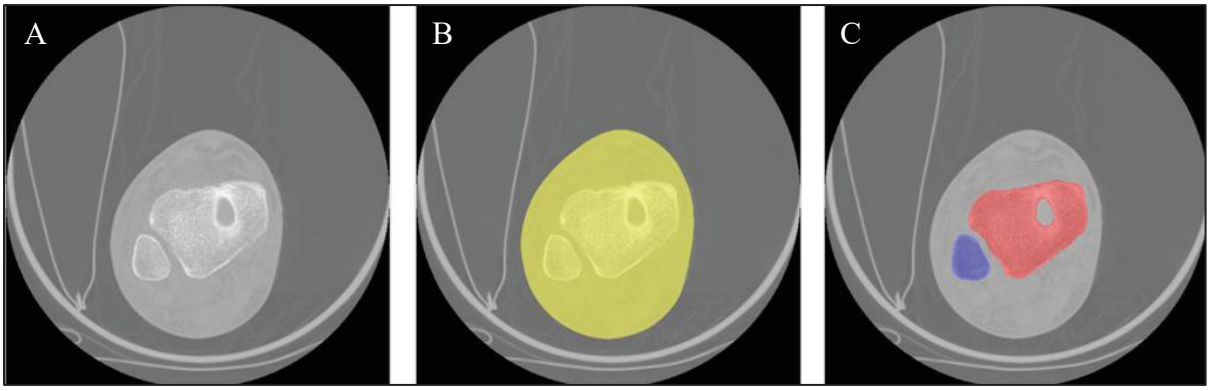

Figure 4.4 A sample of CT scan slice (A), the segmentation of the foot (B) and segmentation of Tibia and Fibula bones in red and blue respectively (C)

The foot segmentation is performed manually for the initial slice using the MATLAB Image Segmenter tool, specifically utilizing the 'Flood Fill' and 'Draw ROIs' functions (MathWorks, 2024). For the following slices, segmentation is automated using the geodesic active contour method (Caselles, Kimmel, & Sapiro, 1997). The purpose of foot segmentation is to isolate the foot within the 3D volume, represented by $ankle_i, i = 1..53$, and to remove any artifacts in the CT slices, such as the CT table, casts, or splints.

The joint segmentation is carried out manually across all slices to ensure accuracy. The MATLAB Image Segmenter tool is again used for this purpose. After segmentation is completed and the volume stored in $joint_i^{vol}, i = 1..53$, the bones are distinguished based on their size, with the distal tibia, $tibia_i^{vol}, i = 1..53$ being larger than the distal fibula, $fibula_i^{vol}, i = 1..53$. Figure 4.5 illustrates the segmented volumes of the foot and joint for a sample case from the dataset, viewed from a sagittal perspective.

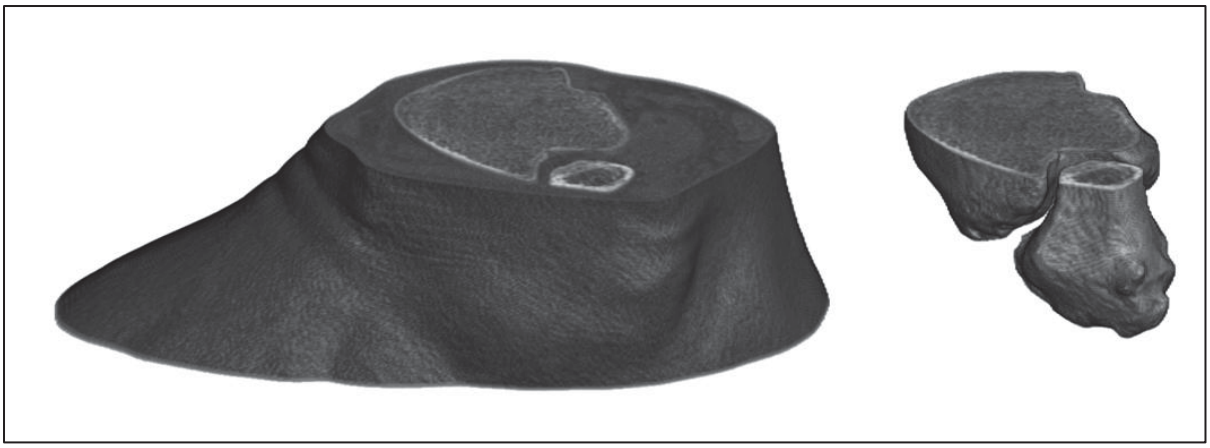

Figure 4.5 Segmented foot and tibiofibular joint volumes

The joint segmentation begins from the selected most superior slice (30 slices above the Tibial plafond) in each case and continues to the tip of the distal fibula, as illustrated in Figure 4.5. In contrast, the foot segmentation is extended an additional 15 slices below the tip of the distal fibula, ensuring a consistent safety margin in the inferior-most slices of the scan as the superior-most slices. The projection of randomly selected segmented volumes for both foot and joint using Voxel Projector is shown in Figure 4.6.

Since the ground truth for the segmentation process is not available, three manually segmented CT scan volumes were randomly selected and the segmentation masks were reviewed and verified by a medical professional in a slice-by-slice manner, ensuring accuracy and reliability. Upon completion of the segmentation process, the prepared dataset is ready for use in the different stages of the pipeline, including building the SSIM, execution of the registration process, and isolation of the joint in radiographs. Each step of the process will utilize specific parts of the dataset, which will be detailed in the following sections.

Figure 4.6 Sagittal view of the segmented foot and ankle projections generated by Voxel Projector (EOS imaging, n.d.)

4.2.3 Meshing the Segmented Joint

The first step is to generate a mesh on the bones. Since the CT scans represent both left and right ankles and the only difference between them is a flip over the sagittal plane, we flip all left ankle scans to ensure that all CT cases correspond to the right ankle structure. To generate meshes over the segmentation volume, we used Marching Cubes (Lorensen & Cline, 1987) to acquire triangular mesh over the surface of the tibia and fibula segmentation masks, tibia ($tibia_i^{vol}, i = 1..53$) and fibula ($fibula_i^{vol}, i = 1..53$), resulting in vertices uniformly distributed in 3D space. The top plane of the meshes is modeled using the circumference vertices and a single central vertex as shown in Figure 4.7. Figure 4.8 shows three random samples of the generated fibula ($fibula_i^{mesh}, i = 1..53$) and tibia meshes ($tibia_i^{mesh}, i = 1..53$). These meshes vary in the number of vertices and faces, as well as in their 3D position and orientation, as the bones differ in size and shape and the location in the CT image space. We also ensure that the mesh sizes, specifically the number of vertices, remain practical for computational efficiency. A high vertex count could significantly impact the performance of subsequent processing steps, including model training and registration. Therefore, an appropriate balance will be maintained to preserve anatomical accuracy while optimizing computational feasibility. The mesh generation method can be controlled for the number of vertices. We aimed for the minimum number of 10K vertices for each joint (Tibia and Fibula). However, dense meshes were also generated to provide an accurate representation of the bone surfaces, as they closely map the CT masks in 3D space and in millimeter representation.

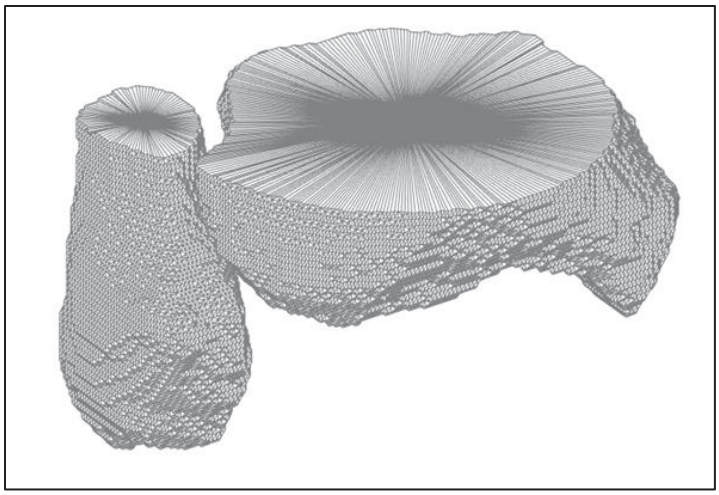

Figure 4.7 Top plane of the mesh modeled by the circumference vertices and one vertex in the center

Figure 4.8 Samples of fibula and tibia meshes, generated on the segmented masks

4.3 Building a 3D Statistical Shape and Intensity Model of the Tibiofibular Joint

Since the CT scan volumes are segmented, and the reconstruction process, along with other sections of the project are guided by the joint model, we constructed a SSIM of the tibiofibular joint. The model is generated based on 50 cases from the dataset, leaving 3 random cases for evaluation purposes. First, we will focus on shape modeling, followed by modeling the intensity of the joint.

4.3.1 Statistical Shape Modeling

Statistical Shape Modeling refers to the mean shape and plausible variations in the joint. The steps involved in generating the model are represented in Figure 4.9. First, the meshes should be matched (Figure 4.9.A) within each category (Tibia and Fibula). Then the meshes in each instance are concatenated to form the joint (Figure 4.9.B). The meshes are not aligned in the same coordinate system (Figure 4.9.C), so we need to align them first, and based on the aligned meshes we may have the mean shape (Figure 4.9.D) and the plausible variations (Figure 4.9.E) of the joint. We discuss these steps in the next sections.

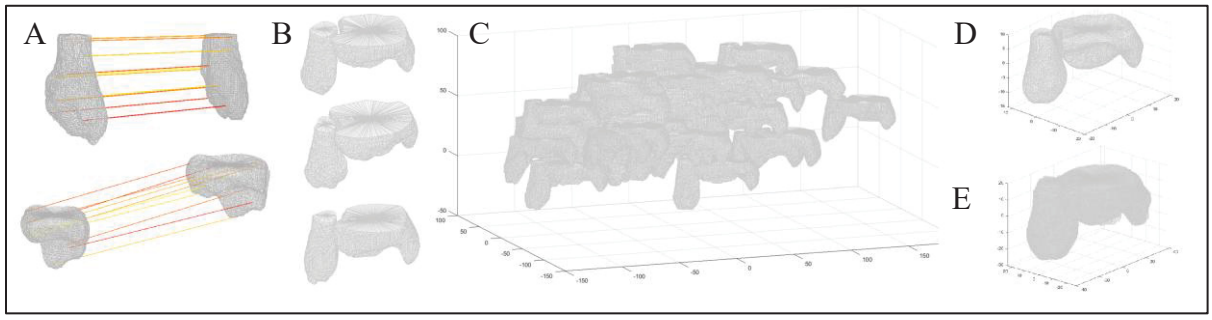

Figure 4.9 Steps for building SSM of the tibiofibular joint

4.3.1.1 Matching the Meshes

As previously mentioned, the meshes are not represented with the same topology due to differences in shape and size of the segmentation mask volumes. Therefore, matching and representing the meshes with the same topology is necessary before proceeding with generating SSM. The Smooth Shell mesh matching algorithm (Eisenberger et al., 2019) is used to find correspondences for all fibula ($fibula_i^{mesh}, i = 1..53$) and tibia meshes ($tibia_i^{mesh}, i = 1..53$) separately. This method results with the least matching error for the shapes representing small non-isometric deformation. For this purpose, the meshes in the smallest joint (based on the number of vertices in both shapes), are selected as $fibula_{target}^{mesh}$ and $tibia_{target}^{mesh}$ to be the reference shapes in which all other meshes are matched to, within their respective categories. Choosing the smallest shapes as the target makes sure that all source vertices go to a common

small target set of vertices. This may result in a vertex from $tibia_{target}^{mesh}$ and $fibula_{target}^{mesh}$ being matched to multiple vertices in their source shapes, as shown in Figure 4.10. When such situation happens (red vertices in Figure 4.11), the source vertex closest to the center of the region formed by all the matched vertices in the source shape (black dot in Figure 4.11) is chosen as the one-to-one matched vertex (marked vertex in Figure 4.11). As a result, all tibia meshes ($tibia_i^{m-mesh}, i = 1..53$) and fibula meshes ($fibula_i^{m-mesh}, i = 1..53$) are now represented with the same topology of vertices, denoted as $tibia_i^V$ and $fibula_i^V, i = 1..53$, respectively. $tibia_i^F$ and $fibula_i^F, i = 1..53$ are set equal to $tibia_{target}^F$ and $fibula_{target}^F$. Here, V and F are representing vertices and faces matrices. Algorithm 4.1 shows these steps.

Algorithm 4.1 Match Tibia and Fibula meshes

Match the Tibia and Fibula meshes through the dataset

Input:	Fibula meshes $fibula_i^{mesh}, i = 1..53$ and tibia meshes $tibia_i^{mesh}, i = 1..53$ with different topologies
Output:	Matched fibula meshes $fibula_i^{mesh}, i = 1..53$ and matched tibia meshes $tibia_i^{mesh}, i = 1..53$ with same topologies
1	Select the smallest Tibia mesh (according to the number of vertices) as $tibia_{source}^{mesh}$
2	Apply matching algorithm from $tibia_{source}^{mesh}$ to $tibia_i^{mesh}, i = 1..53$ where $tibia_{source}^{mesh} \neq tibia_i^{mesh}$ and store the result as $tibia_i^{m-mesh}, i = 1..53$
3	Select the smallest Fibula mesh (according to the number of vertices) as $fibula_{source}^{mesh}$
4	Apply matching algorithm from $fibula_{source}^{mesh}$ to $fibula_i^{mesh}, i = 1..53$ where $fibula_{source}^{mesh} \neq fibula_i^{mesh}$ and store the result as $fibula_i^{m-mesh}, i = 1..53$

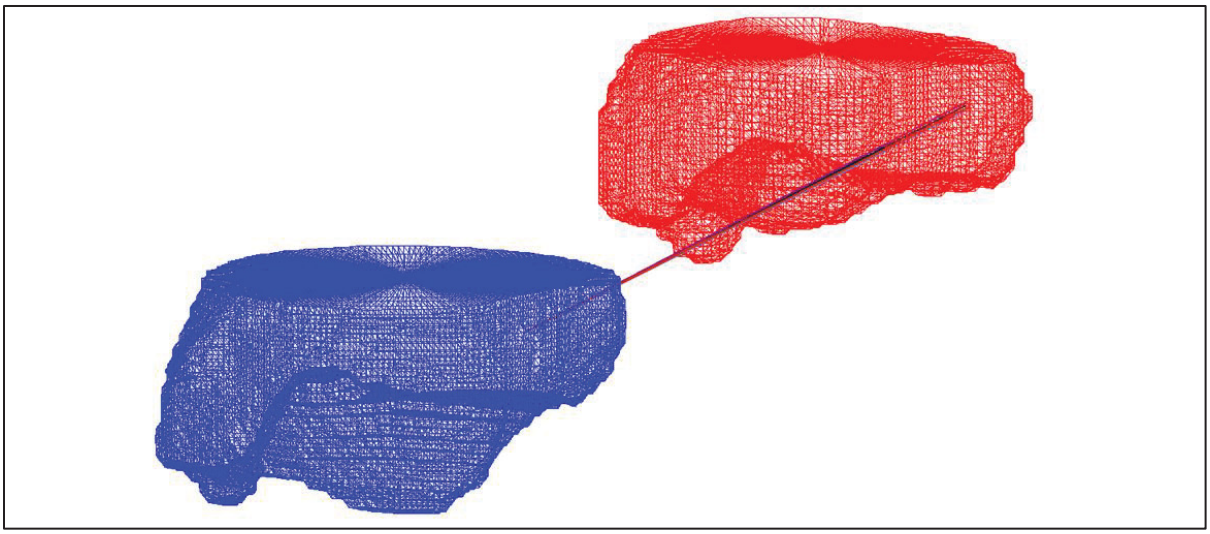

Figure 4.10 Multi-vertex matching from the source tibia (red) to a target tibia (blue)

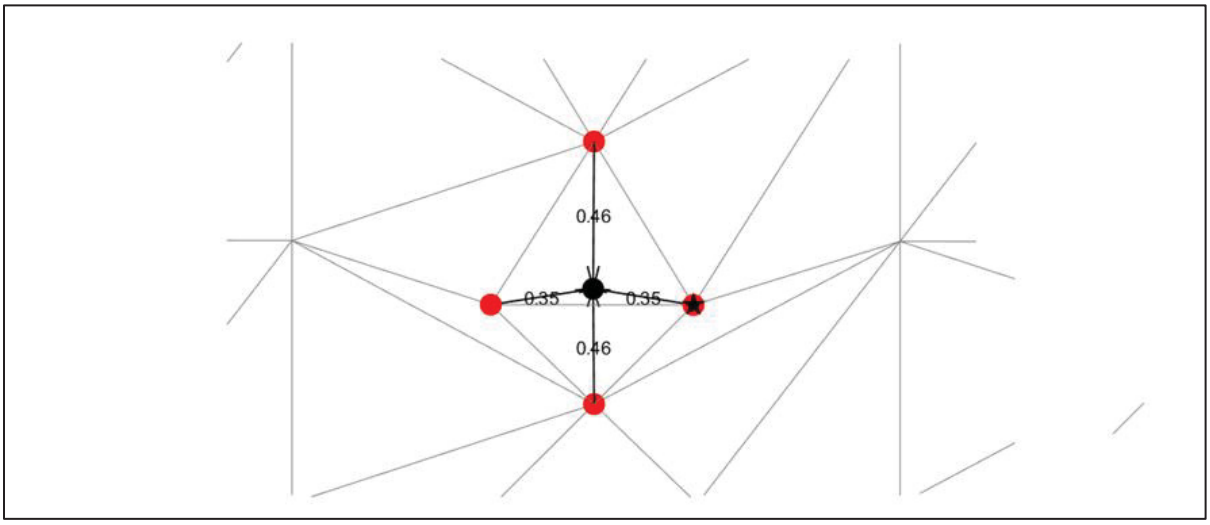

Figure 4.11 Multi-matching on the source mesh (red vertices) and choosing the closest vertex (marked with black asterisk) to the center of the candidate vertices (black dot)

4.3.1.2 Joint Modeling

The dense mesh generation, the simplification of the mesh, and matching the meshes, have not modified the position of the vertices with respect to the CT scan masks. So, the Tibia and Fibula meshes in each case are still in the same relative positioning as in the CT scan masks.

Thus, after obtaining tibia and fibula meshes with the same topology, we concatenate the meshes to form the joint mesh. It is achieved by concatenating $tibia_i^V$ and $fibula_i^V$, as well as $tibia_i^F$ and $fibula_i^F$ for $i = 1..53$ as shown in Equation (4.1). The results are stored in $joint_i^V$ and $joint_i^F$, which together form the $joint_i^{mesh}, i = 1..53$. The vertices matrix consists of vertices' coordinates. Thus, a simple concatenation can make the $joint_i^V$. However, the faces matrix is consists of M rows dedicated to M faces, indicating the vertex indices for each triangle. Thus, for faces we need to add the number of tibia vertices to all indices in the $fibula_i^F$ to complete the concatenation. This model captures the relationship between the tibia and fibula as well.

$$\begin{aligned} joint_i^{mesh} &= \{joint_i^V, joint_i^F\}, i = 1..53 \\ joint_i^V &= [tibia_i^V; fibula_i^V], i = 1..53 \\ joint_i^F &= [tibia_i^F; [fibula_i^F + number\_of\_tibia\_vertices]], i = 1..53 \end{aligned} \tag{4.1}$$

4.3.1.3 Generalized Procrustes Analysis

Although a correspondence mapping is calculated for the 53 meshes in the dataset and they share the same topology, they are not aligned in a common coordinate system. Figure 4.12 shows the joint meshes having different positions and orientations. We applied Generalized Procrustes Analysis (GPA) to align all 53 meshes (Figure 4.13.A) by gradually removing translation, rotation, and scale (

Algorithm 4.2). The mean shape of the aligned joint meshes $joint_i^{mesh^{aligned}}, i = 1..53$ is presented in Figure 4.13.B.

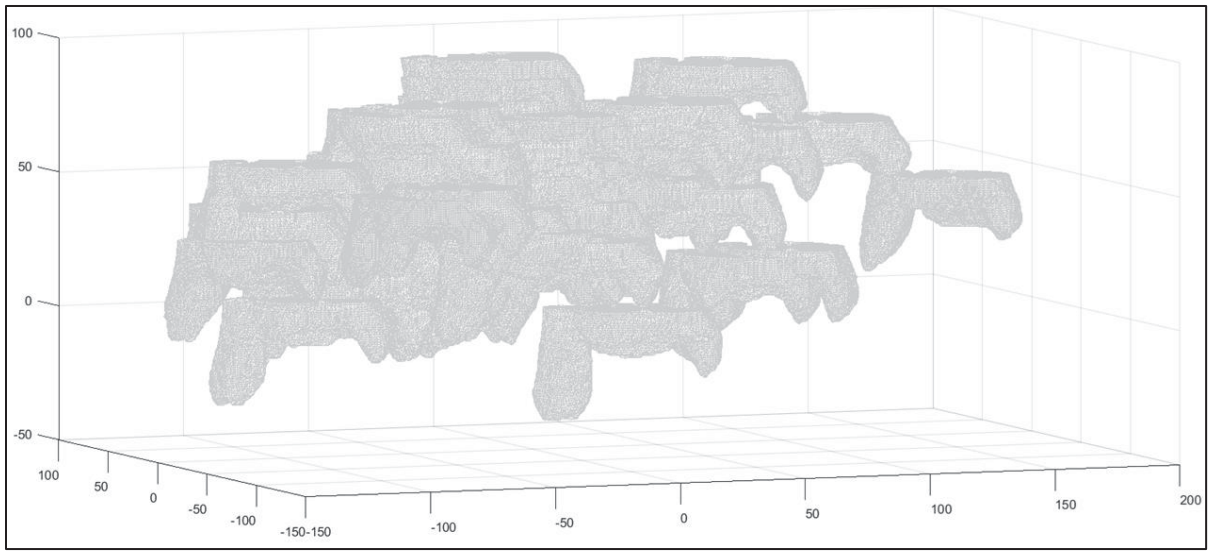

Figure 4.12 Joint meshes before applying GPA

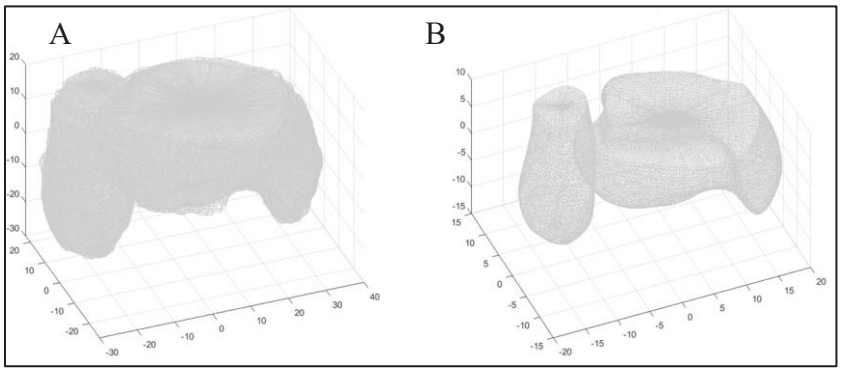

Figure 4.13 All joint meshes while the rotation, translation and scale are removed (A) and their mean shape (B)

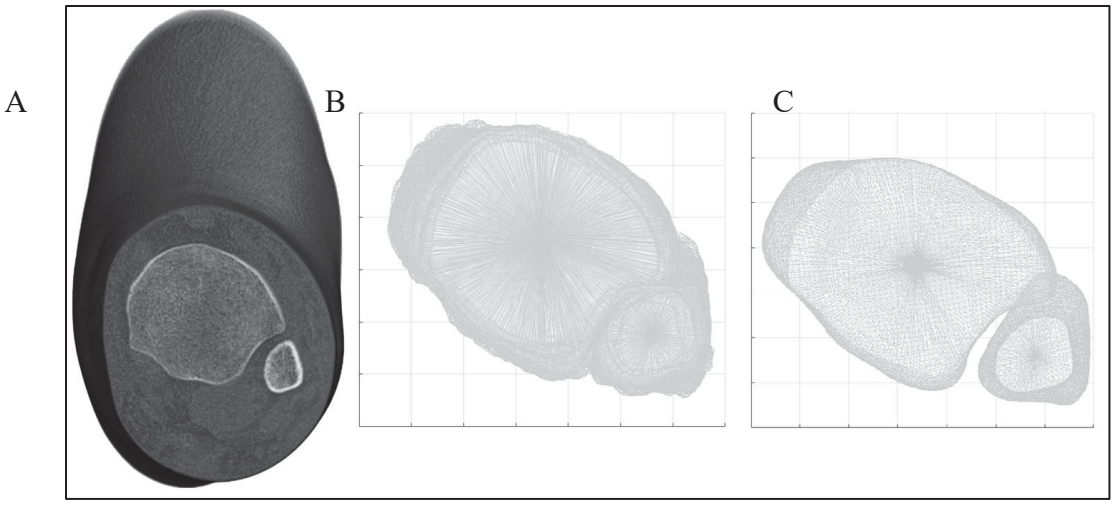

Figure 4.14 Aligning the CT segmented volumes (A) orientation and their corresponding meshes (B) with the anterior-posterior positioning of the foot, similar to the orientation captured in EOS cabin. The mean shape (C) is recalculated on the rotated meshes.

Algorithm 4.2 Apply GPA on meshes and make them in the AP direction

Apply GPA on meshes and make them in the AP direction

Input: Tibiofibular joint meshes $joint_i^V, i = 1..53$

Output: Aligned joint meshes $joint_i^{V^{aligned}}, i = 1..53$

1 Randomly select a mesh as the reference mesh $joint_{ref}^V$

2 $joint_{ref}^V$ is translated to the center of origin and normalized to the unit size

3 **While** $\left|joint_{ref}^V - joint_{mean}^V\right| < \boldsymbol{\varepsilon}$

4 Translate $joint_i^V, i = 1..53$ to $joint_{ref}^V$

5 Normalize $joint_i^V, i = 1..53$ to have unit size

6 Rotate $joint_i^V, i = 1..53$ by Rotation matrix R_i defined by Singular

7 Value Decomposition

8 Compute the $joint_{mean}^V$

9 Scale back the $joint_i^{V^{aligned}}, i = 1..53$ to the average size of the original meshes

10 Rotate $joint_i^{mesh^{aligned}}, i = 1..53$ to the AP direction $joint_i^{mesh^{EOS}}, i = 1..53$

As meshes are employed in the reconstruction process, it is essential that they initially assume a pose similar to the joint observed in the EOS Imaging cabin. As shown in Figure 4.14.A, the foot is positioned in an anterior-posterior orientation, which serves as a reference for alignment. To ensure consistency across the dataset, all meshes are rotated to approximate this pose (Figure 4.14.B). Following this alignment step, the mean mesh is recalculated based on the transformed meshes $joint_i^{mesh^{EOS}}, i = 1..53$ to reflect the updated orientation on shapes (Figure 4.14.C).
Algorithm 4.2 also describes this step.

4.3.1.4 Principal Component Analysis

After aligning the meshes, Principal Component Analysis (PCA) is applied to compactly represent the model. Having all meshes aligned, we calculate the mean mesh $joint_{mean}^{mesh^{EOS}}$ using Equation (4.2).

$$joint_{mean}^{mesh^{EOS}} = \frac{1}{50}\sum_{i=1}^{50} joint_i^{mesh^{EOS}} \tag{4.2}$$

Considering the matrix $Y = [joint_1^{mesh^{EOS}}, joint_2^{mesh^{EOS}}, \dots, joint_{50}^{mesh^{EOS}}] \in R^{3K\times N}$ including all 50 training meshes, we calculated the covariance matrix as in Equation (4.3) and finally, the eigenvectors $\Phi = [\psi_1; \psi_2; \dots; \psi_{N-1}]$ and eigenvalues $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{N-1}]$ are determined using Singular Value Decomposition (SVD), elaborated in Equation (4.4).

$$C = \frac{1}{N-1} YY^T \in R^{3K\times 3K} \tag{4.3}$$

$$C\Phi = \Lambda\Phi \tag{4.4}$$

The joint model is described by the mean shape $joint_{mean}^{mesh^{EOS}}$, along with a combination of eigenvectors ψ_k and shape parameters b_k, outlined in Equation (4.5). The k_max is determined based on the eigenvalues of the eigenvectors, selecting enough components to represent 90% of the variation in the dataset, balancing between capturing sufficient anatomical detail and reducing the model's dimensionality. Shape parameters b_k follow a normal distribution bounded by $\pm 3\sqrt{\lambda_k}$ as explained in Section 2.2.2.2.

$$SSM(Joint) = joint_{mean}^{mesh^{EOS}} + \sum_{k=1}^{k\_max} b_k . \psi_k \tag{4.5}$$

4.3.2 Intensity Modeling

Given the complexity and slow performance associated with intensity modeling using tetrahedral meshing, especially when it comes to repetitive operations, we opted for a simpler approach, similar to modeling cortical and trabecular bones with a single intensity (Julien et al., 2021), to represent the intensity information of the tibia and fibula. Instead of using a detailed tetrahedral model, we decided to represent the tibial intensity using the average intensity value of all tibias in the dataset, applying the same approach to the fibula. The intensity values for the bones are collected using the segmentation mask and the CT scan slices, recording the voxel values. Upon analyzing the histograms of the tibia and fibula intensities (Figure 4.15), we observed that the average intensity values for both bones are very similar. As a result, we simplified the model further by using a single intensity I^{simple} value for both bones. This simplified modeling approach significantly improves the efficiency of iterative processes, such as dataset generation, where the intensity modeling and projection need to be repeated multiple times.

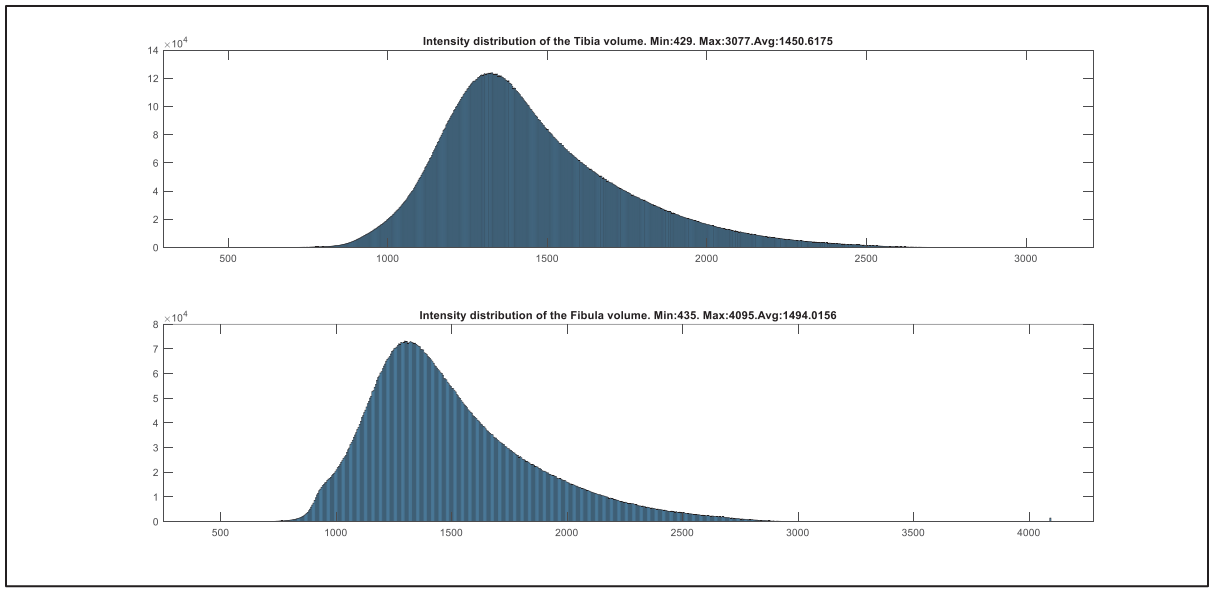

Figure 4.15 Intensity value histogram of Tibia and Fibula volumes for all 53 cases

4.3.3 Statistical Shape and Intensity Modeling

The Statistical Shape and Intensity Model (SSIM) of the tibiofibular joint is constructed using an independent modeling approach, where the shape and intensity components are represented separately. The shape model captures geometric variability through principal shape parameters $b_k, k = 1..k\_max$, while the intensity distribution is modeled by a simplified representation denoted as I^{simple}. As a result, the joint model is defined by the mean shape $joint_{mean}^{mesh^{EOS}}$, the set of shape coefficients b_k, the eigenvector matrix of the shapes Φ, and the associated intensity information I^{simple}, as established in Equation (4.6).

$$SSIM(Joint) = \{joint_{mean}^{mesh^{EOS}}, b_k, \psi_k, I^{simple}\} \tag{4.6}$$

4.4 One-shot 2D/3D Reconstruction of the Distal Tibiofibular Joint from Biplanar Radiographs using Deep Learning Registration

A reconstruction network is trained to take in biplanar X-rays of the tibiofibular joint and output the shape and pose parameters of the SSIM. These parameters enable us to construct a

model that aligns with the input radiographs. The network performs 2D/3D registration in a single step. The following sections describe the network architecture and provide details on the training process.

4.4.1 2D/3D Deep-Learning based Registration Network

The registration network takes the AP and LT images of the joint, concatenated into a single AP-LT image, and outputs the joint model's shape parameters ($b_k, k = 1..k\_max$) along with pose parameters, including 3 rotational parameters around axes X, Y, and Z (rot_x, rot_y, rot_z) and 1 scale parameter (S). The joint is assumed to be centered in the radiographic images. Therefore, translation parameters are excluded from the pose estimation process. This regression network is based on the VGG-16 architecture (Simonyan & Zisserman, 2014), with the soft-max activation function replaced by ReLU in the final layer. The network is trained using the Mean Squared Error (MSE) loss, comparing the shape and pose parameters to the ground truth. Equation (4.7) presents the MSE loss. To prevent overfitting, dropout layers are added to the fully connected layers. Stochastic Gradient Descent (SGD) is employed, consistent with the conventional VGG architecture (Simonyan & Zisserman, 2014) . Figure 4.16 depicts the network architecture.

$$MSE = \frac{1}{k\_max + 4}\Big(\sum_{k=1}^{k\_max} (b_k - \hat{b}_k)^2 + (rot_x - \widehat{rot}_x)^2 + (rot_y - \widehat{rot}_y)^2 + (rot_z - \widehat{rot}_z)^2 + (S - \hat{S})^2\Big) \tag{4.7}$$

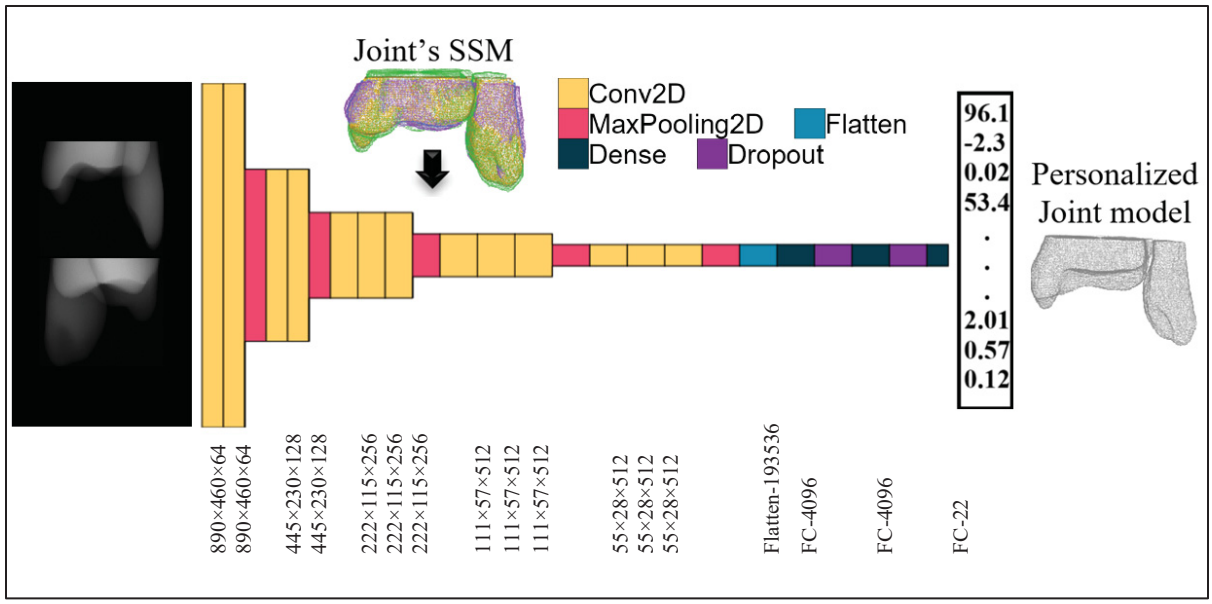

Figure 4.16 2D/3D registration network architecture, trained on the Joint's SSM parameters and pose parameters

Initially, the learning rate is set to 0.1, the batch size is 12, and the dropout rate is 0.2. The initial values are taken from the VGG-architecture implementation (Simonyan & Zisserman, 2014). A grid search (Liashchynskyi & Liashchynskyi, 2019) is performed to find the optimized combination of the hyperparameter, learning rate, dropout rate and batch size. The learning rate is controlled by a scheduler that halves the current learning rate if the validation loss does not improve over 10 consecutive epochs. The minimum learning rate, below which training stops, is set to $1e^{-6}$, making sure is it set low enough that the convergence is completed. Additionally, an early stopping mechanism is employed to halt training if the validation loss fails to decrease over 15 epochs, making sure the training process is converged even after decreasing the learning rate after 10 epochs. The best network parameters are returned based on the lowest validation loss seen. The evaluation of the network is conducted by monitoring the Mean Squared Error (MSE) as in Equation (4.7). However, the generated mesh using the network output parameters can be compared against the mesh generated from the ground truth values by point-to-point Euclidean error (explained in Section 2.4.2) in mm in 3D space. Equation (4.8) calculated this error. The latter metric is preferred, as we aim to reconstruct the

3D model, and it is feasible since both the reconstructed model and the ground truth meshes have the same topology.

$$P2P = \frac{1}{N}\sum_{i=1}^{N}\left\|joint^{V}_{Reconstructed} - joint^{V}_{Ground\_Truth}\right\|^{2} \tag{4.8}$$

4.4.2 Generation of the Dataset for 2D/3D registration network

The dataset is generated based on the SSIM of the joint, as described in Section 4.3.1. New instances are created by randomly selecting the parameter b_k from a normal distribution $N(0, \lambda_k)$ where λ_k is the k-th eigenvalue of the model. Although b_k is drawn from this distribution, it is restricted to the range $(-3 \times \sqrt{\lambda_k}, 3 \times \sqrt{\lambda_k})$ for $k = 1..k\_max$. Once the instance is generated from the model, it is rotated and scaled to create different poses of the instance. The rotational parameters rot_x, rot_y, rot_z and the scale parameter S are stored alongside the shape parameters $b_k, k = 1..k\_max$ for the generated instance. This is done on all generated shapes instead of the first mesh which is generated using all-zero shape parameters, with no rotation or scaling applied. So, the first instance represents the mean shape of the joint.

After the watertight mesh is generated, a ray intersection method (S. Patil & Ravi, 2005) is used to fill the mesh structure in 3D with the intensity value I^{simple}, creating the volume. The volume is treated as a CT scan, which is further processed by the Voxel Projector tool to generate DRRs, the AP and LT radiographs. Algorithm 4.3 explains these steps.

Algorithm 4.3 Generate Training dataset for 2D/3D registration Network

Generate Training dataset for 2D/3D registration Network

Input: SSIM of the tibiofibular Joint, rotation and scale parameters bound

Output: N AP and LT radiographs generated along with their shape and pose parameters

1 **While** $i < N$

2 Choose b_k from a normal distribution bounded by $\pm 3\sqrt{\lambda_k}, k = 1..k\_max$

3 Generate new instance $joint_i^{mesh^{DB}} = joint_{mean}^{mesh^{EOS}} + \sum_{k=1}^{k\_max} b_k^i . \psi_k$

4 Choose rot_x, rot_y, rot_z, s from a uniform distribution

5 Apply rotation and scaling to $joint_i^{mesh^{DB}}$ by rot_x, rot_y, rot_z, s

6 Fill the $joint_i^{mesh^{DB}}$ by I^{simple} and generate the volume $joint_i^{vol^{DB}}$

7 Apply Voxel Projector to $joint_i^{vol^{DB}}$ to have the AP_i^{DB} and LT_i^{DB} images

8 Store AP_i^{DB} and LT_i^{DB} along with $rot_x, rot_y, rot_z, s\ b_k^i, k = 1..k\_max$

To eliminate the impact of different scales in the shape and pose parameters, each parameter $b_k, k = 1..k\_max$ as well as rot_x, rot_y, rot_z, and S are normalized to the range [-1, 1], based on the maximum value of each respective parameter. Equation (4.9) depicts the normalization of the kth shape parameters from ith sample in the dataset. The maximum values of each parameter ($\max(b_k)$) among N samples are used to rescale the normalized values back to their original range.

$$b_k^{i\ normalized} = 2 \times \frac{b_k^i - \min(b_k)}{\max(b_k) - \min(b_k)} - 1, k = 1..k_{max}, i = 1..N \tag{4.9}$$

4.4.2.1 Bounds of the Parameters

We imposed the same limitations on the parameters as described for the JE-Nets in section 4.5.2.1. Specifically, the rotations around the X and Y axes are restricted to ±5 degrees, while the rotation around the Z axis is limited to ±20 degrees. These parameters are experimentally chosen to imitate the variability of the rotations in the EOS cabin. The scale parameter is

limited to ±0.17, as the scale variation observed across all 53 cases in the dataset is approximately 17%. The selection of the parameters is performed using a uniform random distribution.

For the initial cases in the dataset, the mesh is generated with $b_k, k = 1..k\_max$ set to zero, and no rotation or scaling is applied. This creates the mean shape without any pose alterations. Afterward, all meshes are randomly generated as described in the previous section. As a result, the final dataset consists of 30,000 pairs of AP-LT image along with its corresponding shape and pose parameter values. Out of this total, 10% (3,000 pairs of images) are reserved for testing, while the remaining 90% is split into training and validation sets (90% for training and 10% for validation).

4.4.2.2 Post-processing of the Digitally Reconstructed Radiographs

After the images are generated using the Voxel Projector tool, they undergo preprocessing to create a normalized dataset before training the networks, e.g., having the same size for all images or normalized intensity values on the dataset. Figure 4.17 shows the preprocessing steps applied to the original DRRs (Figure 4.17.A). The first step involves applying a simple thresholding process on the DRRs (Figure 4.17.B) to isolate the joint using the vertical and horizontal projection profile method, cropping the image to its smallest possible size (Figure 4.17.C). This step is performed for both the AP and LT images. Next, a cropping process removes the top safety margin (explained in Section 4.2.1), while the bottom safety margin is already removed using the previous step (Figure 4.17.D). The resulting AP and LT images are then concatenated vertically (Figure 4.17.E), with a gap of 5 lines of pixels between the two. Following this, the largest AP-LT image in the dataset is used as a reference, and all other images are padded to match its size. The padded areas are filled with the minimum pixel value from the image being processed. Finally, the pixel intensities of all images are normalized between 0 and 1 using the maximum intensity value throughout the dataset. This helps maintain a consistent bone appearance through the dataset. Figure 4.18 shows sample final images from the generated dataset.

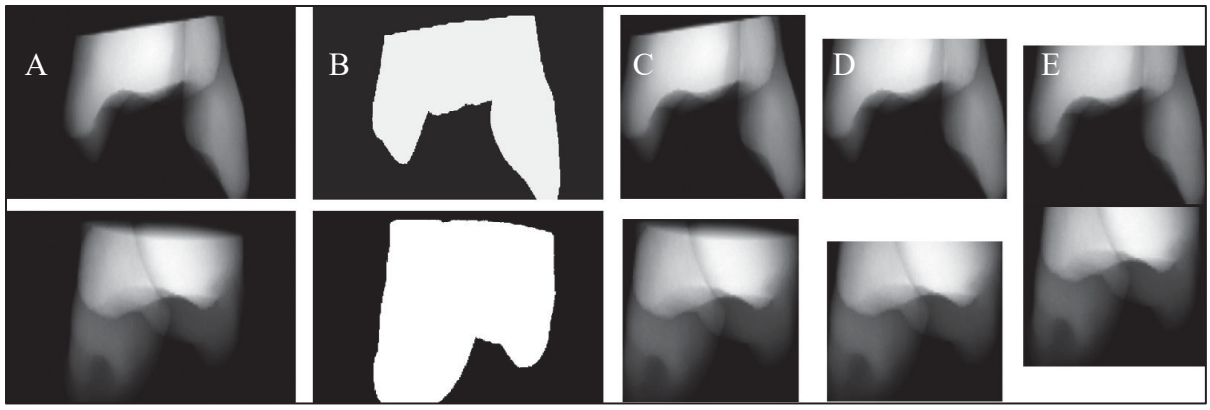

Figure 4.17 preprocessing steps of generated DRRs for Reconstruction Network

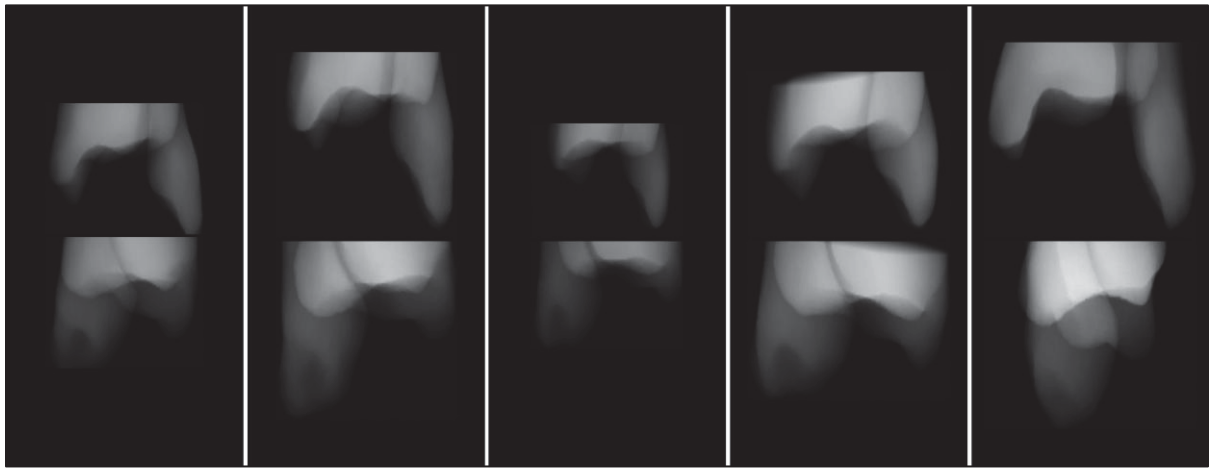

Figure 4.18 AP-LT cases generated to train the registration network

4.5 Isolation of the Tibiofibular Joint skeletal structure in biplanar radiographs

Since the 2D/3D reconstruction network is trained on DRRs that represent only the tibiofibular joint, and real ankle radiographs contain additional skeletal structures and soft tissues, it is essential to isolate the joint by removing surrounding anatomical components. This preprocessing step ensures that the input ankle radiographs closely resemble the training data, thereby improving the compatibility and performance of the reconstruction network. These networks, referred to as Joint Extraction Networks (JE-Net), are designed specifically for this

purpose. The following sections describe the structure of these networks, and the training process involved.

4.5.1 Joint Extraction Deep-Learning based Networks

We utilize two separate JE-Nets for the AP and LT images, as their input sizes differ. These networks are inspired by the pixel-wise structure removal technique for extracting bones from chest X-rays (Kalisz & Marczyk, 2021) which uses a denoising autoencoder architecture. The network architecture is shown in Figure 4.19. In addition to the original network architecture, we have incorporated additional dropout layers to prevent overfitting of the network.

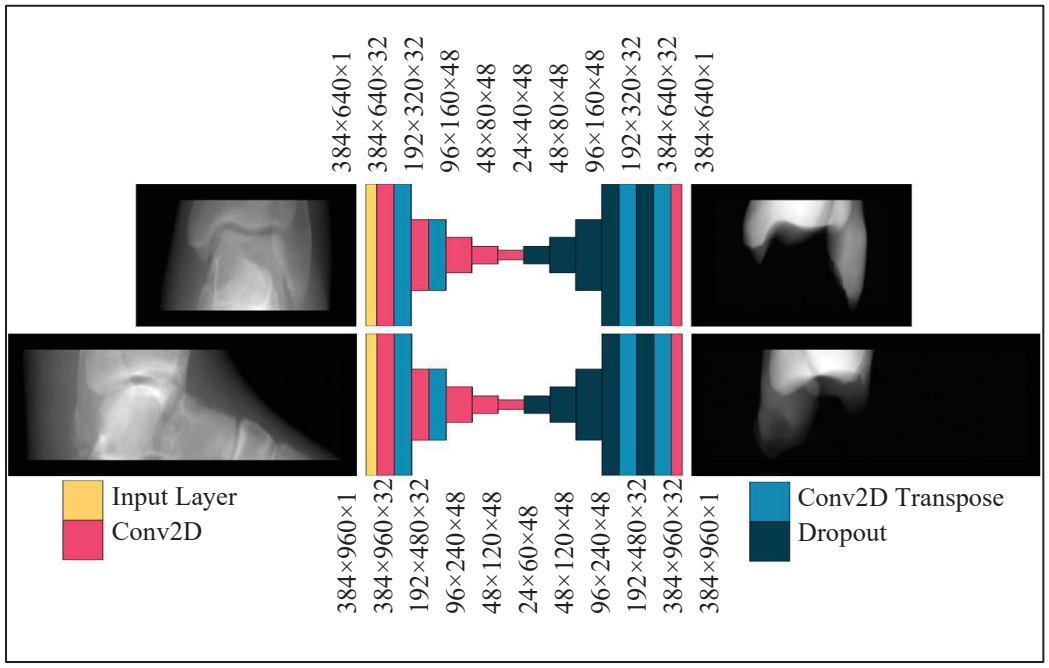

Figure 4.19 Architecture of the Joint Extraction Networks (JE-Net) for AP and LT images

Both networks process the radiographs of the ankle in AP and LT views to extract the tibiofibular joint, effectively removing soft tissues and other skeletal structures. The loss function for the network is defined as a combination of the Structural Similarity Index Measure (SSiM) and Mean Squared Error (MSE) between the output and the ground truth as proposed on the original work (Kalisz & Marczyk, 2021). The relative weighting of these two

components is controlled by the parameter α. The loss function for this network is defined in Equation (4.10).

$$loss = \alpha \times (1 - SSiM(y_{true}, y_{pred})) + (1 - \alpha) \times MSE(y_{true}, y_{pred}) \quad (4.10)$$

The Adam optimizer algorithm is used to train the networks. For the network's hyperparameters, we start with the values provided in the original work (Kalisz & Marczyk, 2021) and a grid search (Liashchynskyi & Liashchynskyi, 2019) is applied for hyperparameter optimization. Initially, α is set to 0.9, the learning rate is 0.001, the batch size is 8, and the dropout rate is 0.3. The learning rate is managed by a scheduler that halves the current learning rate if the validation loss does not improve over 10 consecutive epochs. The minimum learning rate, at which training stops, is set to $1e^{-8}$. Additionally, an early stopping mechanism halts the training if the validation loss fails to decrease over 15 epochs, where the best weights (from the best epoch with the lowest validation loss) are restored.

4.5.2 Generation of the Dataset for Joint Isolation Networks

The dataset used to train the JE-Nets is generated from the segmented ankle volumes and their corresponding segmented joint volumes. However, the segmented joint volumes are modified to align with the representation used in the intensity model. Specifically, all voxel values within the joint volume are adjusted to match the average intensity value determined in the dataset, as described in Section 4.4.2. In addition, since the dataset must represent the ankle volumes in the AP orientation, the first step is to align the segmented ankle volumes $ankle_i$ and the segmented joint volumes $joint_i^{vol}$ with the EOS-oriented joint meshes $joint_i^{mesh^{EOS}}, i = 1..53$. This is achieved by rigidly aligning the joint mesh $joint_i^{mesh}$ to $joint_i^{mesh^{EOS}}$ to determine the rotation and translation parameters. These rotational parameters are then applied to the volumes $ankle_i$ and $joint_i^{vol}$.

After the volumes are aligned to the EOS anterior-posterior orientation, we used Voxel Projector (EOS imaging, n.d.) to simulate the EOS biplanar radiographs from the volumes.

Similar to the creation of the Statistical Shape Model (SSM), three instances are excluded from the 53 cases in the dataset, which are used for evaluation purposes. By applying projection to pairs of $ankle_i^{EOS}$ and $joint_i^{vol^{EOS}}, i = 1..50$ we generate an AP image of the ankle and its corresponding AP image of the joint. The same procedure is applied to generate LT images. Figure 4.20 presents a sample of data produced by the voxel projector from a random volume in the dataset. Algorithm 4.4 explains the steps for generating the training dataset.

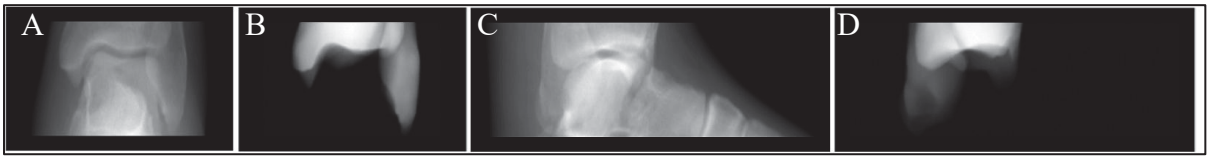

Figure 4.20 Sample images for training JE-Nets. Input of the JE-Net-AP (A), desired output of the JE-Net-AP (B), Input of the JE-Net-LT (C) and desired output of the JE-Net-LT (D).

After generation of the first 50 cases without any rotation applied to the volumes, and since we only have 50 volumes, a data augmentation (Kumar et al., 2025) process is applied to generate additional data for training the networks. A random volume is selected from the 50 cases, and both the ankle and joint volumes are rotated using rot_x, rot_y, rot_z around the X, Y, and Z axes, respectively. The projection process is then applied to the rotated volumes. Although the volumes are rigidly rotated with no deformation applied to the joint or ankle, this augmentation generates different images that help generalize the training process for the networks. Translation is not applied to the volumes as the joint is always assumed to be centered in all radiographs, making any translation be removed from images. Furthermore, since the CT scans represent different sizes of ankles and joints, scale is inherently represented in the dataset. The evaluation of the networks is carried out by monitoring the loss value on the test set and calculating the Peak Signal-to-Noise Ratio (PSNR) between the output images and the ground truth.

Algorithm 4.4 Generate Training dataset for Joint Extraction Networks

Generate Training dataset for Joint Extraction Networks

Input: 50 Segmented CT-scans $ankle_i$ and $join_i^{vol}, i = 1..50$

Output: N AP and LT radiographs of Ankle and its corresponding Joint

Rotate the $ankle_i, joint_i^{vol}, i = 1..50$ to align the $joint_i^{mesh^{EOS}}, i = 1..53$

While $k < N$

    Choose a volume $ankle_i$ from the dataset using uniform distribution

    Choose rot_x, rot_y, rot_z from a uniform distribution

    Apply rotation to $ankle_i$ and $joint_i^{vol}$ by rot_x, rot_y, rot_z

    Apply Voxel Projector to $ankle_i$ and $joint_i^{vol}$ to have the AP_{ankle}^{DB}, AP_{joint}^{DB}, LT_{ankle}^{DB} and LT_{joint}^{DB} radiographs

    Store AP_{ankle}^{DB}, AP_{joint}^{DB}, LT_{ankle}^{DB} and LT_{joint}^{DB}, rot_x, rot_y, rot_z and i

4.5.2.1 Bounds of the Parameters

We imposed specific limitations on the parameters used to generate the dataset. The rotations around the X and Y axes are restricted to ±5 degrees, while the rotation around the Z axis is limited to ±20 degrees. These values are empirically selected to present the volume with the same pose as we may face in the EOS cabin. For the first 50 cases in the dataset, the volumes are projected without any rotation. After that, for each subsequent case, a random volume is selected, and a random rotational degree is applied to each axis within the specified range. The selection of the parameters is performed using a uniform random distribution. As a result, the final dataset consists of 15,000 pairs of images (for each network). Out of this total, 10% (1,500 pairs of images) are set aside for testing, while the remaining 90% is split into training and validation sets (90% for training and 10% for validation).

4.5.2.2 Post-processing of the Digitally Reconstructed Radiographs

After the images are generated using the Voxel Projector tool, they undergo preprocessing to create a normalized dataset before training the networks e.g., having the same size for all images or normalized intensity values in the dataset. The preprocessing is applied on both AP and LT images of ankle (Figure 4.21 A, B) and joint (Figure 4.21 C, D). The first step involves applying a simple thresholding process to isolate the ankle (Figure 4.21 E, F), cropping the AP and LT images of the ankle according to the resulting masks (Figure 4.21 G,H). The same masks are used for cropping joint images (Figure 4.21 I, J). Next, a cropping process removes the safety margin at the top and bottom of the images for both the ankle (Figure 4.21 K, L) and joint images (Figure 4.21 M, N). Following this, the largest images in the dataset are used as a reference, and all other images are padded to match the size of the largest image. The padded areas are filled with the minimum pixel values from the image. Finally, the pixel intensities of all images are normalized between 0 and 1. In some cases, the dataset contains images where medical treatments, such as the insertion of metal pins into the talus or calcaneus, result in distortions and high-intensity pixels (Figure 4.22). To avoid these cases affecting the normalization, such images are excluded when determining the maximum and minimum intensity values for the dataset. This ensures that the pixels in these areas are saturated, and the rest of the image is appropriately normalized.

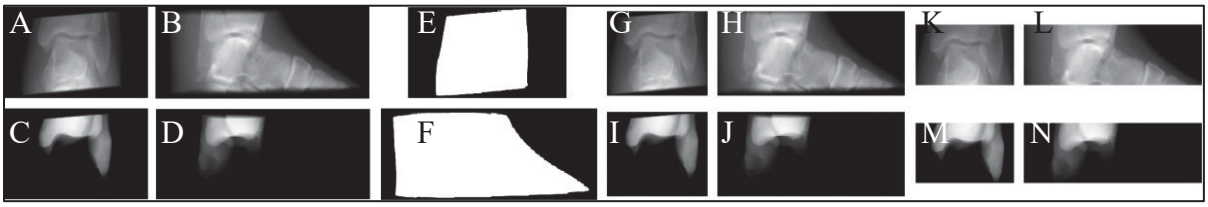

Figure 4.21 Preprocessing steps of generated DRRs for Joint Extraction Networks

Figure 4.22 Sample from the dataset, showing treatment and its effects on the pixel intensities

4.6 Region-of-Interest Selection in Radiographs

The biplanar radiographs are captured with a wider scope than the Region of Interest (ROI) needed for DTF joint reconstruction. Therefore, it is essential to accurately select the specific region corresponding to the relevant part of the DTF joint. The ROI selection begins by manually identifying two anatomical landmarks on the AP EOS image of the ankle: the Tibial Plafond (TP) (Figure 4.23.A.1) and the Distal Fibula (DF) (Figure 4.23.A.2). These landmarks are then used to define the superior and inferior crop lines on the AP image. Selection of the landmarks are reviewed and verified by a medical professional to ensure accuracy and reliability.

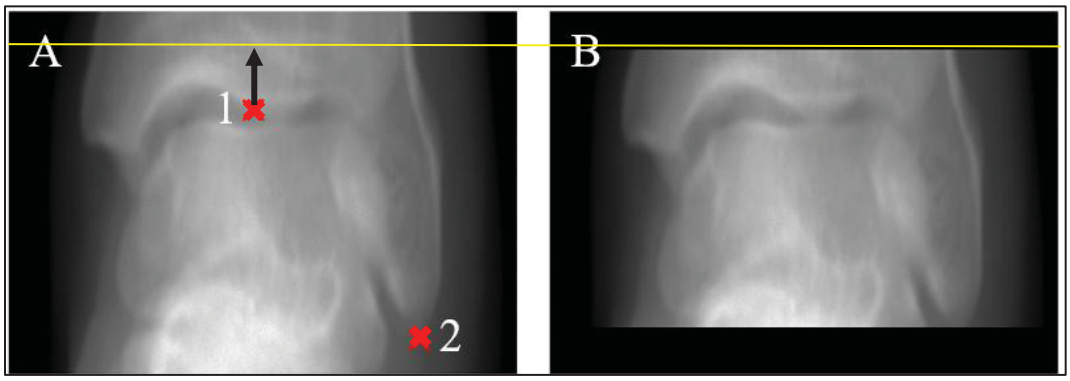

Figure 4.23 Selection of the anatomical landmarks TP (A.1) and DF (A.2) and selecting the ROI based on those landmarks (B)

The Region of Interest (ROI) for the reconstruction process begins 46 pixels above the Tibial Plafond landmark, identified on AP EOS image, which corresponds to 15 slices above the identified Tibial Plafond slice in the CT volume. This is calculated based on the CT slice thickness of 0.625 mm, the slice spacing of 0.5 mm, and the pixel spacing of 0.164 mm in the EOS images. The detailed calculation is provided in Equation (4.11). Figure 4.23.B illustrates the result of this cropping. Since both the AP and LT images are calibrated in EOS, the same crop lines can be applied to the LT image as well.

$$\frac{0.625}{2} + (15 - 1) \times 0.5 + \frac{0.625}{2} = 7.625mm; \frac{7.625}{0.164} = 46.49 \sim 46px \tag{4.11}$$

In addition to establishing the superior and inferior boundaries, it is also essential to determine the left and right lateral crop lines to isolate the ankle within the ROI. These boundaries are automatically identified using a simple thresholding technique applied to both the AP and LT ankle images, with the threshold value determined experimentally.

4.6.1 Simulation of the ROI selection

Since the ROI selection relies on the manual identification of landmarks by the operator, we need to simulate this process for our experiments. The simulation of ROI selection is performed using the segmented joint. Given that the $ankle_i$ and $joint_i$ are in correspondence, the segmented $joint_i$ can be used to detect the superior and inferior crop lines on AP image of $ankle_i$. As the joint is segmented 30 slices above the tibial plafond slice in the CT volume, we can match the same anatomical landmarks as the manual selection by moving 15 slices from the most superior segmented slice toward the distal fibula, aligning it with the relevant slice for the ROI (since 30 slices above the tibial plafond are segmented). These 15 slices in the CT volume correspond to 46 pixels in AP EOS image (Figure 4.24.A) as calculated in Equation (4.1). First, we binarize the AP image of the joint (Figure 4.24.B) using simple thresholding (Figure 4.24.C) with the threshold value determined experimentally. Next, the vertical projection profile method is applied to the binary image. By moving 46 pixels from the most superior boundary toward the distal fibula, the top boundary is defined (Figure 4.24.C), while the most inferior boundary is also determined using the vertical projection profile. The crop lines for the LT image are identical to the lines of the AP image. The left and right lateral boundaries are selected as outlined in the previous section. Figure 4.24 illustrates the automatic selection of the ROI based on the joint segmentation.

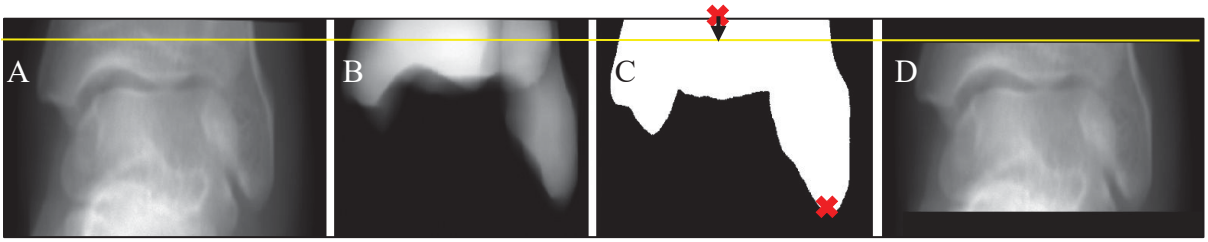

Figure 4.24 Simulation of the ROI selection of the ankle (A) using the segmented joint (B) and its binary image (C) using simple thresholding. The result ROI (D)

4.6.2 Post-processing of the Radiographs

After ROI selection of the radiographs to isolate the ankle for reconstruction process, an additional round of post-processing is applied to prepare the images for the joint isolation process. This step involves padding the images to match the input size of the Joint Isolation networks and normalizing their intensity based on the maximum and minimum values from the dataset used to train the network.

First, image padding is applied until the image reaches the required input size for the network. Since we have separate joint isolation networks for the AP and LT images, each network has its own specific input size. The intensity of the padded regions is set to the minimum intensity value of the input image. Next, the pixel intensities of the image are normalized between 0 and 1 based on the maximum and minimum intensity values specific to each network. A final pair of processed AP and LT ankle images, ready for the next step, is shown in Figure 4.25.

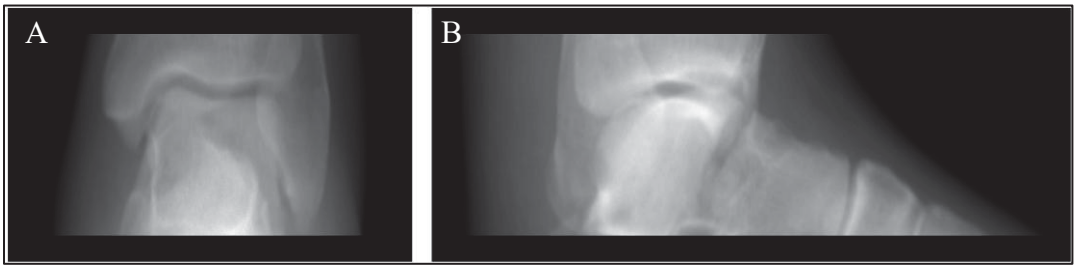

Figure 4.25 a pair of AP (A) and LT (B) images of the ankle preprocessed for the Joint Extraction Network

4.7 Establishing the 2D/3D Reconstruction Pipeline

After all networks are trained, we establish the pipeline to receive the AP and LT biplanar radiographs of a patient's lower limb and generate a personalized 3D model of their tibiofibular joint. As depicted in Figure 4.1, first, the ROI is selected from the biplanar images. The ROI which isolates the ankle is then processed by JE-Nets to extract the joint. The ROI selection is simulated automatically as described in Section 4.6.1. After concatenation of the AP and LT

images of the extracted joint, the 2D/3D registration network reveals the shape and pose parameters, constructing the personalized joint model in 3D space.

The Point-to-Surface (P2S) error metric is used to evaluate the results of the pipeline. The personalized 3D model of the joint, generated from the ankle biplanar radiographs, is compared against the dense mesh obtained from the joint's CT scan segmentation masks to compute the P2S error in mm in 3D space.

In the following section, we describe the generation of the evaluation set for evaluating the pipeline.

4.7.1 Generation of the Dataset for Pipeline Evaluation

The evaluation dataset is generated based on the three left-out instances. The three left-out volumes $ankle_i^{EOS}$ and $joint_i^{vol^{EOS}}, i = 1..3$ are used to generate evaluation instances. A data augmentation process similar to the one explained in Section 4.5.2 is conducted to augment the evaluation data to 1,000 instances. The rotations around the X and Y axes are restricted to ±5 degrees, while the rotation around the Z axis is limited to ±15 degrees. For the first 3 cases in the dataset, the volumes are projected without any rotation. After that, for each subsequent case, a random volume is selected, and a random rotational degree is applied to each axis within the specified range. The selection of the parameters is performed using a uniform random distribution. Different scales are already present in the dataset, no additional scaling is applied. Algorithm 4.5 explains these steps.

Since the selection of the ROI is currently simulated automatically but will eventually be done manually in future applications, we simulated the reproducibility of the operator by adding uniformly distributed noise to the selection of the most superior and most inferior crop lines. The most superior crop line, which is defined based on the tibial plafond, may shift toward the distal fibula, as the dataset does not provide slices on the opposite side. The maximum shift allowed is 12 pixels, which corresponds to approximately 3 mm in the EOS images.

For the most inferior crop line, the movement is allowed in both directions, with the crop line potentially shifting toward the knee or the heel bone. Again, the maximum shift is 12 pixels, but we can also apply negative values, which would move the crop line toward the most

superior boundary. By introducing this noise into the ROI selection process, we generated an additional 1,000 instances, as described earlier, and for each instance, two random noise values are selected. This allows us to assess the impact of noisy ROI selection on the P2S reconstruction error.

Algorithm 4.5 Generate Training dataset for Pipeline Evaluation

Generate Training dataset for Pipeline Evaluation

Input: 3 Segmented CT-scans $ankle_i$

Output: N AP and LT radiographs of Ankle

Rotate the $ankle_i, i = 1..3$ to align the $joint_i^{mesh^{EOS}}, i = 1..3$

While $k < N$

    Choose a volume $ankle_i$ from the dataset using uniform distribution

    Choose rot_x, rot_y, rot_z from a uniform distribution

    Apply rotation to $ankle_i$ by rot_x, rot_y, rot_z

    Apply Voxel Projector to $ankle_i$ to have the AP_{ankle}^{DB}, and LT_{ankle}^{DB} radiographs

    Store AP_{ankle}^{DB}, LT_{ankle}^{DB}, rot_x, rot_y, rot_z and i

4.8 Syndesmosis Clinical Parameters Assessment on 3D model

After generating the personalized 3D model of the tibiofibular joint from ankle biplanar radiographs, we calculate the syndesmosis clinical parameters. The model is presented in a 3D space within the real coordinate system in mm, enabling the calculation of clinical parameters in 3D. Moreover, since the personalized model is constructed with the same topology as our statistical shape model, the identification of key landmarks on the mean shape, which are required for the calculation of clinical parameters facilitates the automatic computation of the

parameters on the personalized model. The following sections detail the process for calculating these parameters.

It is important to note that variations in joint size among patients may affect the anatomical landmarks. For instance, the 15 slices above the tibial plafond may not always correspond to the exact same anatomical structures across different patients. However, for clinical parameter calculations, we adhere to established medical research guidelines, which approximate these anatomical landmarks consistently. The landmark selection is visually inspected and verified by an expert.

To quantify variability across observers, the interobserver variability was computed as the standard deviation of measurements obtained by two experts, reported together with the mean value. The experts' relative variability (%) was then defined as the variability normalized by the reference value (which is the average of the two experts), allowing comparisons across parameters of different magnitudes. Equation (4.12) defines experts' relative variability, where E_1 and E_2 denote the values measured by two independent observers.

$$Experts' \ relative \ variability(\%) = \frac{|E_1 - E_2|}{\frac{E_1 + E_2}{2}} \times 100 \tag{4.12}$$

As mentioned, the reference value Ref for each parameter was obtained as the average of the expert measurements. The automatic method error was evaluated by comparing the automatic measurement ($Auto$) to the reference. This is defined in Equation (4.13).

$$Auto \ vs \ Reference = |Auto - Ref| \tag{4.13}$$

Finally, the relative error (%) was defined as the absolute error normalized by the reference value, providing a scale-independent measure of accuracy (Equation (4.14)).

$$Relative \ Error \ (\%) = \frac{|Auto - Ref|}{Ref} \times 100 \tag{4.14}$$

These complementary metrics enable a comprehensive evaluation: interobserver variability highlights human measurement uncertainty, experts' relative variability expresses measurement reproducibility, reference values serve as the gold standard, and relative errors assess the accuracy of the proposed automatic method.

4.8.1 Selection of the Landmarks on the Statistical Shape Model

To calculate the syndesmosis clinical parameters, as described in Figure 1.7, we need to identify specific landmarks on the mean shape, according to the definition outlined in section 1.4.2. These landmarks are utilized following the reconstruction of the personalized joint model to compute clinical parameters. To reproduce the parameters reported in the literature (Nault et al., 2013), we calculated them in 2D, even though these parameters can also be extended into three-dimensional space, based on the reconstructed personalized model. This helps enabling direct comparison with values estimated by clinical experts on the CT slice (Nault et al., 2013), which are inherently two-dimensional. As mentioned in section 1.4.2, all parameters, except for 'angle2', are calculated 15 slices above the tibial plafond in the CT volume. This corresponds to 7.625 mm above the tibial plafond and 7.625 mm below the most superior vertex in the model (since the model includes 30 slices above the tibial plafond from the CT volumes). The 'angle2' parameter is calculated on the talar dome.
Below is the list of clinical parameters and the landmarks required for their calculation. These landmarks are shown in Figure 4.26.

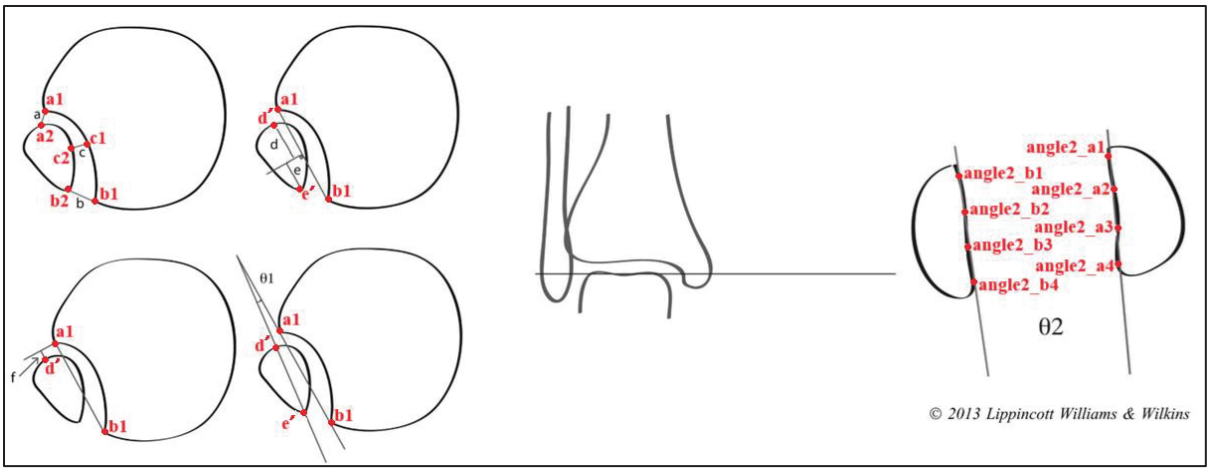

Figure 4.26 Syndesmosis clinical parameters and the landmarks needed for calculating them Taken and Modified from Nault et al. (2013)

- a: parameter ‘a’ is the distance between two landmarks ‘a1’ and ‘a2’.
- b: parameter ‘b’ is the distance between two landmarks ‘b1’ and ‘b2’.
- c: parameter ‘c’ is the distance between two landmarks ‘c1’ and ‘c2’.
- d and e: parameter ‘d’ and ‘e’ are calculated by drawing a line between landmarks ‘a1’ and ‘b1’, finding the center of this line, and then drawing a perpendicular line at the center. The distance from the landmark ‘d′’ (the most anterior point of the fibula) and ‘e′’ (the most posterior point of the fibula) to this perpendicular line defines parameters ‘d’ and ‘e’, respectively.
- f: parameter ‘f’ is calculated by drawing a line between landmarks ‘a1’ and ‘b1’ and then drawing a 2nd line at landmark ‘a1’, perpendicular to the 1st line. The distance between landmark ‘d′’ and the 2nd line defines parameter ‘f’.
- angle1: Parameter ‘angle1’ is the angle between two lines: the first line drawn between landmarks ‘a1’ and ‘b1’, and the second line drawn between landmarks ‘d′’ and ‘e′’.
- angle2: To calculate ‘angle2’, the talar dome is selected on the model, and 8 landmarks are identified: ‘angle2_a1’, ‘angle2_a2’, ‘angle2_a3’ and ‘angle2_a4’ on tibia, and ‘angle2_b1’, ‘angle2_b2’, ‘angle2_b3’ and ‘angle2_b4’ on fibula. The angle between these two lines passing through the selected points is computed to define ‘angle2’. Choosing four points for defining each line makes the calculation more robust to small movements of the vertices. The line is fitted to the 4 landmarks.

In conclusion, the originally defined clinical parameters are automatically calculated using these 16 identified landmarks. Figure 4.27 shows the landmarks and intersecting planes (black line at 7.625 mm above the tibial plafond and the blue line at talar dome). These landmarks are identified and recorded manually on the mean shape in 3D and then used to automatically calculate the clinical parameters after the personalized 3D reconstruction process is completed. To ensure the accuracy and reliability of the landmark selection process, it was reviewed and verified by a medical professional. However, due to the nature of triangular meshes, the selected vertices do not necessarily lie on a single plane (e.g., they may be slightly offset). While selecting vertices directly on the model facilitates the automatic calculation of clinical parameters, they may not provide an exact contour on the intersecting plane, nor identify landmarks as precisely as manual selection after reconstruction. Moreover, when vertices are not confined to a plane parallel to the CT slices, the resulting parameters may deviate from the clinical reference values established by two orthopedic specialists (Nault et al., 2013). To ensure consistency with expert assessments and maintain calculations in 2D, it is therefore necessary to project the model landmarks onto corresponding 2D planes and compute the clinical parameters accordingly. This step was essential to validate the method and confirm that the computed parameters align with those obtained through traditional manual assessment by experts.

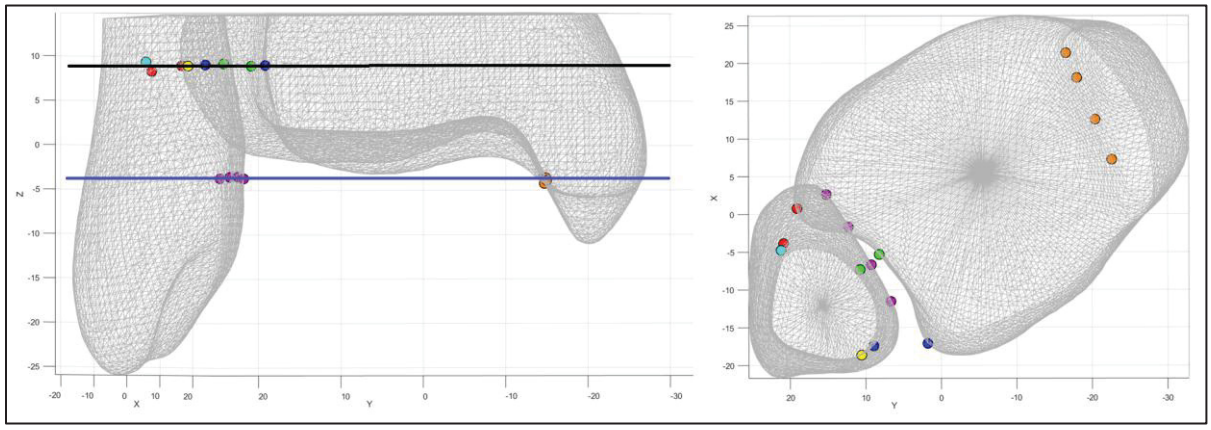

Figure 4.27 Landmarks identified to calculate the syndesmosis clinical parameters on the mean shape in 3D and at two different plane intersections (7.625 above the tibial plafond, indicated by the black line, and the talar dome, indicated by the blue line). ‘a1’ and ‘a2’ shown as red, ‘b1’ and ‘b2’ shown as blue, ‘c1’ and ‘c2’ shown as green, ‘d’ and ‘e’ shown as cyan and yellow respectively, and the 4 vertices for each line of angle2 are shown in orange and violet.

4.8.2 Calculation of the Parameters based on Reconstructed 3D model on a 2D reference plane

To match the clinical parameters calculated in our method with those described in the literature (Nault et al., 2013), which were originally computed in 2D from CT scan slices, we projected all the selected landmarks defined on the 3D model (outlined in previous section) onto their corresponding 2D planes. This projection ensures consistency with the clinical parameters calculated in the 2D space of CT scans.

To facilitate the projection process, a reference plane is defined by three vertices located at the most superior part of the model, as shown in Figure 4.28 by the purple line. This plane is parallel to all CT slices, including the one 7.625 mm above the tibial plafond and the one at the talar dome, and is easier to define as a large number of vertices lie on it. We project all identified 3D landmarks onto this plane and then proceed to calculate the clinical parameters in 2D space. This approach allows us to directly compare our computed parameters with those traditionally derived from CT scan slices, ensuring that the method is consistent with existing clinical practices.

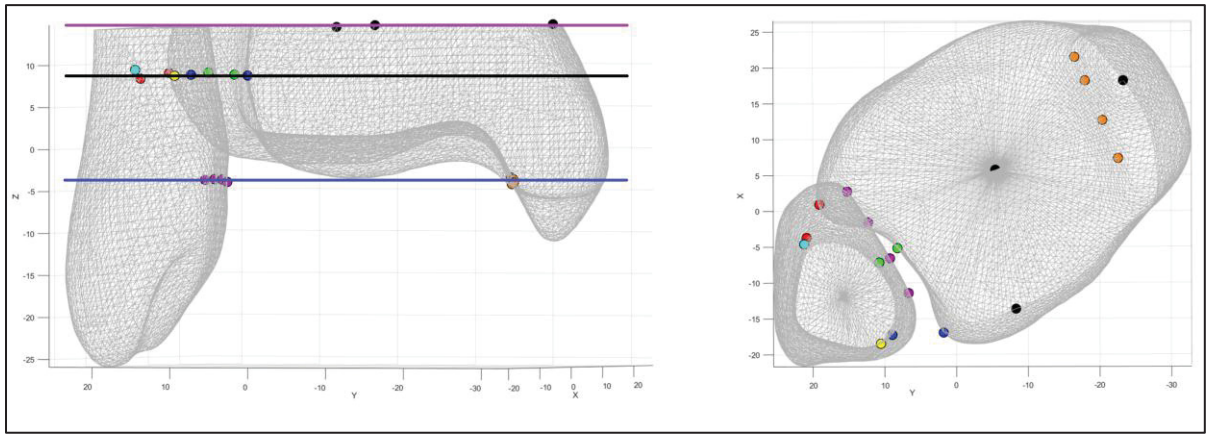

Figure 4.28 Showing existing and new landmarks to define the reference plane. The three landmarks are shown in black.

4.9 Conclusion

This chapter provided a comprehensive explanation of the methodology used to reconstruct a personalized 3D model of the tibiofibular joint from ankle biplanar radiographs. The methodology is structured in several key stages, including dataset cleaning and preprocessing, multi-class segmentation of the CT scan cases, 3D statistical shape and intensity modeling, Joint extraction from the biplanar radiographs using deep learning-based methods, 2D/3D deep learning-based registration, and the assessment of syndesmosis clinical parameters. Each step was carefully designed to ensure accuracy, robustness, and reproducibility in the reconstruction process.

The first step involved dataset cleaning, which was essential for ensuring high-quality input data. Multi-class segmentation was then performed on the CT scan dataset to differentiate the tibia, fibula, and foot from other anatomical structures. This segmentation phase played a critical role in the accuracy of the subsequent 3D modeling and reconstruction steps.

A core part of this methodology was the construction of a 3D SSIM for the tibiofibular joint. This process involved meshing segmented joint structures, matching the meshes, aligning them through Generalized Procrustes Analysis (GPA), and applying Principal Component Analysis (PCA) to capture shape variations. By incorporating intensity information into the model, the reconstruction approach ensured a more precise representation of the bone structures.

Another significant aspect was the extraction of the tibiofibular joint in radiographs. This step aimed to isolate the joint while eliminating other structures. A deep learning-based denoising autoencoder model was trained for this purpose, using a dataset of digitally reconstructed radiographs from the segmented volumes. Post-processing techniques were applied to further refine the outputs.

A deep learning-based one-shot 2D/3D registration approach was then implemented to align the 3D statistical model with joint's biplanar radiographs. This registration process aimed to minimize the reconstruction time while increasing the reconstruction accuracy. By leveraging a CNN-based network for 2D/3D registration for both feature extraction and the registration task, the applied methodology improved automation and reduced reliance on manual interventions.

Finally, syndesmosis clinical parameters were extracted from the personalized reconstructed 3D model. The placement of anatomical landmarks on the mean shape was obtained manually, which is then used to enable automatic calculation of the parameters.

Overall, the methodology presented in this chapter establishes a novel and robust pipeline for reconstructing personalized 3D models of the tibiofibular joint. The combination of the statistical shape modeling and deep learning-based registration ensures high accuracy and reproducibility, leveraging the advances of each technique. The next chapter will present the experimental validation and results, evaluating the effectiveness of this approach through quantitative and qualitative assessments.

CHAPITRE 5

EXPERIMENTS AND RESULTS

5.1 Introduction

In the previous chapter, we introduced the methodology for reconstructing a personalized 3D model of the tibiofibular joint using biplanar radiographs of the ankle. This methodology encompasses multiple stages, including dataset preparation, multi-class segmentation, 3D statistical shape and intensity modeling, Joint extraction from radiographs, deep learning-based 2D/3D reconstruction, and the extraction of clinically relevant syndesmosis parameters. Each stage was designed to generate accurate results independently while ensuring seamless integration with the other steps.

This chapter presents the experimental validation of the proposed methodology, aiming to assess the performance of each stage of the pipeline using both quantitative and qualitative metrics, where applicable. Additionally, the integration of individual components into the complete 2D/3D reconstruction pipeline is evaluated to ensure coherence and accuracy across the system.

The chapter is organized into two main sections: Experiments and Results. The Experiments section details the implementation and testing procedures for each component of the methodology. The Results section presents the outcomes of these experiments along with corresponding analyses.

The Experiments section begins with the procedures applied to the dataset, followed by experiments involving the Statistical Shape and Intensity Models, the 2D/3D reconstruction network, and the Joint Extraction Networks. This is followed by experiments related to region of interest (ROI) selection and the assembly of the full reconstruction pipeline. The final subsection addresses the experimental procedures used for the estimation of clinical parameters.

In the Results section, we begin by analyzing the dataset characteristics and presenting a qualitative evaluation of the segmentation results and mesh matching process. The Statistical Shape and Intensity Models are assessed using both qualitative comparisons and quantitative

metrics. The 2D/3D reconstruction network is evaluated based on its convergence graph, parameter-wise normalized loss values, point-to-point reconstruction accuracy in millimeters, and shape and pose parameter errors. The performance of the Joint Extraction Networks is then analyzed by convergence graphs, loss values, and PSNR between the generated and ground truth images. Finally, we present the results of the complete reconstruction pipeline, demonstrating its effectiveness in producing accurate 3D model reconstruction from ankle biplanar radiograph. An evaluation of the effect of ROI selection on reconstruction accuracy of the pipeline is explored in the next section, considering both ideal and noisy selections of crop lines. The chapter concludes with an evaluation of the estimated clinical parameters, which are compared against reference values from professional physicians specializing in ankle, to validate their accuracy.

The outcomes of these experiments will help determine the feasibility of our approach for clinical applications. By systematically analyzing the performance of each methodological component, this chapter provides a comprehensive evaluation of our method's strengths and limitations. These findings will serve as the basis for the next chapter, where we will discuss the broader implications of our results and explore potential refinements to enhance the methodology.

5.2 Experimental Design and Implementation

5.2.1 Dataset Statistics

The cleaned dataset consists of 53 ankle CT scans collected between 2007 and 2018, covering patients aged 17 to 69 years. Among these, 25 scans correspond to the right ankle and 28 to the left. The majority of cases involve patients diagnosed with Talus or Calcaneus fractures. However, all were classified as non-pathological with respect to the tibiofibular syndesmosis. Some cases include the presence of medical interventions such as pins in the Talus or Calcaneus, as well as arthrography. These procedures did not impact the tibiofibular joint, which is the primary focus of this study.

To ensure consistency across the dataset, scans were analyzed for spatial resolution characteristics. While pixel spacing varies from 0.23 mm to 0.54 mm between cases, the slice thickness and inter-slice spacing are uniformly set at 0.625 mm and 0.5 mm, respectively. These parameters were taken into account during the conversion of segmented CT volumes into 3D meshes. The generated meshes were represented in millimeters to maintain geometric consistency and were subsequently simplified to ensure compatibility with downstream processing stages. The experiments applied for segmentation, mesh conversion, and simplification are detailed in next sections.

The dataset used in this study consists exclusively of non-pathological syndesmosis cases, meaning the proposed pipeline has been trained and validated only under normal syndesmosis anatomical conditions. This restriction limits its current applicability to pathological scenarios.

5.2.2 Segmentation of the Joint and Ankle

The segmentation process was evaluated qualitatively through expert visual inspection. An experienced physician specializing in ankle surgery and clinical assessment manually reviewed three randomly selected cases from the dataset. For each selected case, the physician examined all segmented CT slices of both the ankle and the tibiofibular joint in a consecutive slice-by-slice manner to verify anatomical accuracy and consistency.

Following the segmentation, the 3D masks of the ankle, tibia, and fibula were post-processed using a Gaussian smoothing filter. A 3D Gaussian kernel with a standard deviation of 1.5 was applied to each mask. This value was determined empirically to balance the smoothing effect, reducing voxel-level noise while preserving the original shape and avoiding significant geometric shrinkage. The smoothed masks ensured suitable surface quality for subsequent mesh generation.

5.2.3 Generation of the Meshes on CT scan masks

The 3D surface meshes of the tibia and fibula were generated in both high- and low-density versions to support different stages of the reconstruction pipeline. The dense tibia meshes

consist of approximately 18K to 71K vertices, while the fibula meshes range from 11K to 39K vertices. The low-density versions of tibia and fibula meshes consist of approximately 6K to 24K, and 4K to 13K, respectively. Consequently, the joint sizes for low-density meshes range from 10K to 38K vertices. The high-resolution models retain detailed anatomical structure. Figure 5.1 illustrates the largest and the smallest joints (tibia and fibula) across the dataset.

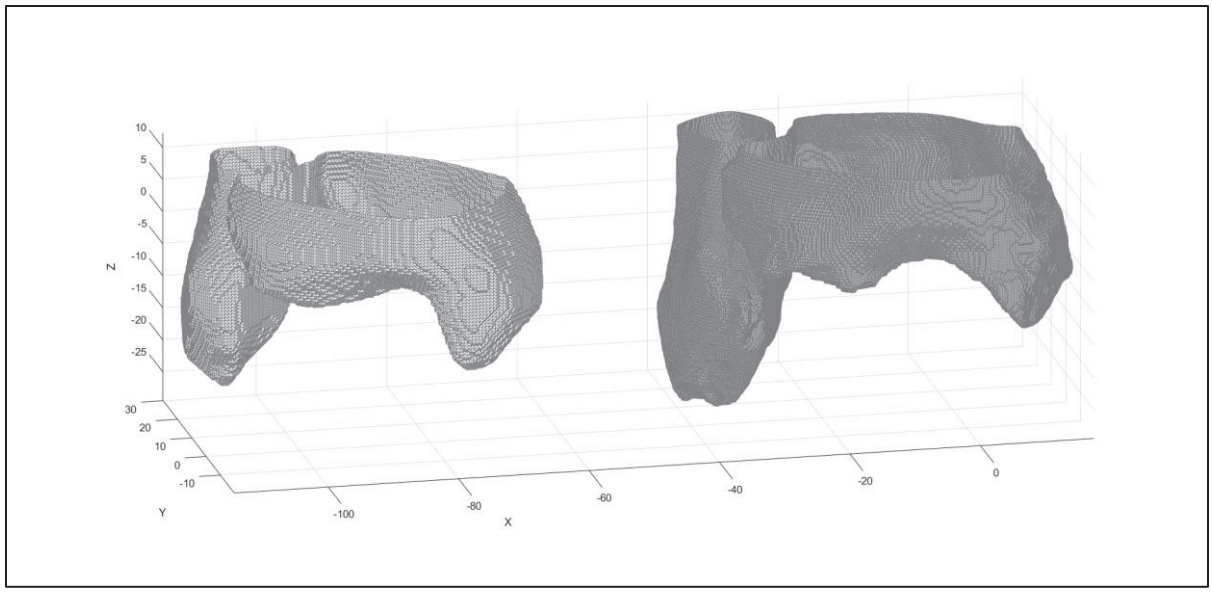

Figure 5.1 The largest and smallest ankle dense meshes in the dataset

The generated meshes are not watertight (Figure 5.1) due to the meshing process. One vertex is added to the center of the top vertices of all low-density tibias and all fibulas within the dataset and corresponding edges are also added to form new faces. Figure 5.2 shows a side-by-side comparison of a high-density mesh and its low-density version. In the representative case, the tibia mesh size was reduced from 35,135 to 12,017 vertices, and the fibula from 20,135 to 6,936 vertices.

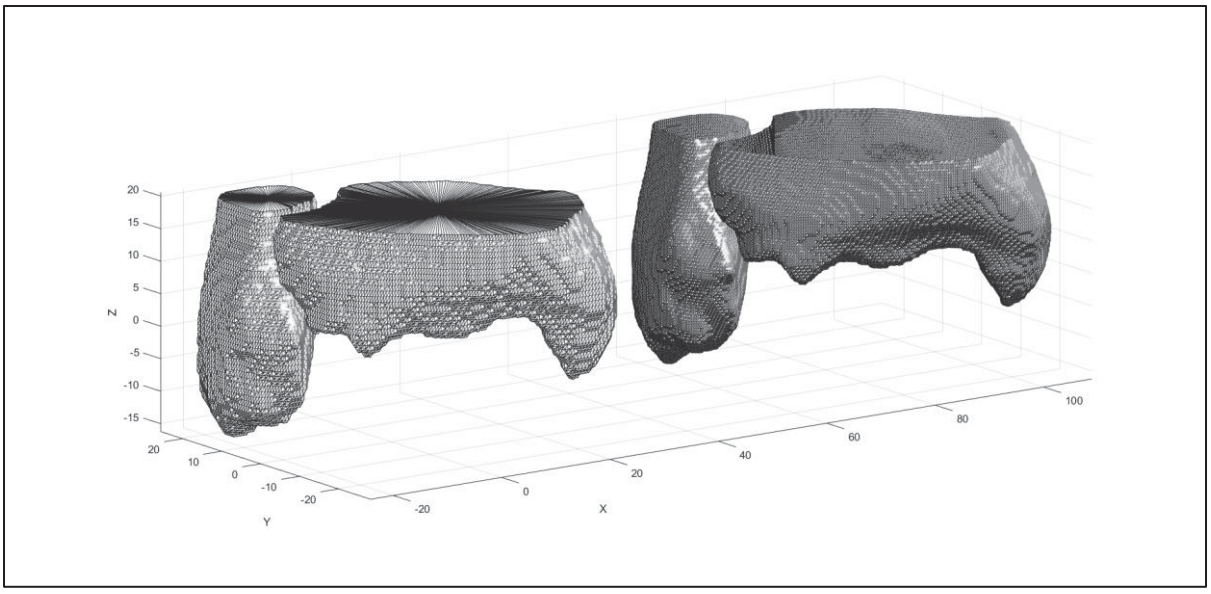

Figure 5.2 The largest ankle dense meshes and their simplified watertight version

5.2.4 Building a 3D Statistical Shape and Intensity Model of the Tibiofibular Joint

After completing the segmentation process, the next step involved matching the meshes and constructing a statistical shape model based on the matched meshes. The following sections present the experiments related to these steps.

5.2.4.1 Matching the Meshes

To enable statistical analysis and ensure that all meshes are aligned within a common coordinate system, a consistent topology across the dataset is required. This means that each mesh within its class (tibia or fibula) must have the same number of vertices and faces, enabling vertex-to-vertex correspondence across subjects. Standardizing the topology ensures that corresponding vertices represent the same anatomical landmarks in each mesh.

To achieve this, a one-to-one registration process was applied by aggregating vertices in one-to-many matches, resulting in standardized meshes. Following this procedure, all tibia meshes were standardized to 6,278 vertices, and all fibula meshes to 3,653 vertices. This uniform

representation is critical for subsequent tasks such as statistical shape modeling and vertex-wise analysis.

5.2.4.2 Generalized Procrustes and Principal Components Analysis

Given that all meshes share a consistent topology, Generalized Procrustes Analysis (GPA) was applied to eliminate variations due to rotation, translation, and scaling. This normalization step is essential for performing Principal Component Analysis (PCA) on the vertex data.

Prior to applying GPA and PCA, the tibiofibular joint was modeled by combining the vertices and faces of the tibia and fibula meshes for each subject. For each case, the joint vertex matrix $joint_i^V$ consists of 9,931 vertices, obtained by concatenating the tibia and fibula meshes. GPA was then applied to all joint meshes $joint_i^V, j = 1..53$ to bring them into a common coordinate frame.

Following alignment, PCA was applied to the matrix of joint vertices. To represent 90% of the total shape variation, the first 18 principal components (eigenvectors) were retained. This reduced representation was used to evaluate the performance of the Statistical Shape Model (SSM) of the tibiofibular joint.

5.2.4.3 Intensity Modeling

To establish a baseline for intensity modeling, we computed the average voxel intensity of the tibia and fibula volumes across all cases in the dataset. This basic approach serves as a simplified intensity representation when generating volumetric data from the Statistical Shape Model, facilitating the repetitive process.

5.2.5 One-shot 2D/3D Reconstruction of the Distal Tibiofibular Joint from Biplanar Radiographs using Deep Learning Registration

The 2D/3D reconstruction network was trained using DRRs generated from the Statistical Shape Model by systematically varying its parameters. This section describes the experimental

setup used to train and test the network, including the generation of input data and the design of the training procedure.

5.2.5.1 Generated Dataset for 2D/3D Reconstruction Network

A total of 30,000 synthetic instances were generated using the method described in Section 4.4.2 to train the 2D/3D registration network. Each instance includes an AP-LT radiograph of the tibiofibular joint, 18 shape parameters, 3 rotational parameters, and one scaling parameter. All parameters were normalized relative to their respective maximum values to ensure consistency and stability during training. These maximum values were retained to allow for accurate rescaling to the original parameter ranges when needed.

The synthetic DRRs were preprocessed to produce uniform input dimensions, with the final image size set to 460 × 890 pixels after all post-processing steps.

5.2.5.2 Hyperparameter Optimization

To identify the optimal combination of hyperparameters for the 2D/3D registration network, a grid search strategy was employed. The following hyperparameters and value sets were tested:

- Learning rate: {0.15, 0. 1, 0.05, 0.01}
- Dropout rate: {0.1, 0.2, 0.3, 0.4}
- Batch size: {8, 12, 16}

Each configuration was trained for five epochs to allow for rapid evaluation. To further reduce computational overhead, only the first 5,000 instances from the dataset were used during this optimization stage.

Model performance was monitored using the loss values, which are directly correlated with the accuracy of the predicted parameters and, consequently, the precision of the reconstructed 3D shapes. For evaluation purposes, the dataset was split into training (80%), validation (10%), and test (10%) subsets, with the test set comprising 500 instances.

5.2.5.3 Training, Validation, and Testing of the Network

With the optimal hyperparameters identified, the 2D/3D reconstruction network was trained using the full dataset of 30,000 synthetic instances. From this dataset, 10% (3,000 instances) were reserved as a test set, while the remaining data were further divided into training (90%) and validation (10%) subsets.
The training was conducted using early stopping based on validation loss, and the learning rate was scheduled to automatically reduce when no improvement was observed.

5.2.5.4 Evaluation of the 3D reconstruction accuracy of the Shape

While error metrics in the parameter space provide useful insights into network prediction performance, they do not fully reflect the accuracy of the reconstructed 3D geometry. Given that the reconstructed meshes generated from shape parameters share the same topology as the statistical shape model, a more direct geometric evaluation was conducted in 3D space.
For each test instance, a mesh was reconstructed using the network-predicted shape parameters. This mesh was then rotated and scaled based on the predicted pose and scaling parameters. A corresponding ground truth mesh was reconstructed using the recorded (true) shape and pose parameters. The two meshes were then compared using point-to-point error in 3D space (in millimeters), with errors computed over all corresponding vertices.

5.2.6 Isolation of the Tibiofibular Joint skeletal structure in biplanar radiographs

This section describes the experimental setup for training and testing the two Joint Extraction Networks (JE-Nets), which are responsible for isolating the tibia and fibula from other structures in the input images.

5.2.6.1 Generated Dataset for Joint Extraction Networks

As described in Section 4.5.2 , the 3D volumes were rotated randomly using a set of pose parameters to synthetically augment the training dataset. The parameters are taken from uniform distributions. This technique is used to simulate anatomical variability and to increase the diversity of the training data.

To generate 2D training images, the voxel projector process was applied to the rotated volumes to create DRRs of the ankle and joint in AP and LT directions. This process resulted in a total of 15,000 AP-LT image pairs for use in training and evaluation.

After generating the DRRs, post-processing was applied to standardize the image dimensions across the dataset. The final image resolution was set to 620 × 400 pixels for AP images and 980 × 400 pixels for LT images.

5.2.6.2 Hyperparameter Optimization

A grid search was conducted to determine the optimal combination of hyperparameters for the Joint Extraction Networks (JE-Nets), which were trained separately for AP and LT networks. The following hyperparameter values were tested:

- Learning rate: {0.005, 0.001, 0.0005, 0.0001, 0.00005}
- Dropout rate: {0.1, 0.2, 0.3, 0.4}
- Batch size: {8, 12, 16}
- Alpha: {0.95, 0.9, 0.8, 0.7}, where alpha controls the balance between the network's loss components.

Each combination was trained for 5 epochs to enable rapid evaluation. To further accelerate the optimization, only the first 5,000 instances from the dataset were used for training.

Instead of relying solely on training or validation loss, Peak Signal-to-Noise Ratio (PSNR) was used as the evaluation metric, as it directly reflects the visual quality of the predicted outputs. The dataset was partitioned into training (80%), validation (10%), and test (10%), with 500 test instances used for PSNR evaluation across all configurations.

5.2.6.3 Training, Validation, and Testing the Joint Extraction Networks

With the optimal hyperparameters identified, both JE-Net-AP and JE-Net-LT were trained using the complete dataset of 15,000 instances. Of this dataset, 10% (1,500 instances) were reserved for testing, while the remaining data were divided into training (90%) and validation (10%) subsets.

Early stopping based on validation loss was applied to prevent overfitting. Additionally, a learning rate reduction strategy was used, where the learning rate was halved if no improvement in validation loss was observed. Training and validation loss values were monitored throughout the training process for both networks.

5.2.7 Establishing the 2D/3D Reconstruction Pipeline

Since the components of the pipeline were trained independently, the next phase involved integrating them into a complete system and evaluating its overall performance on tibiofibular joint reconstruction from ankle radiographs. For this evaluation, three left-out CT cases were used, and cases are augmented to 1,000 instances through random rotations of both ankle and joint volumes. This augmentation was designed to assess the rotational robustness of the pipeline.

Once the dataset is generated and saved in DICOM format, the voxel projector was applied to generate AP and LT DRRs of ankle for each instance. A perfect ROI selection was performed automatically using the joint segmentation masks to localize the region of interest, simulating the ROI selection by the operator.

The resulting AP and LT images were then processed to ensure compatibility with the input dimensions expected by the JE-Nets. These networks extract the joint structures and adjust pixel intensity values according to the intensity information stored in the model. The output images were then refined by cropping to isolate the joint regions in both AP and LT views, following by a concatenation process to form an AP-LT image of the joint.

Finally, each AP-LT image was passed through the trained 2D/3D reconstruction network, which predicted the shape and pose parameters required to reconstruct the corresponding 3D mesh of the joint. All images were prepared to match the size and intensity characteristics of the network's training data.

To assess the accuracy of the 3D reconstruction produced by the pipeline, the point-to-surface (P2S) distance was used as the evaluation metric. Each reconstructed mesh was compared to its corresponding dense ground truth mesh, which accurately represents the surface of the segmented CT data. For every vertex in the reconstructed mesh, the minimum distance to the surface (represented by dense vertices) of the ground truth mesh was computed.

5.2.8 Reproducibility Error on Selection of the Region of Interest

To simulate real-world conditions where region of interest (ROI) selection will likely be performed manually, an additional evaluation dataset was created. This dataset was derived from the same 1,000 augmented instances used in the previous experiment. A uniform randomly selected noise of up to 12 pixels was introduced to the automatically selected ROI, following the procedure outlined in the methodology section. The aim is to analyze the effect of the non-perfect ROI selection on the reconstruction accuracy.

Noise was applied independently to the tibial plafond and distal fibula selected landmarks. For the distal fibula, the noise ranged from −12 to +12 pixels, as the landmark could shift in either direction. For the tibial plafond, only positive noise values were used, since upward shifts (toward the knee) are constrained by the limited number of slices available in the image volume.

5.2.9 Syndesmosis Clinical Parameters calculation on reconstructed 3D model

The evaluation of syndesmosis clinical parameters was performed on the three left-out CT instances used in the full pipeline assessment. Since the 1,000 augmented cases involved only rotational transformations and did not introduce shape deformation, the clinical parameters

remained unchanged across these augmented instances. Consequently, only the original three base cases were retained for evaluating the accuracy of clinical measurements.

To establish ground-truth values, two fellowship-trained ankle specialists independently measured the relevant anatomical parameters on the reference CT data. Their measurements served as the clinical reference values against which the reconstructed meshes were compared. This dual-observer evaluation not only provided a reliable gold standard but also allowed us to quantify the interobserver variability, which reflects the inherent uncertainty and variability between expert clinicians.

Building on these reference values, several complementary metrics were calculated to comprehensively assess the accuracy of clinical parameter estimation from the reconstructed 3D models. These include the interobserver variability, which quantifies agreement between the two experts, the experts' relative variability, which normalizes variability by the parameter magnitude for non-scaled comparison, the reference values themselves, representing the gold standard, the absolute error between the automatic method and the reference (Auto vs. Ref), which indicates the raw deviation of our approach, and finally the relative error, which normalizes this deviation to enable scale-independent comparisons across parameters of different units and magnitudes.

By integrating these metrics, the evaluation provides both a measure of the reliability of expert-derived parameters and a detailed assessment of the accuracy of the proposed automatic reconstruction method. This multi-metric analysis ensures that the results capture not only the average error but also the variability introduced by human observers, thereby offering a more clinically relevant perspective on the robustness and applicability of the pipeline.

5.3 Evaluation Results and Analysis

5.3.1 Dataset Statistics

Although all CT scans were classified as non-pathological, certain cases exhibited treatment-induced distortions that affected the quality of intensity information. Artifacts introduced by medical interventions, such as metallic pins, were visible in a subset of cases and posed

potential challenges for the intensity-based learning stages of the pipeline, particularly in the training of the Joint Extraction Networks (JE-Nets). Nonetheless, as the tibia and fibula remained structurally unaffected, the primary anatomical targets for reconstruction and analysis were preserved.

Given that the methodology is intended for application on anteroposterior (AP) and lateral (LT) radiographs acquired in a weight-bearing and neutral position, the ideal reference would be CT scans captured under similar biomechanical conditions. Although the current dataset lacks such simulations, future data acquisition could incorporate external forces during CT to replicate weight-bearing postures.

It is also noted that minor positional variations of the tibia and fibula during plantarflexion and dorsiflexion are negligible. However, other structures like the talus are more sensitive to such movements. To minimize alignment variability and improve the robustness and clinical relevance of the reconstructed 3D models, it is recommended that all CT scans be acquired with the foot positioned neutrally.

5.3.2 Segmentation of the joint and Ankle

The results of the segmentation process are illustrated in Figure 5.3, which shows 22 consecutive CT slices with multi-class segmentation. The tibia is labeled in red, the fibula in blue, and the ankle and foot structures in yellow. As the slices move toward the distal end of the fibula, the tibia, initially segmented as a single continuous structure, begins to separate into multiple regions as the talus becomes visible in the center. Eventually, the tibia disappears, and in contrast, the distal fibula remains present over several additional slices before gradually fading as the image approaches the heel bone.

To ensure full anatomical coverage and incorporate a safety margin, the ankle segmentation is extended by an additional 15 slices beyond the slice where the fibula completely disappears. Slice numbers are included in the corner of each image for reference. A full example of the segmented slice sequence is provided in ANNEX I.

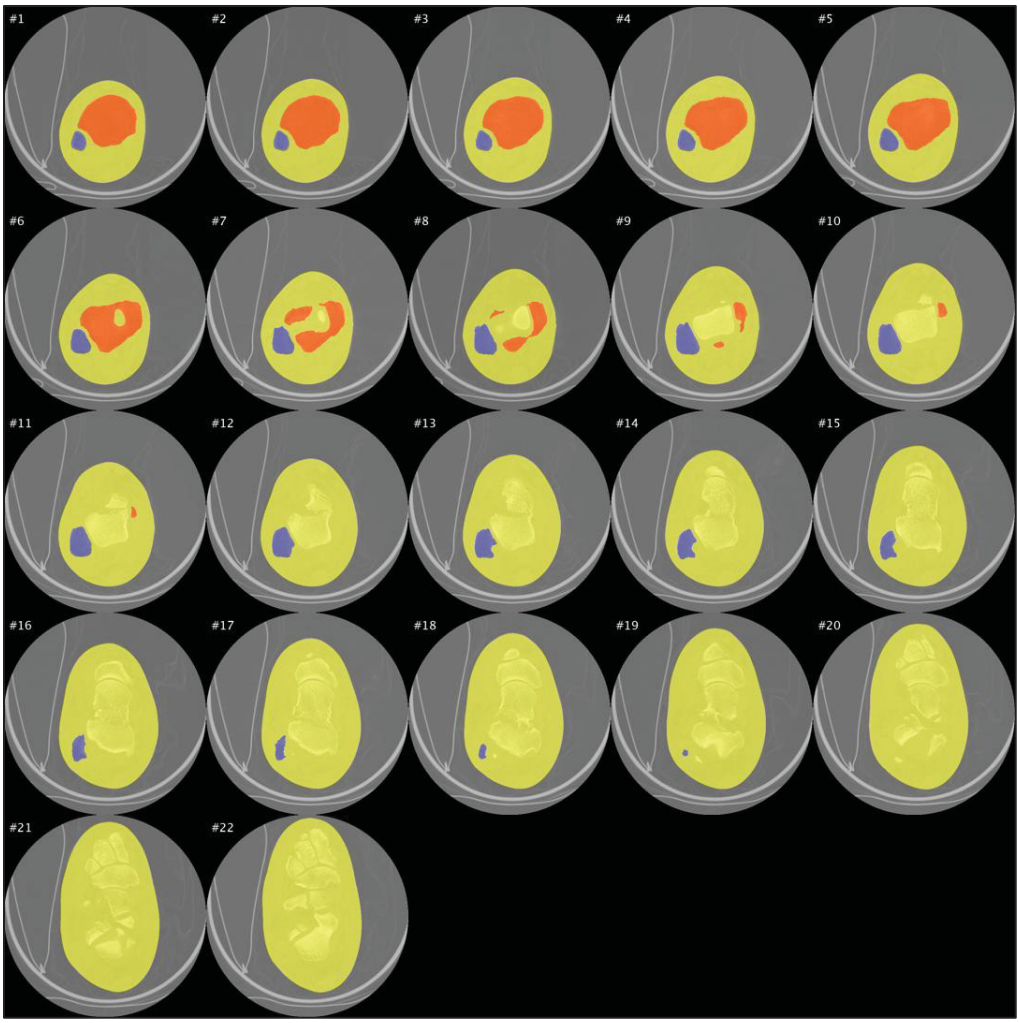

Figure 5.3 multi-class segmented CT slices, indicating tibia (in red), fibula (in blue) and ankle and foot (in yellow)

5.3.3 Generation of Meshes on CT scan masks

The impact of mesh simplification on surface fidelity was evaluated by calculating the Point-to-Surface (P2S) errors in millimeters between the dense and simplified meshes. Figure 5.4 and Figure 5.5 illustrate the distribution of 3D surface errors for the tibia and fibula, respectively, across all cases in the dataset. These figures provide a per-case visualization of surface deviation due to mesh simplification.

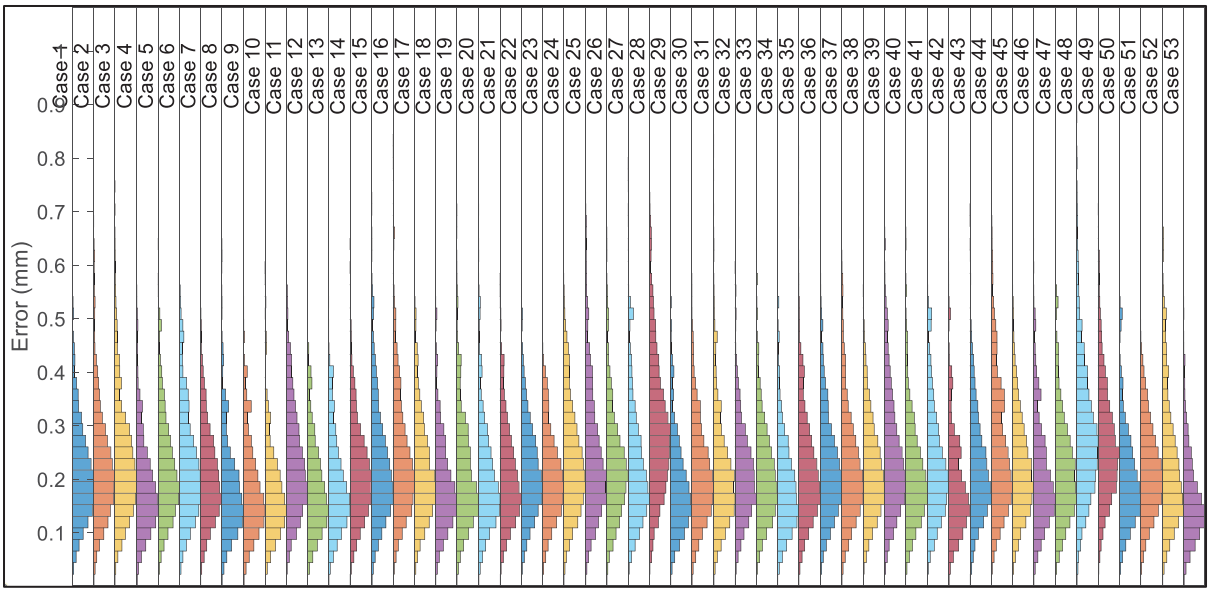

Figure 5.4 Distribution of the errors in 3D space caused by simplification of the tibia meshes

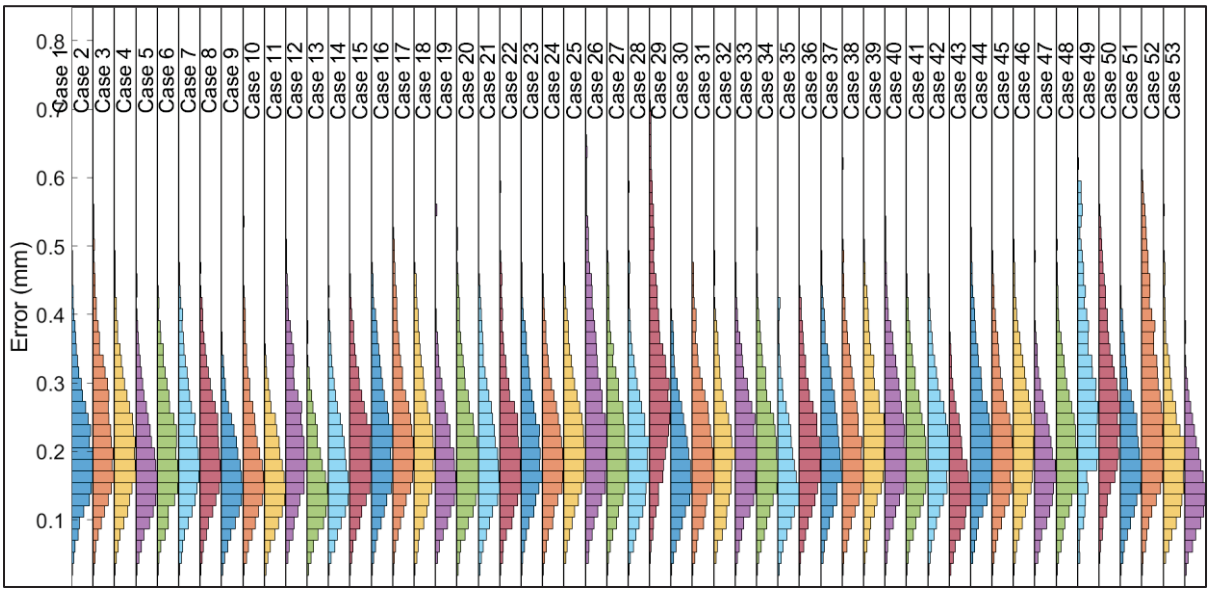

Figure 5.5 Distribution of the errors in 3D space caused by simplification of the fibula meshes

Summary statistics of the P2S errors are reported in Table 5.1. These results confirm that the simplification process preserves anatomical shape fidelity with sub-millimeter accuracy, making the low-density meshes suitable for further modeling and analysis.

Table 5.1 Mesh simplification error statistics

Shape	Mean ± STD (mm)	Min (mm)	Max (mm)
Tibia	0.22 ± 0.10	0.00	1.08
Fibula	0.21 ± 0.09	0.00	0.85

5.3.4 Building a 3D Statistical Shape and Intensity Model of the Tibiofibular Joint

5.3.4.1 Match the Meshes

The consistency of mesh alignment and vertex correspondence was qualitatively evaluated using color-coded visualizations. Figure 5.6 and Figure 5.7 display four randomly selected tibia and fibula meshes, respectively, with color maps applied based on vertex index IDs. Although ground truth correspondences are unavailable, the visual consistency of color patterns across different meshes provides a qualitative measure of registration accuracy. Similar color distributions across anatomically equivalent regions indicate that the meshes are well-aligned and that corresponding vertices represent the same anatomical landmarks across the dataset. In addition, an indirect quantitative measure, proving the quality of the correspondences are the quality metrics of the SSM built on the matched meshes, which is presented in the following sections.

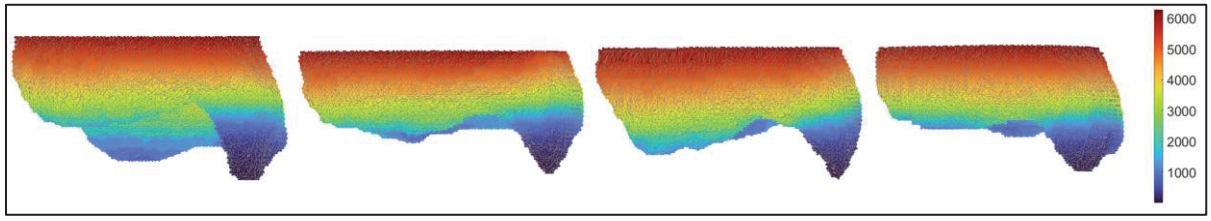

Figure 5.6 Randomly selected tibia meshes, colorized by the vertex indices, showing visual consistency through shapes

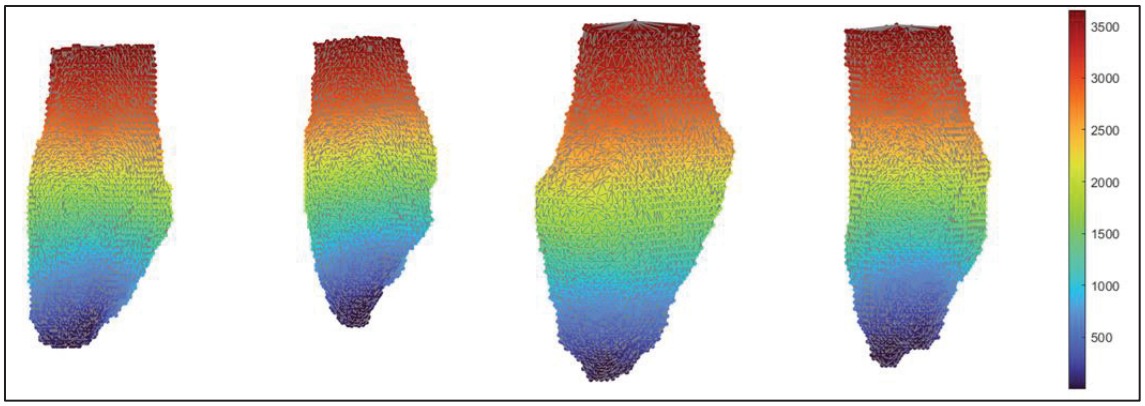

Figure 5.7 Randomly selected fibula meshes, colorized by the vertex indices, showing visual consistency through shapes

5.3.4.2 Generalized Procrustes and Principal Components Analysis

The performance of the Statistical Shape Model was evaluated using four standard criteria: Compactness, Generalization, Specificity, and Accuracy. These metrics assess different aspects of model quality:

- Compactness measures how efficiently the model captures shape variation with a limited number of components.
- Generalization assesses how well the model can represent unseen shapes.
- Specificity evaluates how realistic the shapes generated by the model are.
- Accuracy quantifies how well the model reconstructs known shapes.

The corresponding results are illustrated in Figure 5.8, Figure 5.9, Figure 5.10 and Figure 5.11. For the specificity analysis, 1,000 synthetic samples were generated using the learned shape parameters.

Strong performance across all four metrics also serves as an indirect validation of the mesh matching and alignment process, indicating that the statistical model was constructed from consistently aligned and topologically correct data.

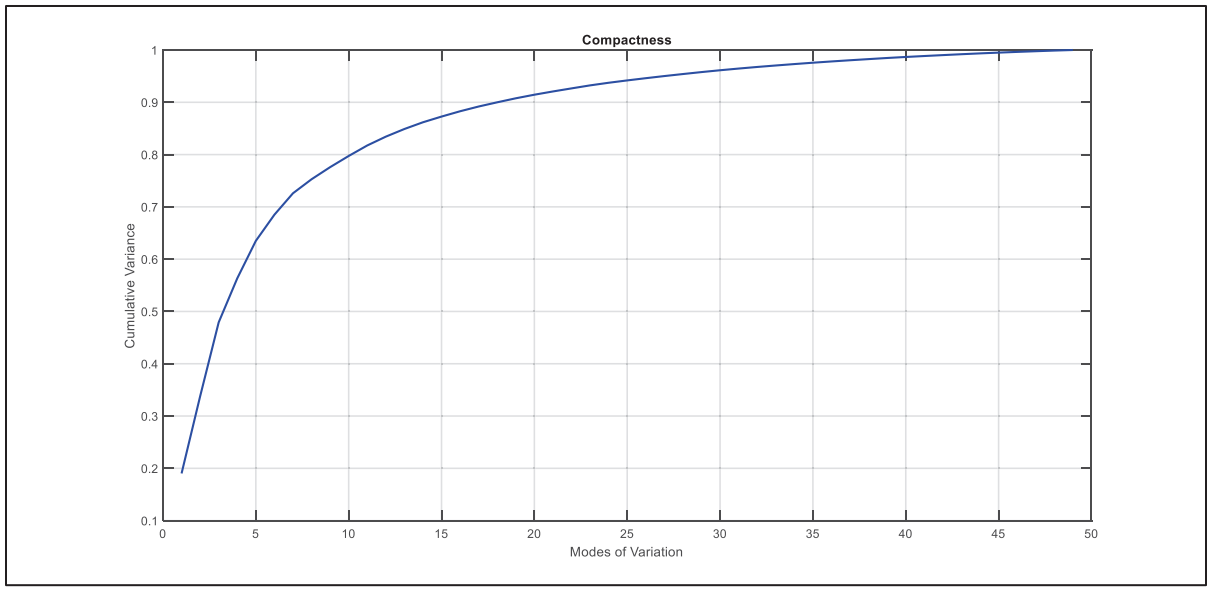

Figure 5.8 Compactness of the joint model

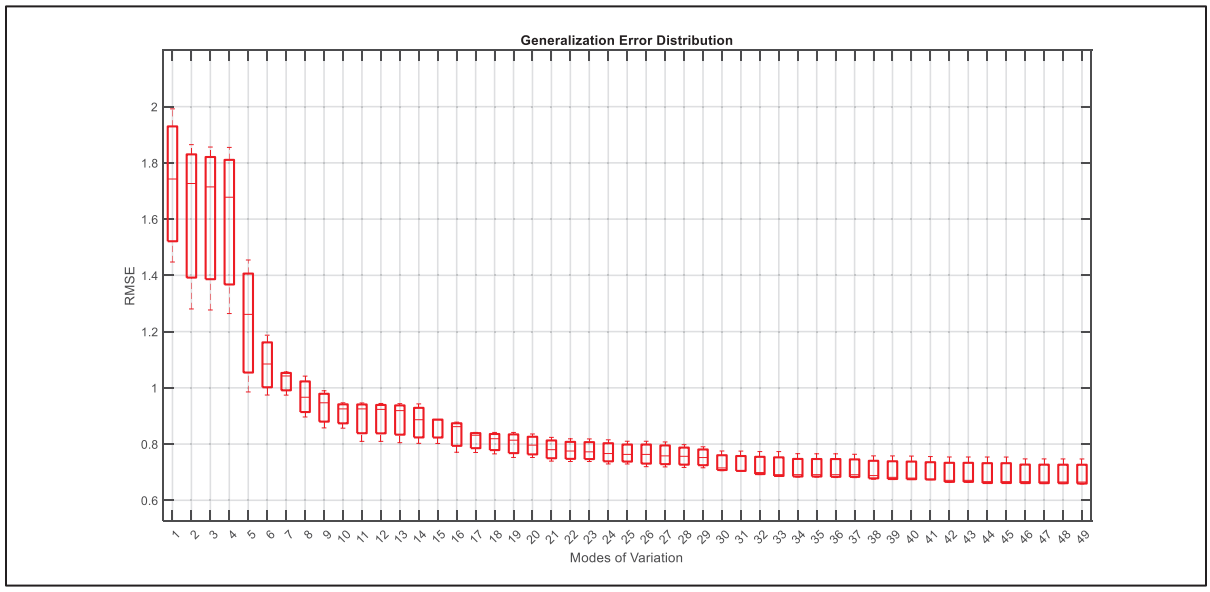

Figure 5.9 Generalization of the joint model (mm)

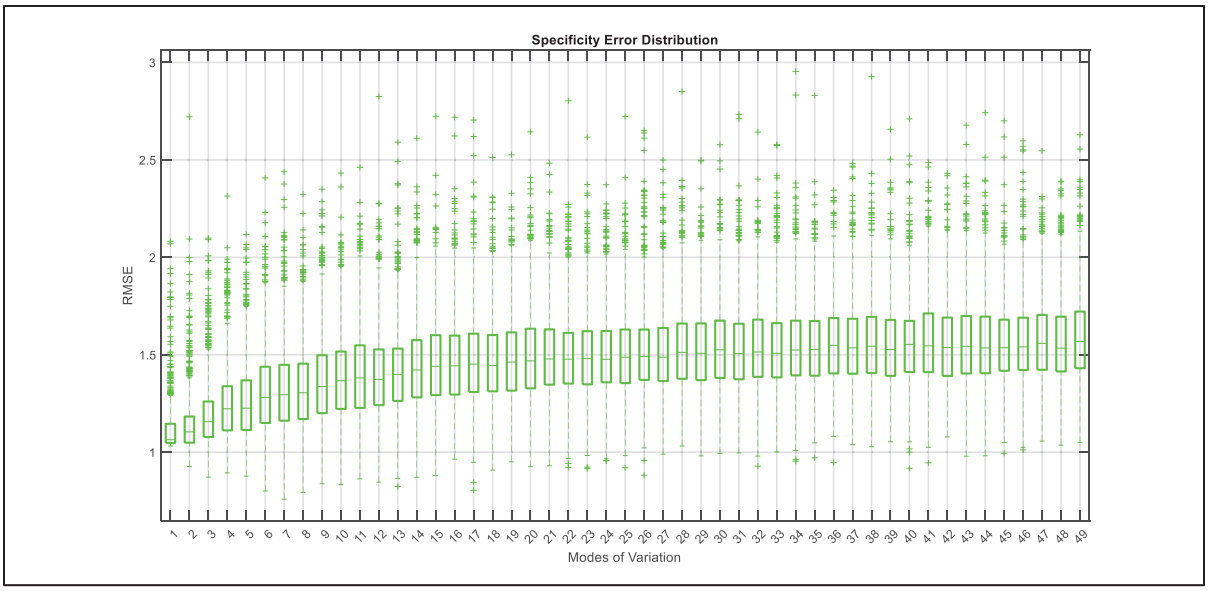

Figure 5.10 Specificity of the joint model (mm)

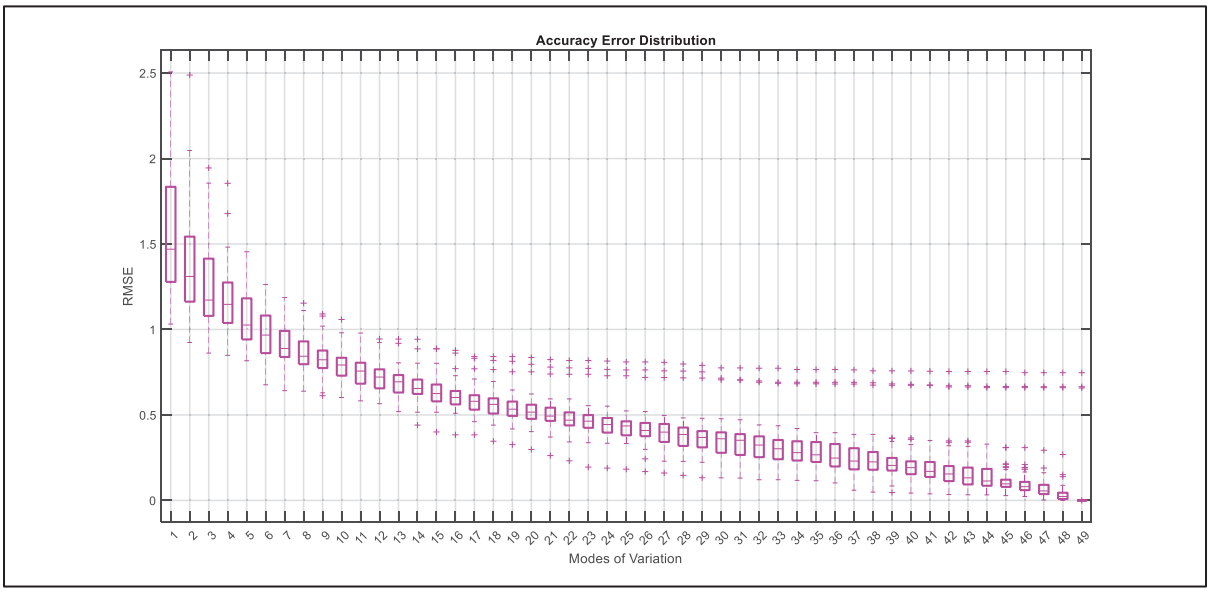

Figure 5.11 Accuracy of the joint model (mm)

5.3.4.3 Intensity Modeling

The resulting average intensity, denoted as I^{simple}, was computed to be 1321 in a 16-bit image space. This single intensity value is used to fill the voxels when generating synthetic volumes from the model, providing a uniform approximation of bone density.

5.3.5 One-shot 2D/3D Reconstruction of the Distal Tibiofibular Joint from Biplanar Radiographs using Deep Learning Registration

5.3.5.1 Generated Dataset for 2D/3D Registration Network

To better understand the variability and coverage of the generated training data, the distributions of the shape and pose parameters were visualized using histograms. Figure 5.12 presents the histograms for the 18 shape parameters, illustrating the range and density of shape variation across the synthetic dataset. Figure 5.13 displays the distributions of the rotational parameters (in radians) and the scaling parameter, confirming a diverse and well-balanced representation of pose and scale variations for training.

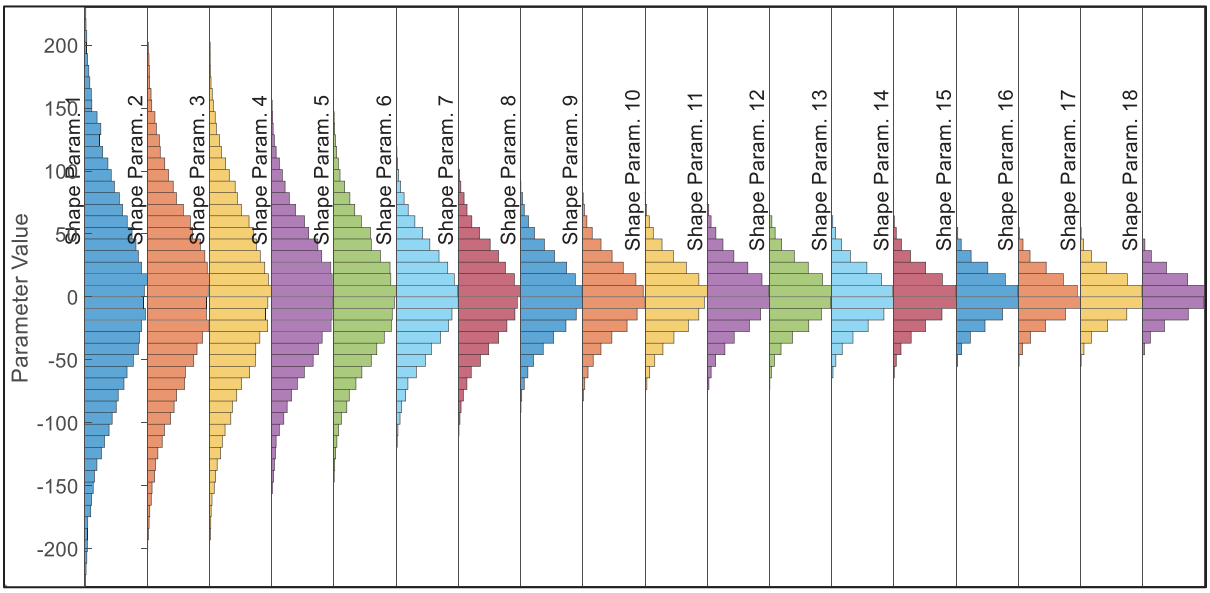

Figure 5.12 Histogram of Shape parameters in the generated dataset

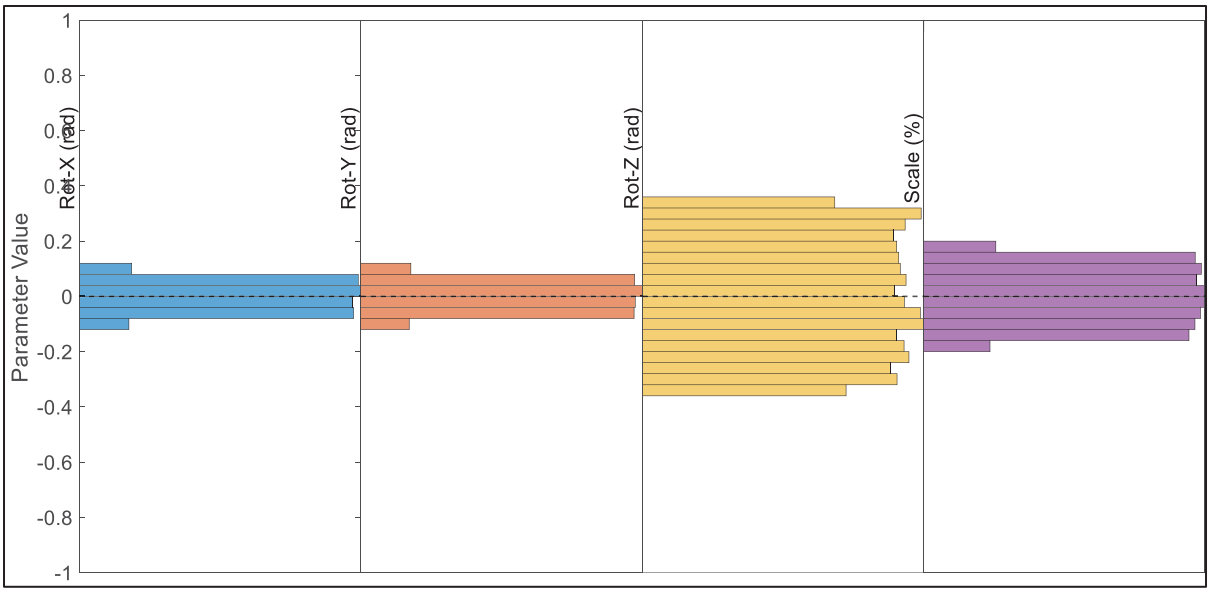

Figure 5.13 Histogram of Pose and Scale parameters in the generated dataset

5.3.5.2 Hyperparameter Optimization

Based on the evaluation of the loss values across all tested configurations, the optimal hyperparameter combination for the 2D/3D registration network was determined as follows:

- Learning rate: 0.15
- Dropout rate: 0.1
- Batch size: 8

Figure 5.14 presents the heatmaps of the test set loss values for all combinations, illustrating the influence of each hyperparameter on model performance. These results were used to select the final configuration for subsequent full training and evaluation of the network.

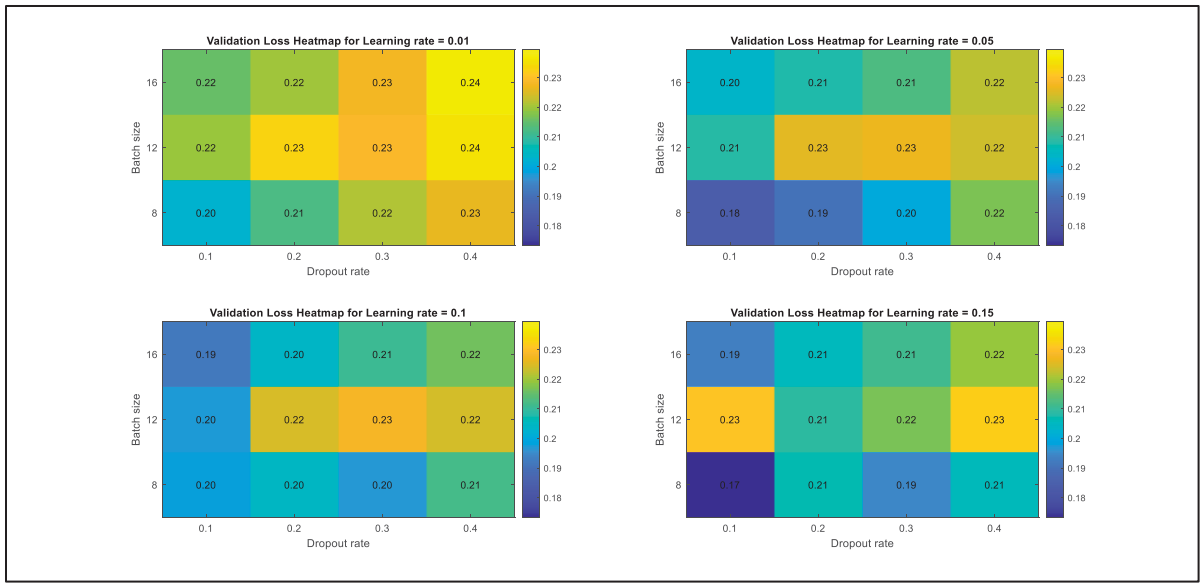

Figure 5.14 2D/3D registration network hyperparameter optimization heatmaps

5.3.5.3 Training, Validation and Testing of the Network

Figure 5.15 illustrates the convergence of the training and validation losses throughout the training process.

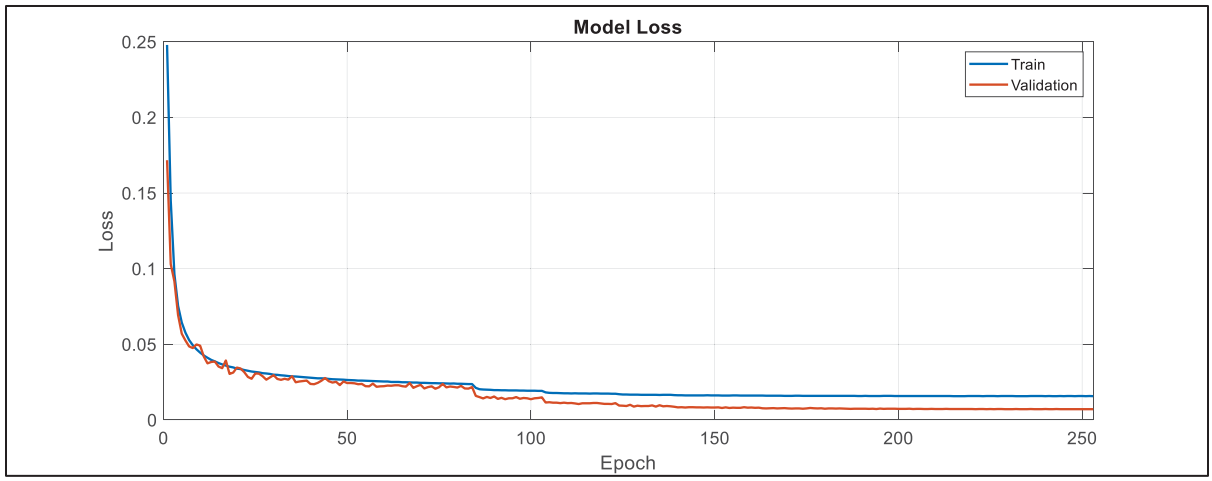

Figure 5.15 Convergence plot of the 2D/3D registration network

The training process concluded at epoch 253 due to the lack of further improvement in validation loss. At this point, the final training loss was 0.0157, and the validation loss reached

0.0071. During training, the learning rate was reduced 11 times, halving at each step, and ultimately reaching 7.32×10^{-5} in the final epoch.

On the test set, the model achieved a loss of 0.018, demonstrating good generalization to unseen data. The distribution of prediction errors for each of the 18 shape parameters is shown in Figure 5.16, while the error distributions for the three pose parameters and the scaling parameter are presented in Figure 5.17. Table 5.2 presents the statistical summary of the error distribution for each parameter.

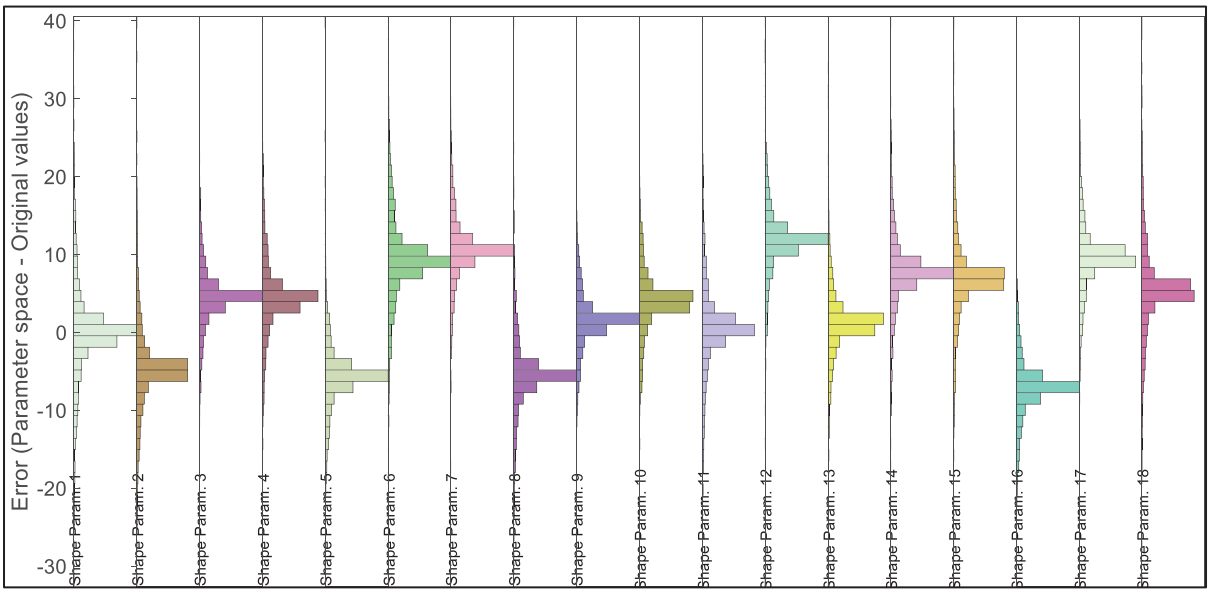

Figure 5.16 Shape parameters error histogram

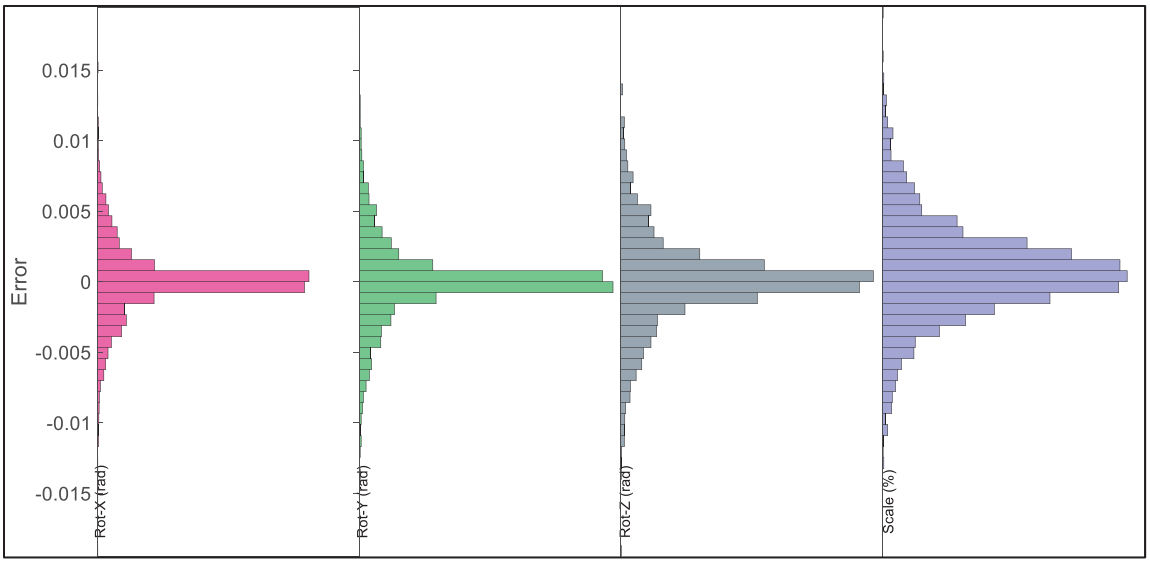

Figure 5.17 Pose and Scale parameters error histogram

Table 5.2 Reconstruction network output parameters error statistics

Parameter	Mean ± STD	Min	Max	Parameter Range
Shape Parameter 1	-0.14 ± 6.75	-32.6118	40.49249	±228.091
Shape Parameter 2	0.22 ± 6.71	-35.036	57.38702	±202.793
Shape Parameter 3	0.02 ± 6.4	-49.8786	48.09783	±198.117
Shape Parameter 4	0.11 ± 4.53	-30.8344	30.1882	±151.834
Shape Parameter 5	-0.19 ± 4.44	-23.7041	40.17647	±140.541
Shape Parameter 6	0.11 ± 3.26	-20.6052	15.38627	±115.319
Shape Parameter 7	0.09 ± 3.47	-31.8972	22.26545	±102.671
Shape Parameter 8	0.07 ± 3.07	-16.4077	27.82063	±85.299
Shape Parameter 9	0.04 ± 2.62	-19.6449	22.45785	±79.202
Shape Parameter 10	-0.08 ± 2.55	-18.9686	18.91436	±76.273
Shape Parameter 11	0.005 ± 2.23	-11.5804	13.97083	±74.669
Shape Parameter 12	-0.04 ± 2.25	-19.5303	12.599	±65.850
Shape Parameter 13	-0.01 ± 2.14	-13.647	15.96725	±61.310
Shape Parameter 14	0.06 ± 2.12	-15.93	12.9922	±59.214
Shape Parameter 15	-0.004 ± 2.06	-15.1865	12.98679	±51.602

Parameter	Mean ± STD	Min	Max	Parameter Range
Shape Parameter 16	-0.03 ± 1.94	-10.0156	18.62456	±52.835
Shape Parameter 17	-0.04 ± 1.80	-14.8145	10.83103	±48.695
Shape Parameter 18	0.04 ± 1.79	-12.7798	11.84031	±45.794
Rotational Parameter X	0.0917 ± 0.1176	4.7539e-05	1.1159	±5
Rotational Parameter Y	0.1041 ± 0.1372	6.9484e-06	1.1209	±5
Rotational Parameter Z	0.2018 ± 0.2267	7.3751e-06	1.8530	±20
Scale Parameter	0.32 ± 0.31	3.7156e-04	2.50	±17

5.3.5.4 Evaluation of the 3D reconstruction accuracy of the Shape

The average 3D reconstruction error across all vertices and all test shapes was 0.173 ± 0.159 mm (mean ± standard deviation). The minimum and maximum reconstruction errors observed were 0.0001 mm and 2.80 mm, respectively. Figure 5.18 illustrates the histogram of reconstruction errors aggregated across all vertices and all test instances.

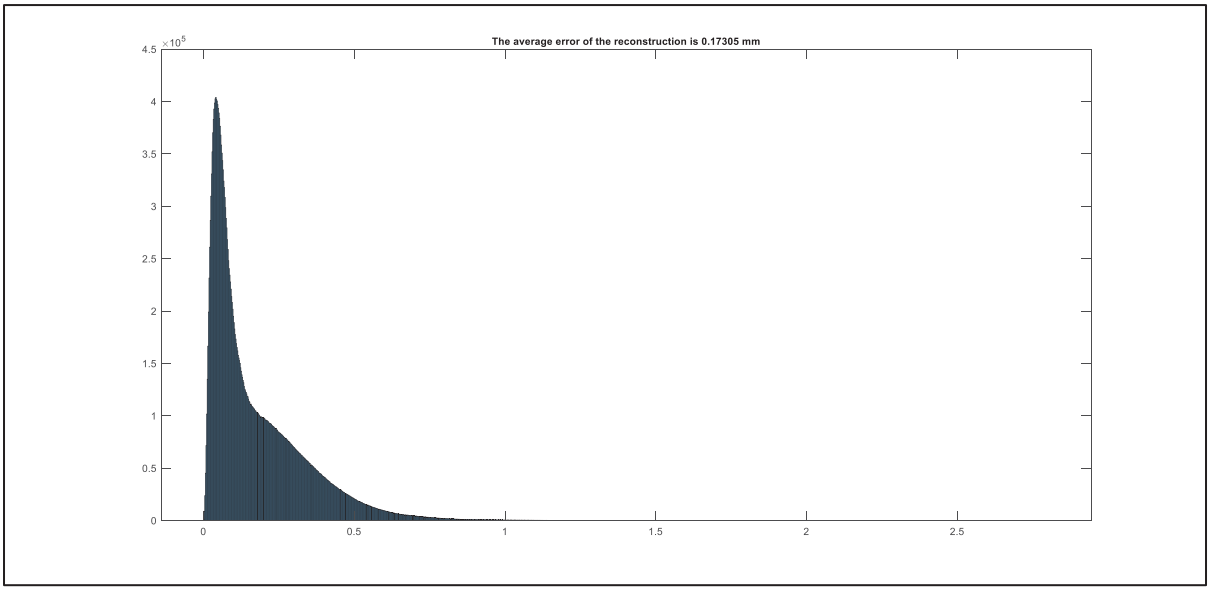

Figure 5.18 Histogram of the point-to-point reconstruction error for the test set

To qualitatively assess the reconstruction, Figure 5.19 shows an example test case, including the AP and LT radiographs, a comparison between the predicted and ground truth values for

the first three shape parameters, as well as the rotational and scaling parameters. The 3D visualization includes the mean mesh (black), the network-generated mesh (blue) after scaling and rotation, and the ground truth mesh (red). This case corresponds to an average reconstruction error of 0.075 mm, demonstrating the network's precision in deforming the mean shape in a single forward pass.

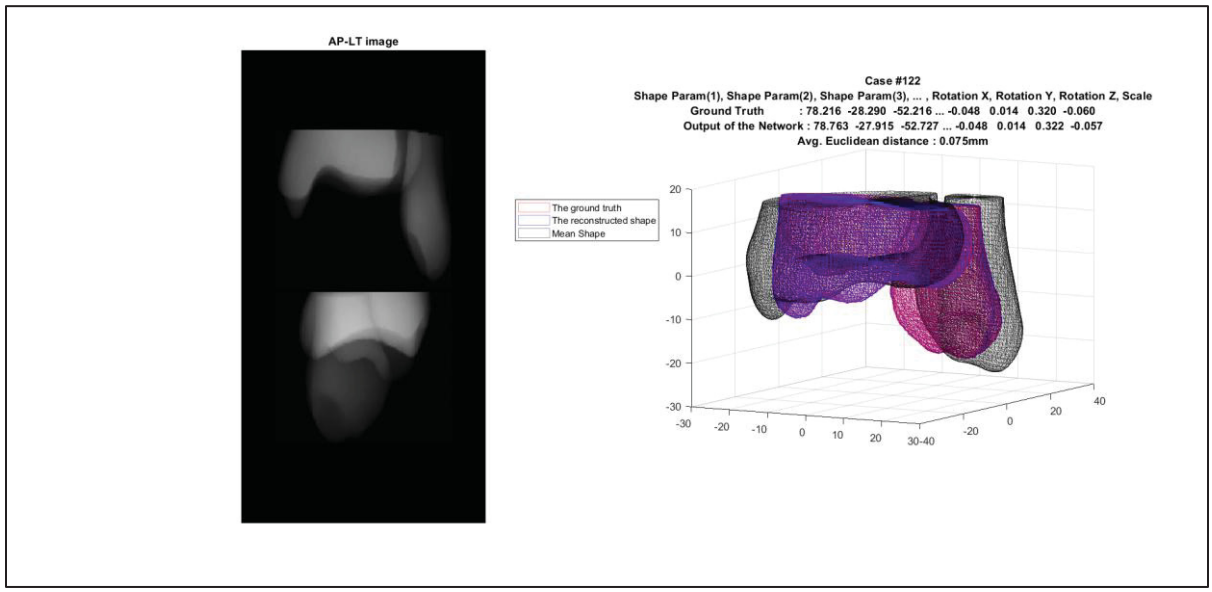

Figure 5.19 Input image of the case #122 from the test set, along with its reconstructed surface (in blue) using the predicted parameters, ground truth mesh (in red) and the mean mesh of the SSIM (in black).

Additionally, Figure 5.21 shows distance map visualizations for the test samples with the highest and lowest reconstruction errors, having average vertex-wise errors of 1.149 mm and 0.016 mm, respectively.

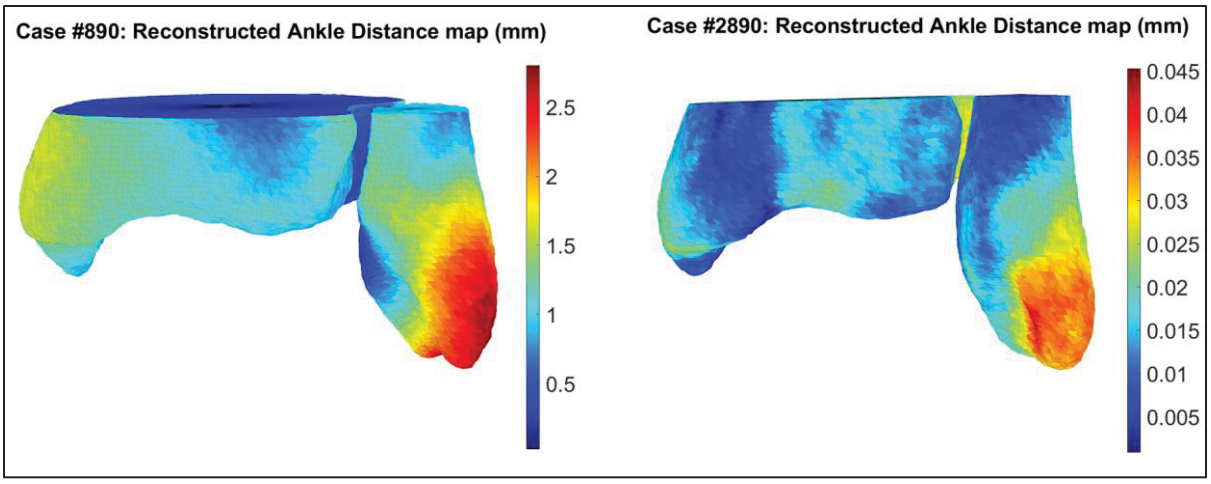

Figure 5.20 Distance map for the error reconstruction of the case #890 with the largest error and case #2890 with the smallest error in the test set

5.3.6 Isolation of the Tibiofibular Joint skeletal structure in biplanar radiographs

In this section, we present the training and testing results of two JE-Nets. They are trained independently from the other stages of the pipeline.

5.3.6.1 Generated Dataset for Joint Extraction Networks

The distribution of randomly selected rotational parameters is shown in Figure 5.21.

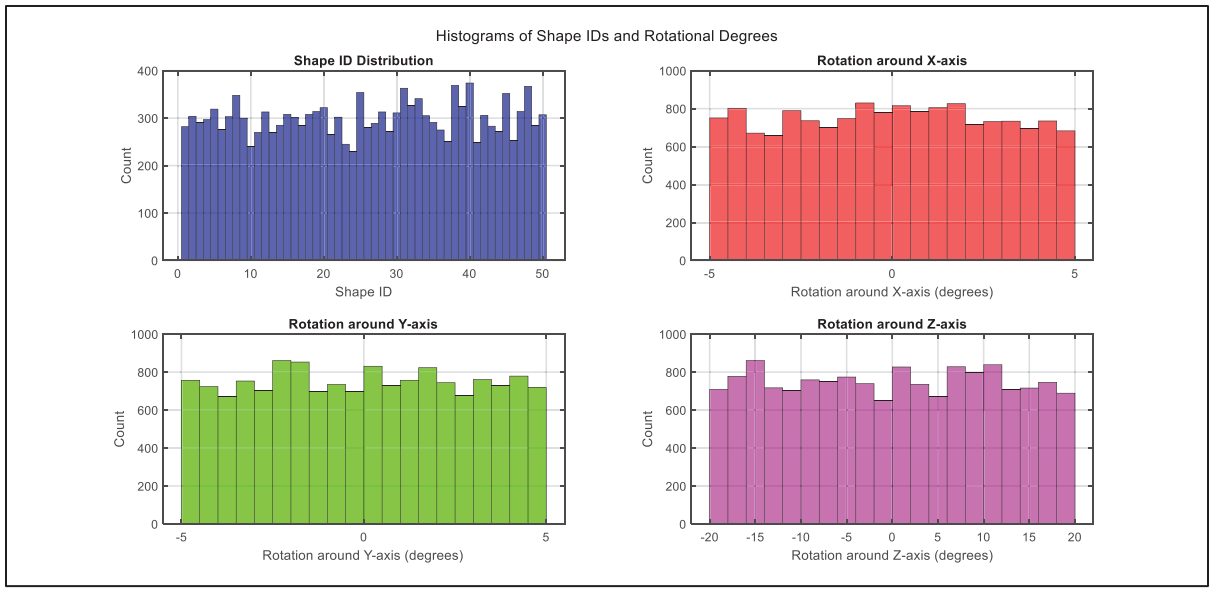

Figure 5.21 Distribution of the selected shape ID and rotational degrees around axes X, Y and Z for generating dataset for JE-Nets

Sample training and testing images for both the AP and LT networks are presented in Figure 5.22, Figure 5.23, Figure 5.24 and Figure 5.25. These samples illustrate the consistency of image quality and anatomical alignment across the foot and joint radiographs.

Figure 5.22 Sample train data for JE-Net-LT

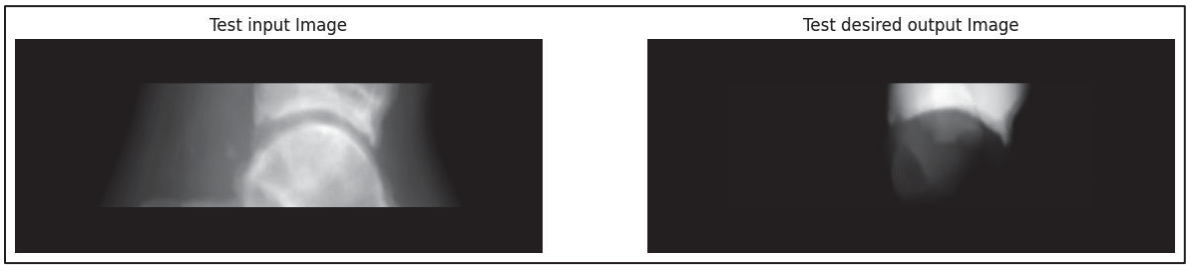

Figure 5.23 Sample test data for JE-Net-LT

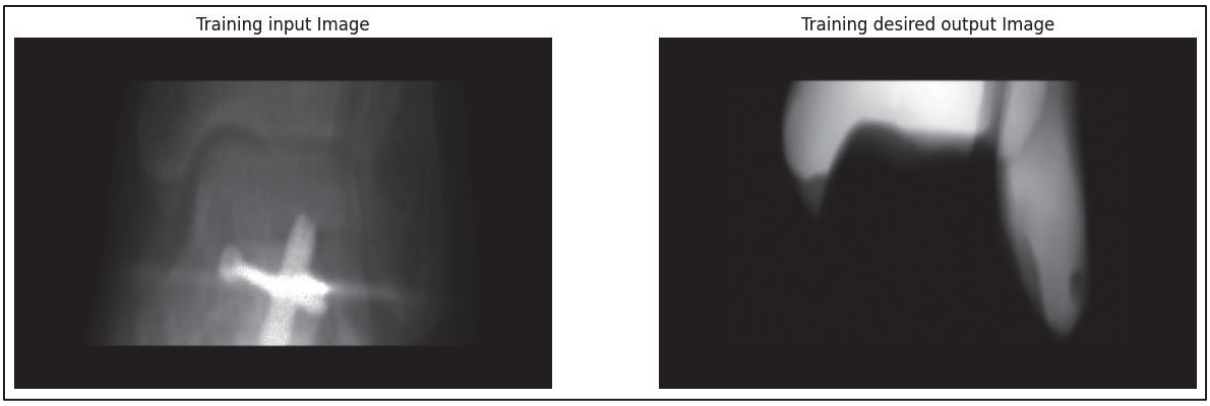

Figure 5.24 Sample train data for JE-Net-AP

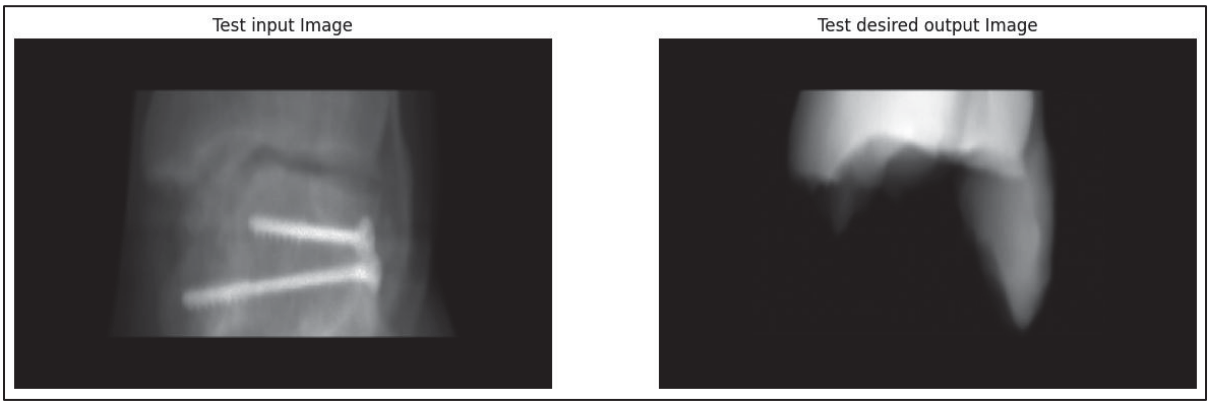

Figure 5.25 Sample test data for JE-Net-AP

5.3.6.2 Hyperparameter Optimization

Based on PSNR evaluation across all hyperparameter configurations, the following combination was found to yield the best performance for both JE-Net-AP and JE-Net-LT:

- Learning rate: 0.001
- Dropout rate: 0.3
- Batch size: 8
- Alpha: 0.9

Figure 5.26 presents the PSNR heatmaps for JE-Net-AP, while Figure 5.27 shows the corresponding results for JE-Net-LT. These visualizations demonstrate how different hyperparameter combinations influenced performance, ultimately supporting the selected

configuration for final training. A detailed summary of all hyperparameter combinations tested for the AP and LT network optimization is provided in ANNEX II and ANNEX III.

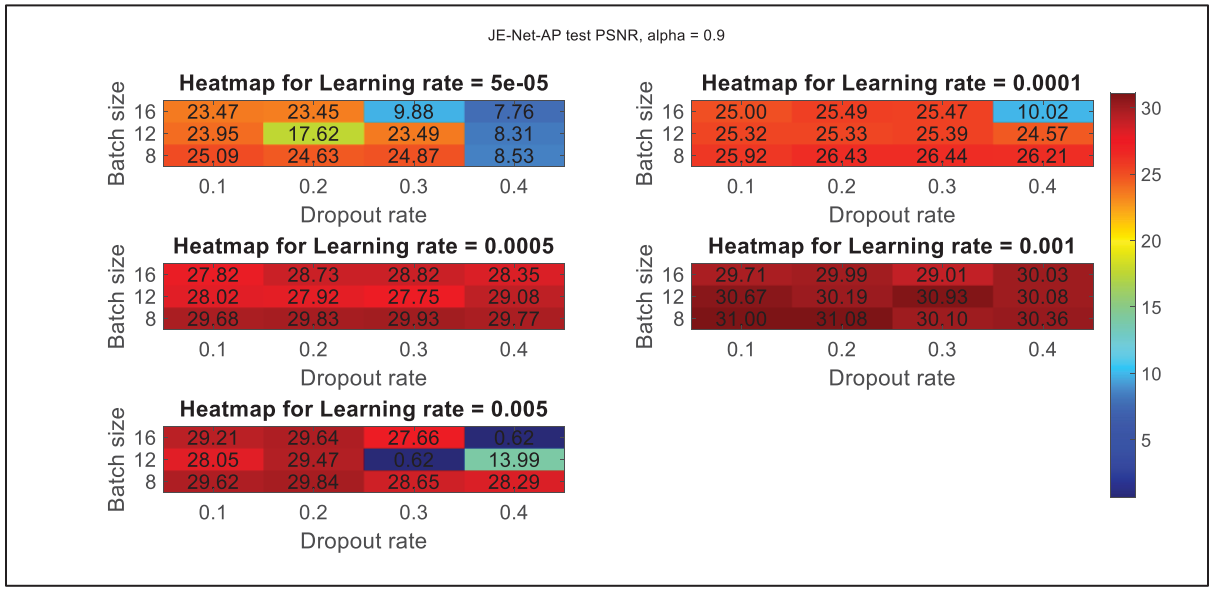

Figure 5.26 JE-Net-AP hyperparameter optimization heatmaps for alpha=0.9

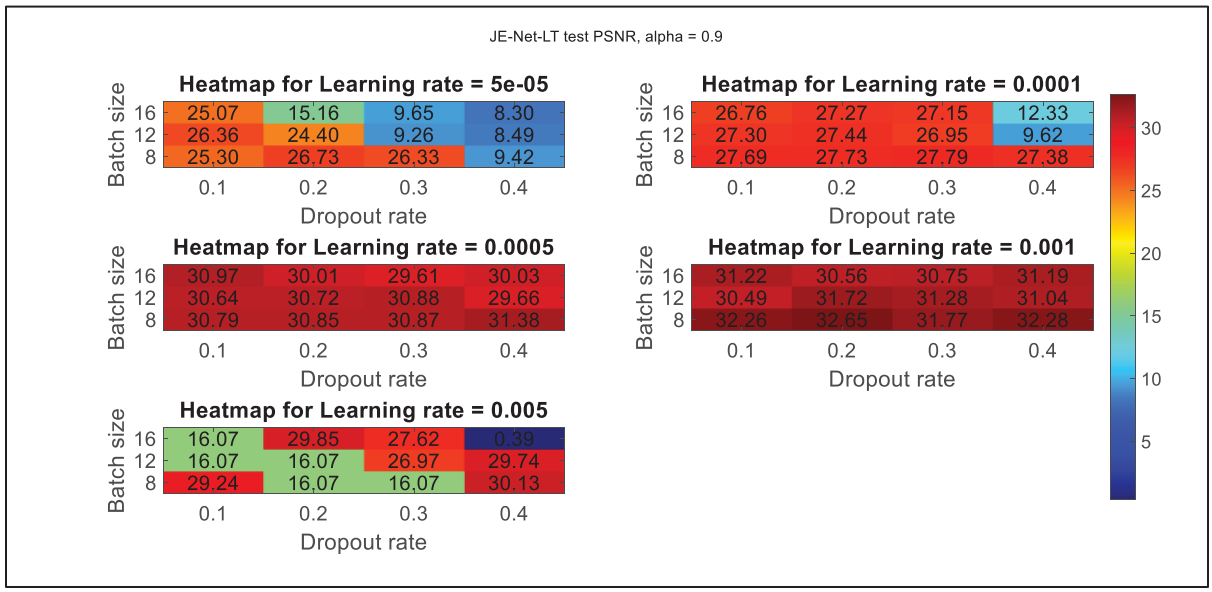

Figure 5.27 JE-Net-LT hyperparameter optimization heatmaps for alpha=0.9

5.3.6.3 Training, Validation and Testing the Joint Extraction Networks

Figure 5.28 and Figure 5.29 illustrate the convergence plots for JE-Net-AP and JE-Net-LT, respectively, showing the evolution of training and validation loss across epochs.

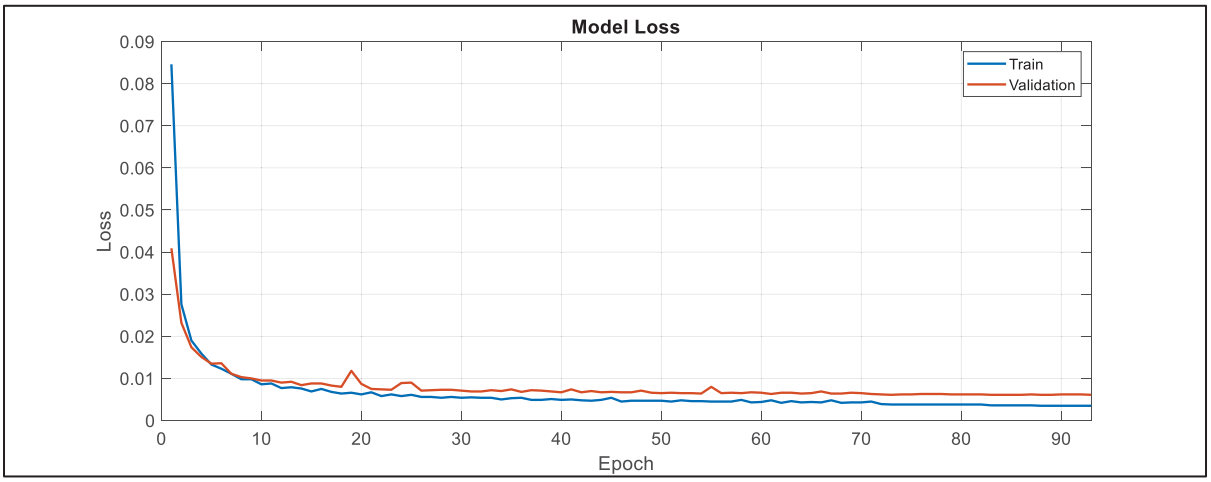

Figure 5.28 Loss convergence plot for JE-Net-AP

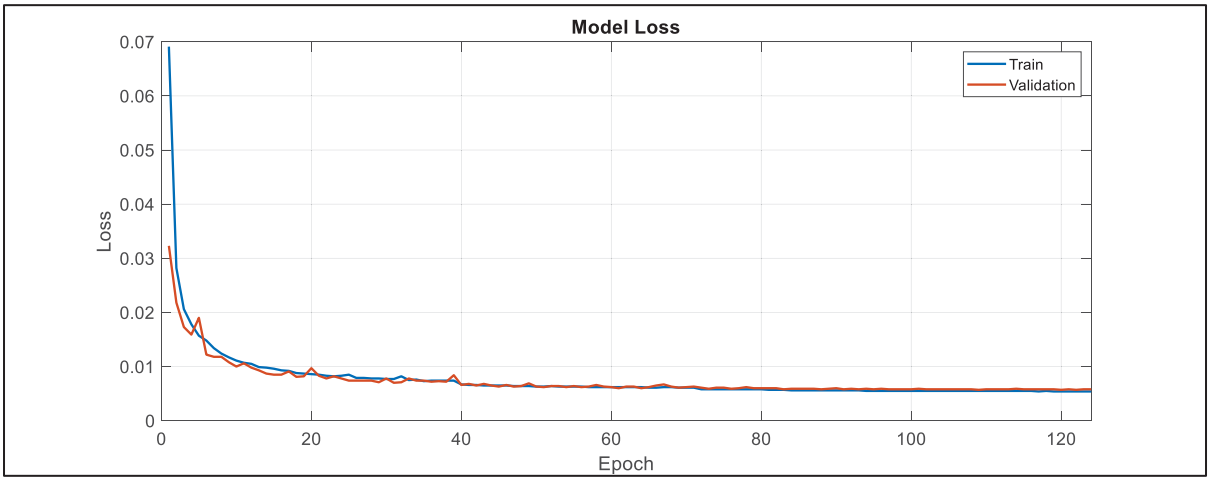

Figure 5.29 Loss convergence plot for JE-Net-LT

For JE-Net-AP, training concluded at epoch 93 due to lack of improvement in validation loss. The final training loss was 0.0035, while the validation loss reached 0.0061. The learning rate was reduced twice, halving each time, and reached a final value of 2.5×10^{-4}. On the test set,

the model achieved a loss of 0.0048 and an average PSNR of 38.75. Figure 5.30 presents an example test image, the network's output, and the corresponding ground truth.

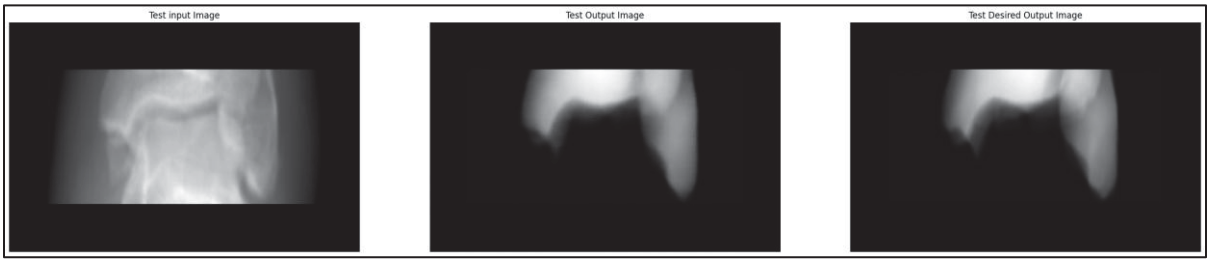

Figure 5.30 Sample test input image, the output of the JE-Net-AP and the ground truth

For JE-Net-LT, training concluded at epoch 124, with a final training loss of 0.0054 and validation loss of 0.0058. The learning rate was reduced six times, reaching a final value of 1.56×10^{-5}. On the test set, the model achieved a loss of 0.0051 and a PSNR of 38.86. Figure 5.31 displays the output of the LT network for a representative test sample, alongside the ground truth.

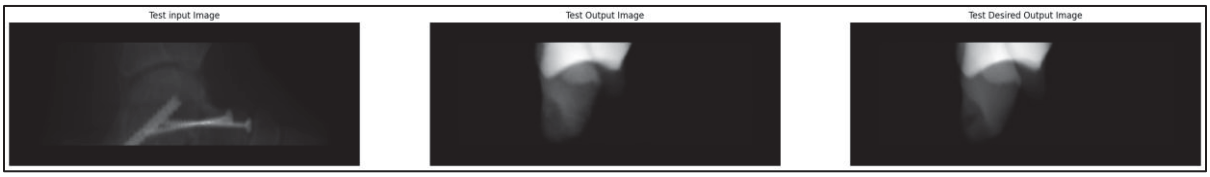

Figure 5.31 Sample test input image, the output of the JE-Net-LT and the ground truth

5.3.7 Establishing the 2D/3D Reconstruction Pipeline

Figure 5.32 shows the distribution of the shape IDs used to generate the augmented dataset, along with the range of rotational degrees applied around the X, Y, and Z axes.

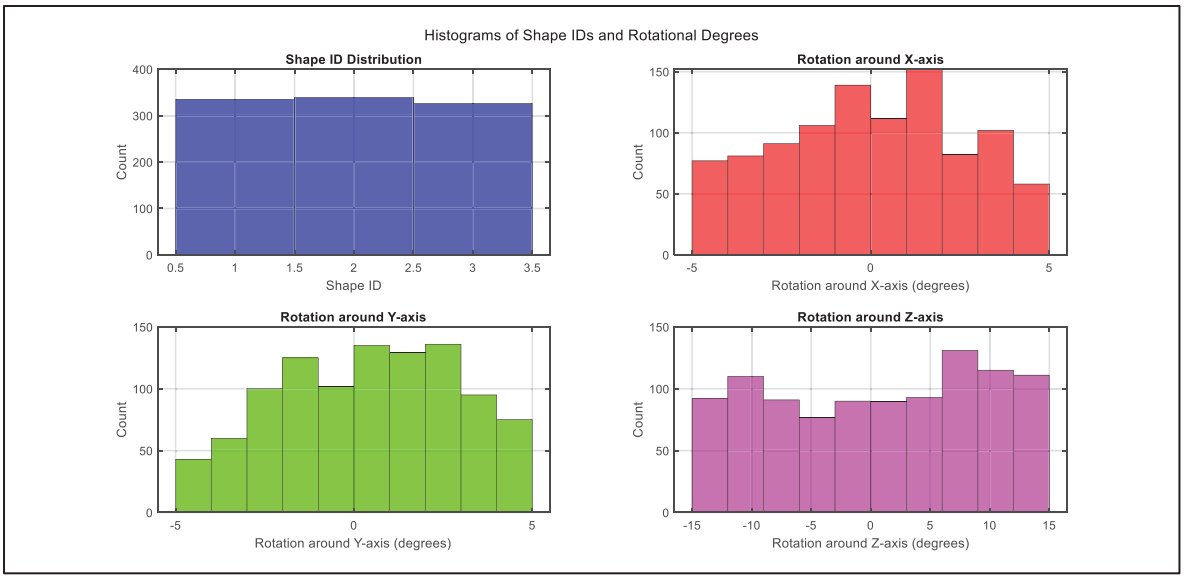

Figure 5.32 Distribution of the shape ID used to generate the evaluation dataset along with the used rotational degrees around axes X, Y and Z

Table 5.3 presents the statistical summary of reconstruction errors across the three original CT cases and the 1,000 augmented instances. In addition, Figure 5.33 shows the histogram of the error distribution across vertices for the 1,000 test cases, illustrating the spatial spread of reconstruction accuracy across the dataset.

Table 5.3 Reconstruction error statistics for Evaluation of the pipeline

Parameter	Mean ± STD (mm)	Min (mm)	Max (mm)
Left-out instance 1	0.55 ± 0.34	0.00	2.14
Left-out instance 2	0.64 ± 0.53	0.00	3.13
Left-out instance 3	0.83 ± 0.51	0.00	2.76
All Left-out instances	0.68 ± 0.49	0.00	3.13
1,000 rotated instances	0.75 ± 0.56	0.00	4.79

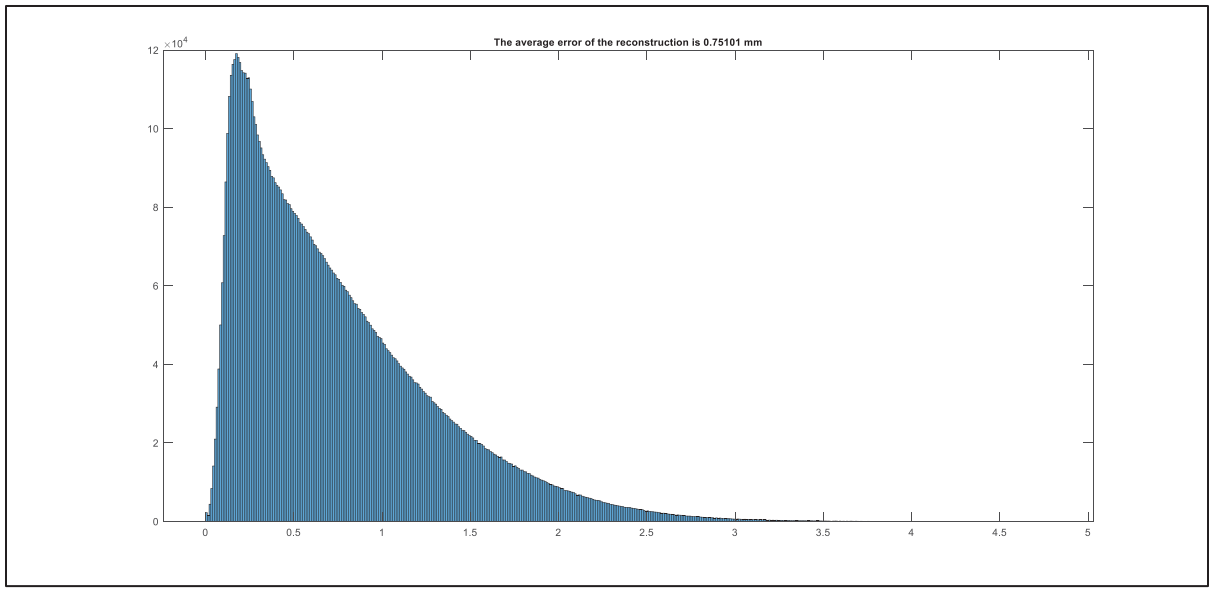

Figure 5.33 Histogram of the point-to-surface reconstruction error for the reconstruction pipeline evaluation set

5.3.8 Reproducibility Error on Selection of the Region of Interest

Figure 5.34 presents the distribution of the noise values applied to both landmarks across the dataset.

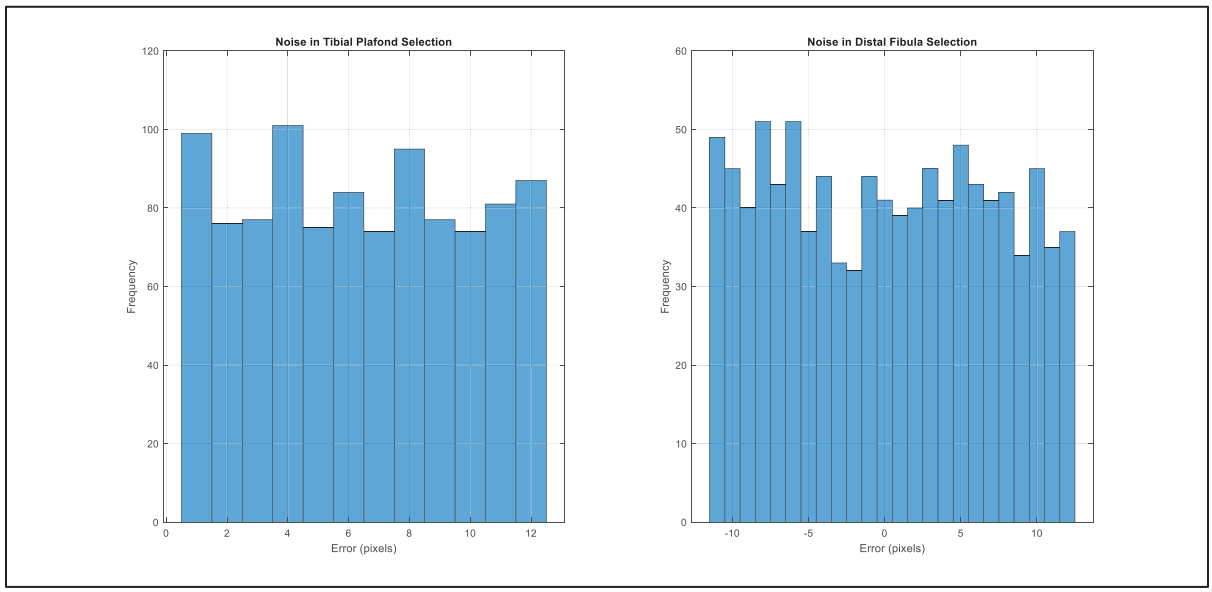

Figure 5.34 Histogram of the noise selection for the distal fibula and tibial plafond

The effect of ROI selection noise on 3D reconstruction performance was found to be minimal.

Table 5.4 compares the reconstruction errors between the perfect and noisy ROI selection scenarios. On average, introducing noise resulted in a 0.04 mm increase in the reconstruction error, indicating strong robustness of the pipeline to ROI localization variation.
Figure 5.35 shows the histogram of the vertex-wise reconstruction errors for all cases in the noisy ROI dataset.

Table 5.4 Reconstruction error statistics for Evaluation of the pipeline with noisy ROI selection and the difference to the perfect ROI selection

Parameter	Mean ± STD (mm)	Min (mm)	Max (mm)	Difference in Mean ± STD
Left-out instance 1	0.58 ± 0.39	0.00	2.21	0.03 ± 0.05
Left-out instance 2	0.67 ± 0.54	0.00	3.39	0.03 ± 0.01
Left-out instance 3	0.93 ± 0.52	0.00	2.79	0.1 ± 0.01
All Left-out instances	0.72 ± 0.49	0.00	3.39	0.04 ± 0.00
1,000 rotated instances	0.79 ± 0.57	0.00	6.12	0.04 ± 0.01

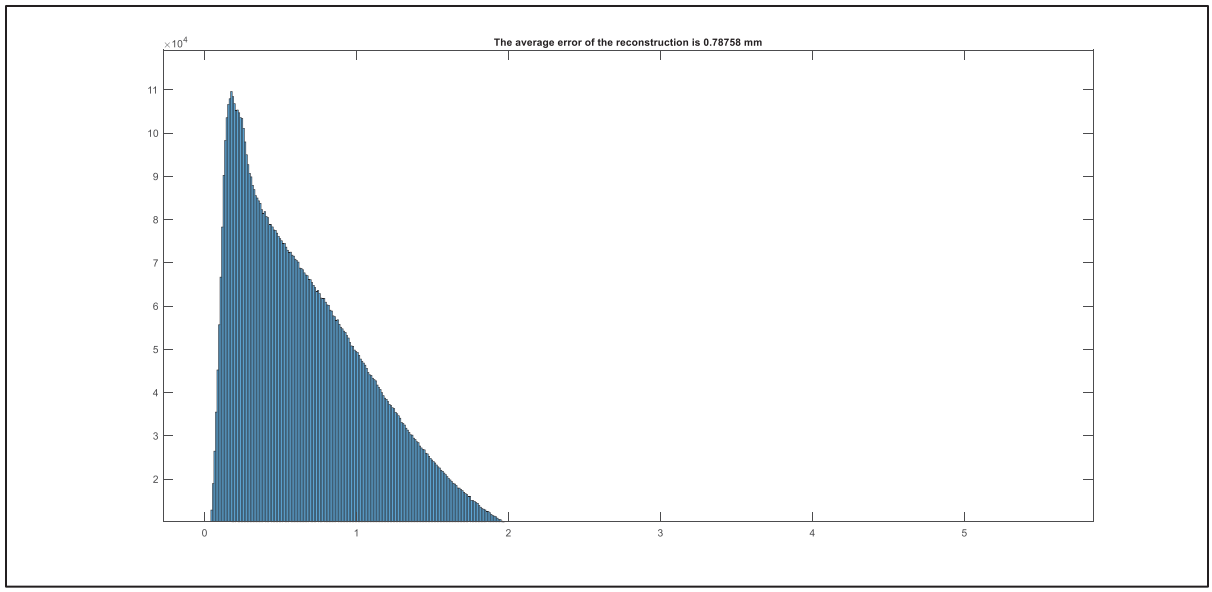

Figure 5.35 Histogram of the point-to-surface reconstruction error for the reconstruction pipeline evaluation set having noisy ROI selection

5.3.9 Syndesmosis Clinical Parameters calculation on reconstructed 3D model

Table 5.5 presents the comparison between the clinical parameters extracted from the reconstructed meshes and the corresponding ground truth values provided by the two expert observers. The table also reports the variability between experts, the absolute difference between the automatic and reference values, and the relative error percentage for each parameter. This combined presentation highlights both the reliability of expert-derived ground truth and the accuracy of the automatic pipeline.

Across most linear parameters (A, B, C, D, E, and F), the relative error remained below 20%, with several values, some showing errors close to or below 15%. These results indicate that the automatic reconstruction preserved the relative proportions of the tibiofibular joint with high fidelity. The ratio D/E exhibited higher error levels (~55%), reflecting the sensitivity of ratio-based metrics to small variations in absolute distances. Similarly, the angular measurements displayed larger variability. Angle1 showed a wide error distribution, strongly influenced by interobserver variability, while Angle2 demonstrated more stable estimates, with an error magnitude similar to that observed between the expert measurements.

In conclusion, the results demonstrate that the proposed reconstruction pipeline is capable of preserving clinically significant anatomical features with acceptable accuracy. For most parameters, the error remained within a range that is consistent with clinical relevance, typically less than 20%, while highlighting specific cases where angular or ratio-based measures are more affected by reconstruction variability.

Table 5.5 Clinical parameters measurement error of three evaluation cases

Parameter	Interobserver Variability (mean ± SD)	Experts' relative variability (%) (mean ± SD)	Reference (mean ± SD)	Auto vs. Reference (mean ± SD)	Relative error (%) (mean ± SD)
A (mm)	0.900 ± 0.624	27.83 ± 24.62	3.883 ± 1.333	0.649 ± 0.465	16.73 ± 13.45
B (mm)	0.867 ± 1.069	10.40 ± 11.43	7.800 ± 2.035	1.076 ± 1.113	17.54 ± 21.78
A/B	0.086 ± 0.025	19.28 ± 12.04	0.511 ± 0.167	0.059 ± 0.016	12.41 ± 4.69
B – A (mm)	0.500 ± 0.100	14.66 ± 5.66	3.917 ± 1.994	0.930 ± 0.322	30.73 ± 25.36
C (mm)	0.600 ± 0.624	23.29 ± 26.44	3.067 ± 0.785	0.437 ± 0.312	13.98 ± 9.16
D (mm)	0.500 ± 0.500	4.78 ± 4.41	10.78 ± 1.693	1.545 ± 0.931	15.50 ± 11.91
E(mm)	1.367 ± 0.306	19.54 ± 2.58	6.950 ± 0.794	1.268 ± 0.956	18.49 ± 13.54
D/E	1.867 ± 0.379	49.82 ± 10.18	3.833 ± 0.900	2.171 ± 0.770	55.45 ± 8.98
F (mm)	0.433 ± 0.404	16.93 ± 17.20	2.783 ± 0.621	0.419 ± 0.470	17.05 ± 19.34
Angle1 (°)	2.883 ± 2.232	35.56 ± 62.98	1.175 ± 8.213	6.164 ± 2.559	280.5 ± 476.46
Angle2 (°)	4.380 ± 2.867	94.38 ± 60.69	4.627 ± 0.125	4.676 ± 1.017	101.11 ± 22.18

CHAPITRE 6

DISCUSSION

6.1 Overview

This chapter provides an in-depth discussion of the findings presented in Chapter 5, analyzing the effectiveness of the proposed methodology for 3D reconstruction of the tibiofibular joint from biplanar radiographs and calculating the syndesmosis clinical parameters. The implications of the experimental results are examined in the context of medical imaging and clinical applications. Additionally, we discuss the strengths and limitations of the proposed approach and highlight potential areas for future research and improvement.

6.2 Key Findings

6.2.1 Dataset, Segmentation and Meshing performance

Although the model and the dataset for JE-Nets built on fifty cases demonstrated strong performance based on quantitative metrics, increasing the number of instances could enhance the generalization of the pipeline to unseen cases. In addition, the number of cases available for pipeline evaluation was limited. While we were able to generate numerous instances with varying rigid transformations, the dataset lacked sufficient representation of non-rigid transformations. This limitation restricts the model's ability to generalize to cases involving anatomical variations or deformities. Future work should focus on incorporating a broader range of non-rigid transformations to improve the robustness and adaptability of the reconstruction pipeline.

The segmentation process was not quantitatively evaluated due to the lack of ground truth data; instead, a qualitative assessment was performed on three randomly selected cases. Given the constraints in ground truth availability, this was the most feasible approach, and the evaluation conducted by an expert provided valuable insights into segmentation quality.

The meshing process preserves the CT mask details at its finest resolution. However, simplifying the mesh introduces some level of error in 3D space between the dense and simplified versions. This trade-off must be carefully considered, as denser meshes increase the computational load throughout the entire pipeline, including instance generation, volume creation, and reconstruction error calculation. In this study, we found that the error introduced by mesh simplification remained within an acceptable range while significantly improving computational efficiency.

The matching process also plays a critical role in the reconstruction pipeline. Since ground truth data for matching was not available, and manually annotating correspondences across the dataset would have been an extremely time- and labor-intensive task, an indirect evaluation approach was adopted, where the quality of the resulting statistical shape model served as an indicator of matching accuracy. The consistency of the generated models suggests that the matching process was effective.

6.2.2 Statistical Shape and Intensity Model

The SSIM was designed to capture the natural variability in tibiofibular joint anatomy. Due to the presence of multiple overlapping anatomical structures in ankle radiographs, incorporating prior knowledge that encodes the variability of the shape of interest improved the reconstruction process by providing a structural guide. Additionally, modeling the multi-object tibiofibular joint through concatenation avoids the complexity of articulated or parametric models, which require explicit geometric modeling of joint kinematics, while still preserving the spatial relationships between the tibia and fibula. The accuracy of the calculated syndesmosis clinical parameters ‘a’, ‘b’, and ‘c’, which represent inter-bone distances, serves as a validation of the model’s ability to encode joint relationships effectively. Furthermore, using simple intensity information reduces computational overhead during volume reconstruction, a trade-off that is compensated by integrating JE-Net to comply with the simple intensity modeling.

The evaluation metrics confirm that the model effectively captures the underlying shape variability, exhibiting strong performance across all key criteria: compactness (Figure 5.8),

generalization (Figure 5.9), specificity (Figure 5.10), and accuracy (Figure 5.11). Using the first 18 modes of variation, the model retains 90% of the total shape variance, demonstrating high compactness. The generalization error is approximately 0.8 mm, while the specificity and accuracy errors are 1.4 mm and 0.5 mm, respectively, all within acceptable ranges reported in the literature for anatomical modeling. However, expanding the training dataset could further improve its adaptability, allowing it to better accommodate unseen anatomical variations.

Modeling intensity in a more detailed manner, such as using tetrahedral meshing within the volume, would enhance the model's realism by providing a more anatomically accurate representation of internal bone density. However, this approach introduces significant challenges, including the complexity of matching tetrahedral elements across datasets and the increased computational overhead associated with processing the model. Some repeating steps in the pipeline, such as generating new instances and creating volumes out of that would become more computationally intensive with a richer intensity model.

To balance accuracy and efficiency, we opted for a simplified approach, using a single intensity value to represent bone density while compensating for lost information through pixel altering process by JE-Nets. These networks adjusted pixel intensities in the AP and LT images to better reflect the model projections, ensuring that it maintained its effectiveness while keeping computational demands manageable. This trade-off allowed for efficient processing while still preserving the information needed for the registration process.

6.2.3 2D/3D Registration Network

This work integrates a deep learning architecture capable of reconstructing the model in a single forward pass, thereby eliminating the need for iterative optimization and effectively addressing the local minima issues commonly associated with traditional optimization-based methods.

The registration network exhibited reliable alignment between the 3D model and the input radiographs, with point-to-point reconstruction accuracy averaging 0.173 mm. This level of accuracy is sufficient for clinical applications, as the inter-bone distances in the tibiofibular joint exceed 1 mm. However, we have to consider that an average error of 0.22 mm due to

mesh simplification is present between the model and the dense meshes, representing the CT masks. Still, we have lower reconstruction accuracy than 1 mm.

This network leverages the advantage of directly reconstructing a mesh rather than a volumetric representation (Kasten et al., 2020), thereby facilitating the calculation of clinical parameters on the resulting 3D model and eliminating the additional step of anatomical landmark identification on a reconstructed volume. In addition, incorporating the SSIM enhances the generalizability of the reconstruction process by explicitly modeling shape variability within the pipeline. In contrast, learning such variability directly from volumetric data (Kasten et al., 2020), requires a large number of CT scans that is often limited or unavailable in clinical practice. To the best of our knowledge, no prior study has estimated shape parameters directly from radiographs, making this the first work to do so (Hashemibakhtiar et al., 2024). However, direct comparison with existing literature is limited, as previous studies typically reconstruct models from complete radiographs that include additional anatomical structures and soft tissues, which differ from the isolated joint-focused radiographs used in our approach.

Similar to the JE-Nets, the hyperparameter optimization process was incorporated to identify the optimal combination of parameters for the registration network. By systematically testing different configurations, this approach aimed to enhance the network's performance and stability. Through a controlled search strategy, we ensured that the registration network achieved the best possible accuracy while maintaining computational efficiency.

6.2.4 Joint Extraction Networks

The JE-Nets successfully extracted the tibiofibular joint from ankle biplanar radiographs, with evaluation metrics such as PSNR and loss values between the output and the ground truth, indicating stable convergence. This demonstrates the effectiveness of the network in accurately extracting the joint structures from radiographic images, while modifying the intensity values of the pixel to reflect the information stored in the SSIM.

Since the training dataset was generated using segmented CT volumes, it was limited to fifty distinct cases, each representing different non-rigid variations of the tibiofibular joint. To

compensate for this limitation, dataset augmentation was performed by introducing rotational variations of the volumes. However, while this approach increased the dataset size, it did not fully capture the diversity of real-world anatomical variations. Expanding the dataset with additional cases, would improve the generalization capability of the model and enhance its robustness for clinical applications.

The hyperparameter optimization process using grid search is computationally intensive but allows for a thorough exploration of different hyperparameter combinations. By systematically testing multiple configurations, this approach helps identify the optimal settings for improved model performance.

To manage the high computational cost associated with a large number of hyperparameter combinations, we addressed this challenge by reducing the dataset size and limiting the number of training epochs. This strategy enabled a more efficient search process while still providing meaningful insights into the impact of different hyperparameter choices.

6.2.5 Evaluation of the Integrated Reconstruction Pipeline

After training all the networks, a complete pipeline was established, integrating these components to process AP and LT radiographs of the ankle and reconstruct a 3D model of the tibiofibular joint. Given that the test cases were previously unseen by all networks, we anticipated an increase in reconstruction error compared to the individual 2D/3D registration network. Additionally, the non-perfect outputs of the JE-Nets were expected to slightly impact reconstruction accuracy.

The pipeline was evaluated three left-out cases and 1,000 instances augmented from the three instances to assess the rotational robustness of the pipeline. Having only three distinct CT scans limited the variability in bone shapes, while we may assess the rotational robustness of the pipeline. Increasing the number of distinct CT cases in the dataset would further enhance the generalizability of the evaluation.

Despite these constraints, the pipeline achieved an average reconstruction error of 0.68 mm, which is less than 1 mm, demonstrating enough accuracy for the tibiofibular joint 3D

reconstruction. The rotational robustness test confirmed that pipeline can handle different poses of the joint as the average reconstruction accuracy increased only for 0.07 mm. The impact of selection of the ROI on reconstruction accuracy was found to be minimal, introducing only an additional 0.04 mm of average error. This result highlights the pipeline's resilience, maintaining reliable reconstruction performance even under imperfect ROI selection conditions. Table 6.1 shows the quantitative comparison of reconstruction accuracy to other works in the literature.

Table 6.1 Comparison of reconstruction accuracy with recent works in the literature

Method	Organ	Error (mm)
Kasten et al. (Kasten et al., 2020)	leg bones	1.29 ± 0.28
Chênes et al. (Chênes & Schmid, 2021)	femur	1.02 ± 0.89
Aubert et al. (B. Aubert et al., 2019)	spine, vertebra	2 ± 1.4
Tchinde Fotsin et al. (Tchinde Fotsin et al., 2019)	knee	0.8 ± 0.3
Houtte et al. (Van Houtte, Audenaert, Zheng, & Sijbers, 2022)	femur	1.29 ± 0.21
Roth et al. (Arn Roth et al., 2024)	proximal tibia	1.42 ± 0.56
Pan et al. (Pan et al., 2023)	tibia and fibula	1.16 ± 0.12
Ours	ankle (DTF)	0.68 ± 0.49

6.2.6 Clinical Parameter Estimation

The reconstructed 3D models in the evaluation set were used to compute syndesmosis clinical parameters, which were then compared against ground truth values annotated by expert physicians. As shown in Table 5.5, the proposed approach was able to reproduce clinically relevant parameters with acceptable accuracy, with most linear measures exhibiting relative errors below 20%. This level of accuracy indicates that the reconstruction pipeline is sufficiently robust to preserve anatomical relationships that are critical for syndesmosis

assessment, supporting its potential as a complementary tool to traditional CT-based evaluations.

Among the assessed parameters, Angle 1 exhibited the highest variability. This outcome can be attributed to the sensitivity of angular definitions to minor vertex displacements during reconstruction, where even small shifts in landmark positions lead to large differences in angular values. By contrast, Angle 2 demonstrated greater robustness. Because it is defined using eight vertices to establish the reference lines, the redundancy reduces sensitivity to local variations and yields error levels comparable to those observed between expert observers. This suggests that angular parameters can be used in reconstruction-based assessments, but their reliability depends strongly on the geometric definitions employed.

Importantly, the parameters most relevant to clinical diagnosis, A, B, and C, were reproduced with an accuracy that falls within clinically acceptable ranges. Maintaining reliability in these key indicators is essential for the method's applicability in real-world syndesmosis evaluations. While some ratio-based and angular parameters showed higher errors, their contribution is secondary compared to the primary measurements of tibiofibular spacing and alignment.

One limitation of the present study is that the dataset consisted exclusively of non-pathological cases. As a result, the reported accuracy may not directly translate to pathological conditions where bone alignment is altered. Future work should therefore extend validation to both pathological and non-pathological cohorts, ideally using CT scans obtained in standardized foot positions. Such inclusion would enhance the robustness and clinical relevance of the pipeline, ensuring that it can support comprehensive assessments of syndesmosis stability in diverse patient populations.

6.3 Strengths of the Proposed Methodology

We outline the following strength of the proposed methodology.

- High Accuracy and Robustness: The deep learning-based registration network effectively integrates rich shape information in the SSIM with one-shot learning approaches, enabling highly accurate 2D/3D reconstruction. By leveraging deep learning techniques, the model optimizes both shape and pose parameters, ensuring a

reliable and precise reconstruction process. In comparison to the recent works in 3D reconstruction from biplanar images (Table 6.1), our method has lowest reconstruction error.

- Extendable to other Anatomical Structures: Since the model is based on triangular meshes and is not specifically limited to the tibiofibular joint, the registration approach can be adapted for other anatomical structures. This is particularly beneficial for regions with tightly connected shapes, where the reconstruction error remains low, achieving an accuracy of approximately 0.68 mm. The JE-Nets are also not limited to the ankle joint and can be applied to extract other anatomical structures from medical images, making it a flexible and adaptable tool for various applications.
- Automation and Efficiency: The integration of AI-driven techniques significantly reduces manual intervention, minimizing processing time and potential errors. Although data augmentation and network training require considerable time, these steps are performed only once, making the operational phase highly efficient and streamlined. Choosing a simplified intensity representation reduces the model's complexity and processing time, while differences between the model's projection and the joint's radiograph are handled through the intensity-altering capabilities of JE-Nets.
- Clinical Applicability: The accuracy of clinical parameter estimation suggests that the method operates with less than 10% error and could be integrated into medical workflows for non-invasive syndesmosis assessment.

6.4 Limitations and Areas of Improvement

The primary limitation of this study is the dataset, which consists of only 53 cases, all of which are non-pathological syndesmosis cases. Additionally, some cases in the dataset were acquired with the foot in dorsiflexion, which, although a minor factor, could slightly alter the relative positioning of the tibia and fibula. Since pathological cases are not represented, the network has not learned the variations associated with injured syndesmosis, which may lead to challenges in reconstructing such cases. Expanding the dataset to include more CT scans,

preferably acquired with the foot in a neutral position and ideally under weight-bearing conditions, would significantly enhance the pipeline's generalizability.

Another limitation is the lack of dataset annotations for segmentation, meshing, and matching processes. The absence of ground truth data at various stages required us to rely on qualitative evaluations instead of more robust quantitative metrics. This could be improved by having expert annotations for segmentation and landmark detection in a subset of cases, enabling a more precise quantitative assessment of segmentation and matching accuracy.

The limited number of cases has also affected the generalization capability of all trained networks. Increasing the dataset size would improve the training process and enable the networks to better handle anatomical variations.

For the assessment of clinical parameters, additional evaluations by multiple experts would provide a better understanding of the reproducibility error, further validating the reliability of the proposed method.

Another limitation is the age diversity of the dataset, as it consists exclusively of adult cases. The networks have not been trained on pediatric cases, but the method can be extended to younger patients by incorporating additional CT scans of children.

Currently, ROI selection in the AP radiograph is performed by manually localizing two landmarks: the distal tibia and the tibial plafond. Automating this process would reduce the reproducibility error and enable a fully automated workflow, enhancing the efficiency of the pipeline.

As noted in the clinical parameter evaluation results, the method may not be fully reliable for parameter Angle 1, as small variations in vertex positioning can lead to large deviations in angle measurements. However, the primary syndesmosis assessment parameters (A, B, and C) were computed with high accuracy.

A technical limitation is to localize the ankle in lateral radiographs, where both ankles are superimposed when a patient stands in the anteroposterior position. To address this, a 45-degree rotation of patient could be introduced to ensure that only one ankle is visible in both AP and LT images. Implementing this adjustment would require modifying the model and volume rotations accordingly and retraining the networks with the newly generated data.

Finally, to obtain statistically valid results for the pipeline, multiple rounds of randomly selecting three cases to be left-out were required. In each iteration, the SSIM is built using the remaining fifty cases, datasets for registration and JE-Nets are generated, the networks are retrained, an evaluation set is created from the three left-out cases, and the entire pipeline is assessed. Repeating this process multiple times would allow for a more generalized assessment of reconstruction accuracy.

However, due to the significant computational time required which is approximately two months per full iteration, it was not feasible to apply this approach within the scope of this study. Future work could explore efficient optimization techniques or leverage higher computational resources to make such an extensive evaluation process more practical.

By addressing these limitations through dataset expansion, annotation improvements, automation, and methodological refinements, the proposed pipeline can be further optimized for broader clinical applications.

CONCLUSION

This thesis presented a novel methodology for reconstructing a personalized 3D model of the tibiofibular joint using biplanar radiographs. A key takeaway from this study is the feasibility of combining the Statistical Shape and Intensity Model (SSIM) and deep learning-based techniques for accurate 3D reconstruction of the tibiofibular joint. The methodology demonstrated a high level of accuracy in reconstructing joint structures, with an average reconstruction error of 0.68 mm, which is sufficient for many clinical applications. The methodology was systematically validated through a series of experiments, demonstrating its potential for clinical applications, particularly in the assessment of syndesmosis injuries.

The key contributions of this research include:

- The integration of a Statistical Shape and Intensity Model using a deep learning A robust registration network that aligns the reconstructed model with input radiographs in a one-shot process, ensuring accurate shape and pose parameters estimation and automatic syndesmosis clinical parameter extraction from the reconstructed model.
- A deep learning-based denoising framework capable of accurately isolate the tibiofibular joint within radiographs.
- The development of an AI-driven 2D/3D reconstruction pipeline that achieves an average reconstruction error of 0.68 mm, which is clinically acceptable for syndesmosis assessment.

Experimental results validated the feasibility of the proposed approach, confirming its accuracy, robustness, and applicability in real-world clinical scenarios. Furthermore, robustness of the methodology was further validated by analyzing its performance under a non-perfect selection of the ROI in the ankle AP radiograph. The results showed that the system maintained an acceptable level of accuracy even when the ROI selection was suboptimal, indicating the resilience of the approach to minor input variations. This feature is crucial for real-world clinical applications, where variability in image acquisition and manual selection errors is inevitable.

Furthermore, the automatic calculation of syndesmosis clinical parameters confirmed that the calculated parameters on the reconstructed model aligns well with expert-annotated ground truth values, making them a viable alternative to traditional CT-based syndesmosis assessments

Despite these strengths, the study also identified some limitations. The primary constraint is the dataset size and diversity, which affects the generalization capability of the trained networks. Increasing the dataset to include a more diverse range of anatomical variations, including pathological cases, would enhance the robustness and applicability of the model. Additionally, automating the ROI selection process would reduce the dependency on manual intervention, further improving the reproducibility of results.

Another important consideration is the computational cost associated with training and evaluating the deep learning models. While the proposed pipeline integrates various machine learning components effectively, the dataset augmentation process remains computationally intensive. Optimizing the methods to generate the data could help reduce processing time and improve scalability.

In conclusion, this study presents a novel and promising methodology for 3D reconstruction of the tibiofibular joint from 2D biplanar radiographs of ankle. The results validate the effectiveness of the approach and demonstrate its potential for clinical applications, particularly in the assessment of syndesmosis injuries. By addressing the identified limitations and expanding the dataset, future work can further generalize the method, paving the way for its integration into clinical workflows.

PUBLICATIONS

The research presented in this thesis has resulted in several publications in peer-reviewed conferences and journals, along with abstracts submitted to various scientific events. These publications highlight the contributions of this research to the fields of image processing, medical imaging, deep learning, 2D/3D registration and assessment of syndesmosis clinical parameters, disseminating the findings to both academic and clinical communities. Future publications may focus on further optimizations, extended clinical evaluations, and potential applications in other anatomical regions.

1- Dense Point-to-Point Correspondences Between Genus-Zero Shapes Using Cubic Mapping and Horn-Schunck Optical Flow
   Authors: Pejman Hashemibakhtiar, T. Cresson, J. De Guise, and C. Vázquez
   Conference: 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - GRAPP
   Year: 2023 ISBN: 978-989-758-634-7 ISSN: 2184-4321 Publisher: SciTePress Pages: 196-205 DOI: 10.5220/0011674900003417
2- 2D/3D Reconstruction of The Distal Tibiofibular Joint from Biplanar Radiographs Using Deep Learning Registration and Statistical Shape and Intensity Model
   Authors: Pejman Hashemibakhtiar, T. Cresson, M.-L. Nault, J. De Guise, and C. Vázquez
   Conference: 2024 IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece
   Year: 2024 Pages: 1-4 DOI: 10.1109/ISBI56570.2024.10635619
3- Automatically Identifying Anatomical Parameters of the Distal Tibiofibular Relationship from Biplanar X-rays
   Authors: Pejman Hashemibakhtiar, T. Cresson, M.-L. Nault, J. De Guise, and C. Vázquez
   Conference: Canadian Orthopedic Association Annual Meeting, Vancouver, BC
   Year: 2025

Abstract Number: 442

Presentation Type: E-Poster Presentation

4- Automatically Identifying Anatomical Parameters of the Distal Tibiofibular Relationship from Biplanar X-rays

Authors: Pejman Hashemibakhtiar, Thierry Cresson, Marie-Lyne Nault, Jacques de Guise, and Carlos Vázquez

Conference: 43e Journée de la recherche du POES et de la Division d'orthopédie de l'Université de Montréal, CHU Sainte-Justine, Montreal, QC

Year: 2025

5- High-Precision 3D Surface Reconstruction of the Distal Tibiofibular Joint via VGG-Based and image-translation Deep Networks

Authors: Pejman Hashemibakhtiar, Thierry Cresson, Marie-Lyne Nault, Jacques de Guise, and Carlos Vázquez

Submitted to IEEE Transactions on Medical Imaging

6- One-shot Automatic Assessment of the Syndesmosis from Ankle Biplanar radiographs

Authors: Pejman Hashemibakhtiar, Thierry Cresson, Marie-Lyne Nault, Jacques de Guise, and Carlos Vázquez

Submitted to Journal of Orthopaedic Trauma

RECOMMENDATIONS

Based on the limitations identified in this study, the following recommendations are proposed for future research:

- Dataset Expansion: Increasing the number of both pathological and non-pathological CT cases to enhance model generalizability. Ideally, these scans should be acquired in a neutral foot position and under weight-bearing conditions to better reflect real-world clinical scenarios.
- Automated ROI Selection: Developing an automated approach for selecting the region of interest (ROI) in radiographs to minimize manual intervention and improve the reproducibility of the reconstruction pipeline.
- Comprehensive Pipeline Evaluation: Repeating the training and evaluation process multiple times with different datasets and conditions to establish a statistically validated reconstruction accuracy, leading to a more generalized assessment of the pipeline's performance.
- Computational Efficiency: Optimizing the data generation and augmentation methods to reduce processing time and improve scalability, making the methodology more suitable for real-time clinical applications.

ANNEX I

Complete CT scan Segmentation for One Random Case

Following is the slice-by-slice segmentation for one CT-Scan case, the tibia (in red), fibula (in blue) and foot (in yellow) in consecutive slices. First, tibia disappears, and then, gradually fibula disappears up to a point that only the foot is remaining.

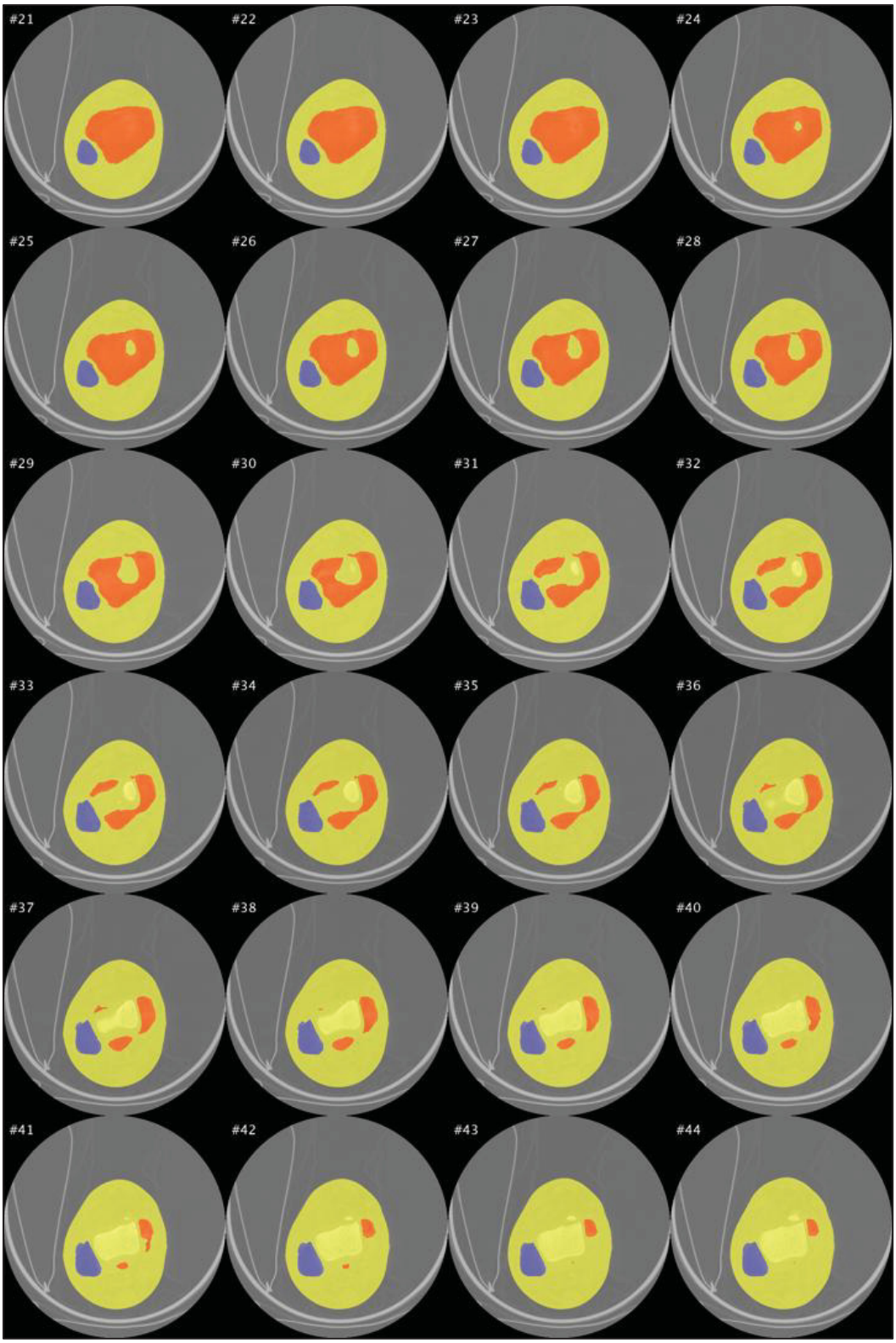

Figure-A I-1: Slice-by-slice segmentation for a CT-Scan sample, slice 1 to 44

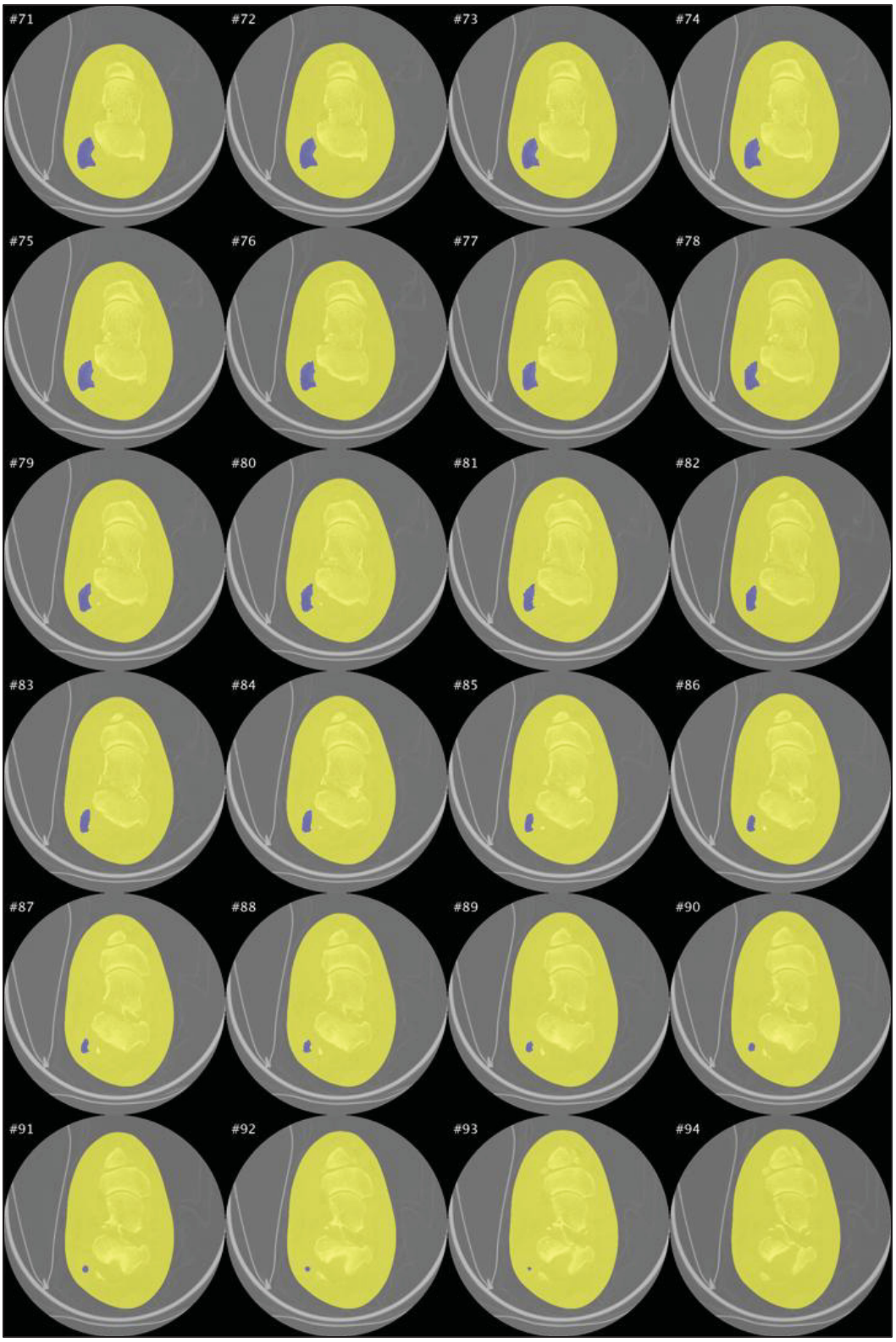

Figure-A I-2: Slice-by-slice segmentation for a CT-Scan sample, slice 71 to 94

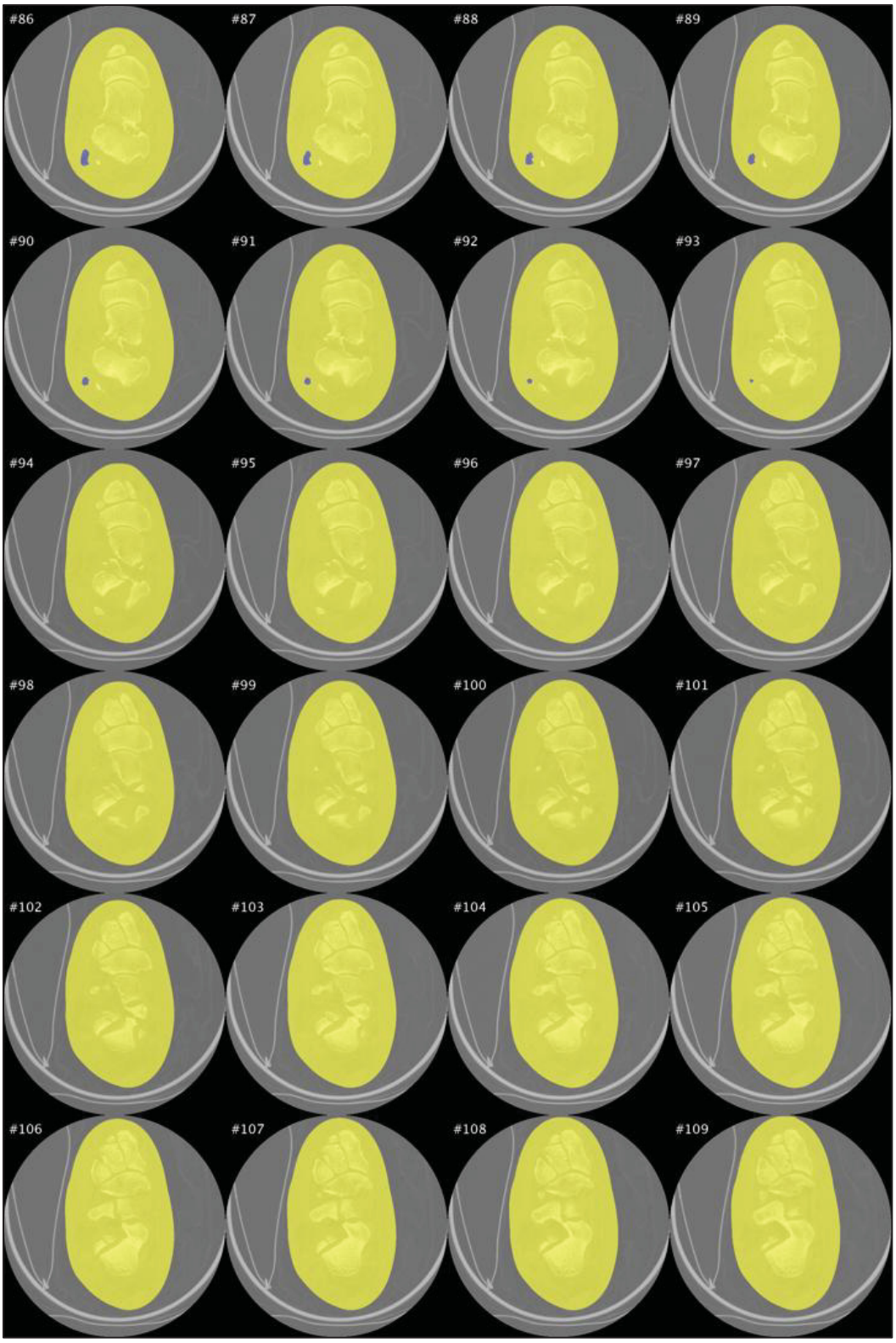

Figure-A I-3: Slice-by-slice segmentation for a CT-Scan sample, slice 86 to 109

ANNEX II

Hyperparameter Optimization Grid Search Heatmaps for JE-Net-AP

In this section you find all the heatmaps for the grid search, applied to optimize the hyperparameters of the JE-Net-AP.

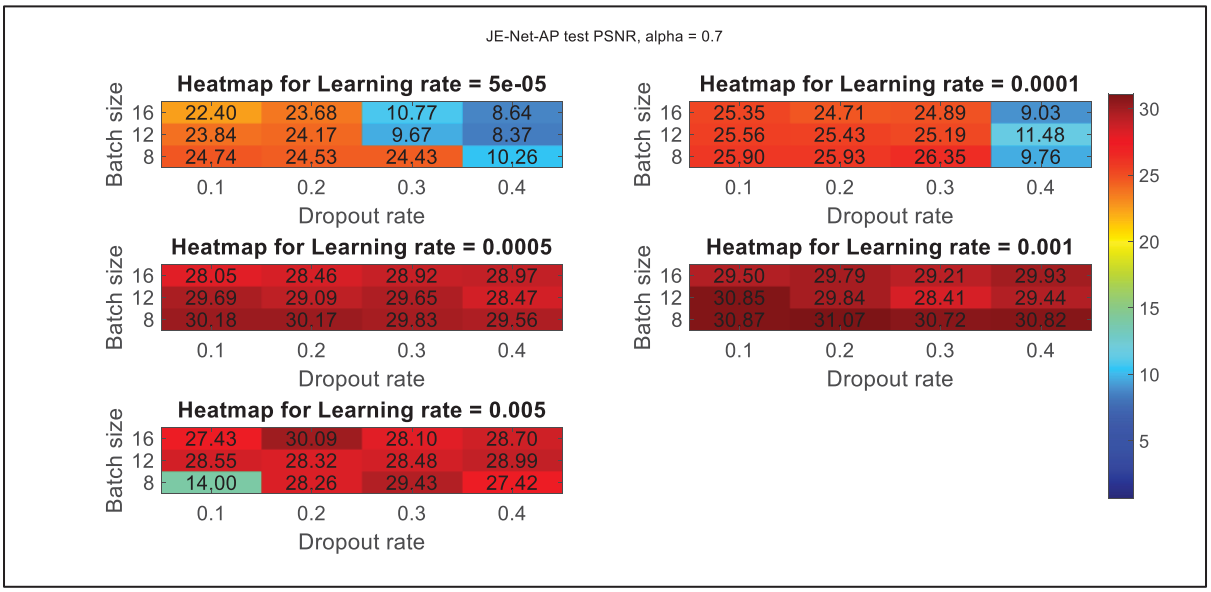

Figure-A II-1: Hyperparameter optimization grid search for JE-Net-AP, Alpha=0.7

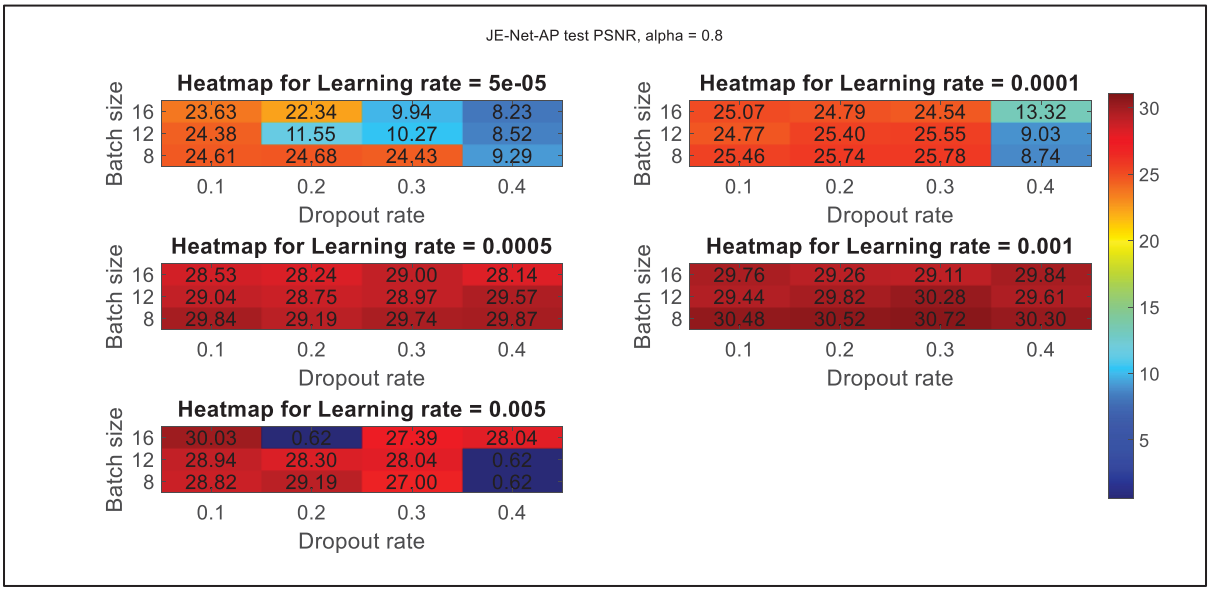

Figure-A II-2: Hyperparameter optimization grid search for JE-Net-AP, Alpha=0.8

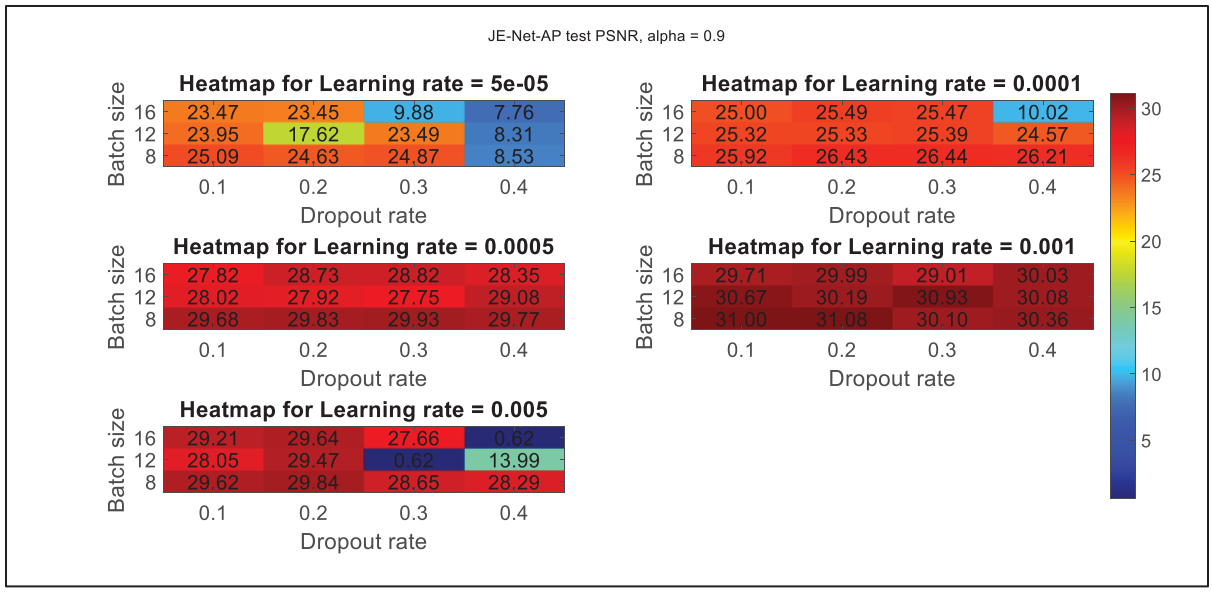

Figure-A II-3: Hyperparameter optimization grid search for JE-Net-AP, Alpha=0.9

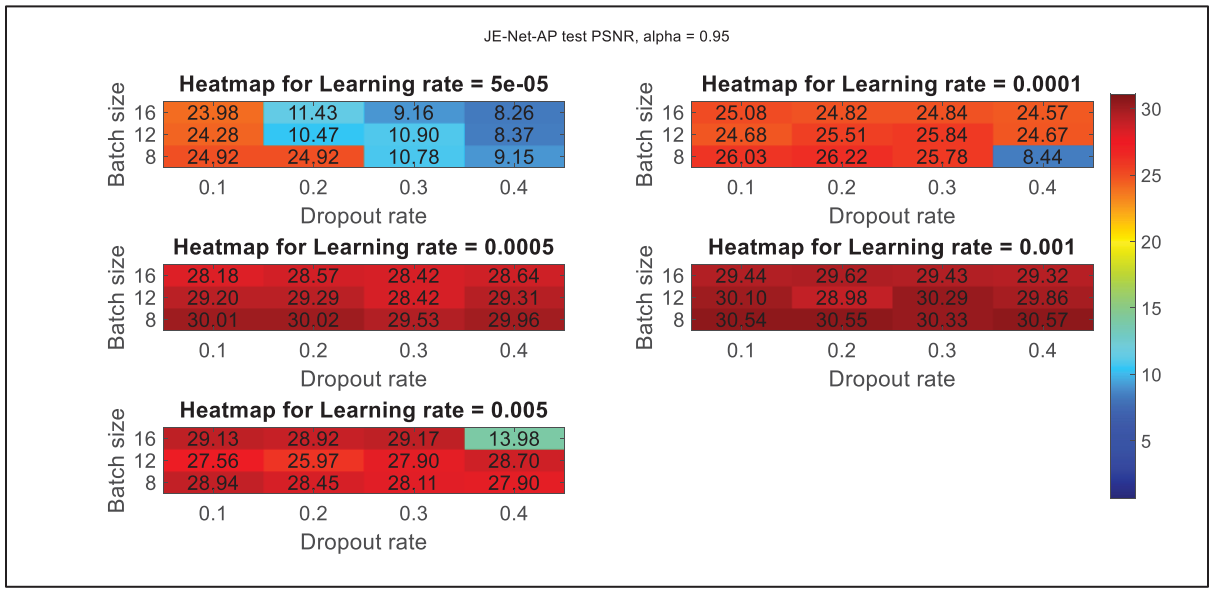

Figure-A II-4: Hyperparameter optimization grid search for JE-Net-AP, Alpha=0.95

ANNEX III

Hyperparameter Optimization Grid Search Heatmaps for JE-Net-LT

In this section you find all the heatmaps for the grid search, applied to optimize the hyperparameters of the JE-Net-LT.

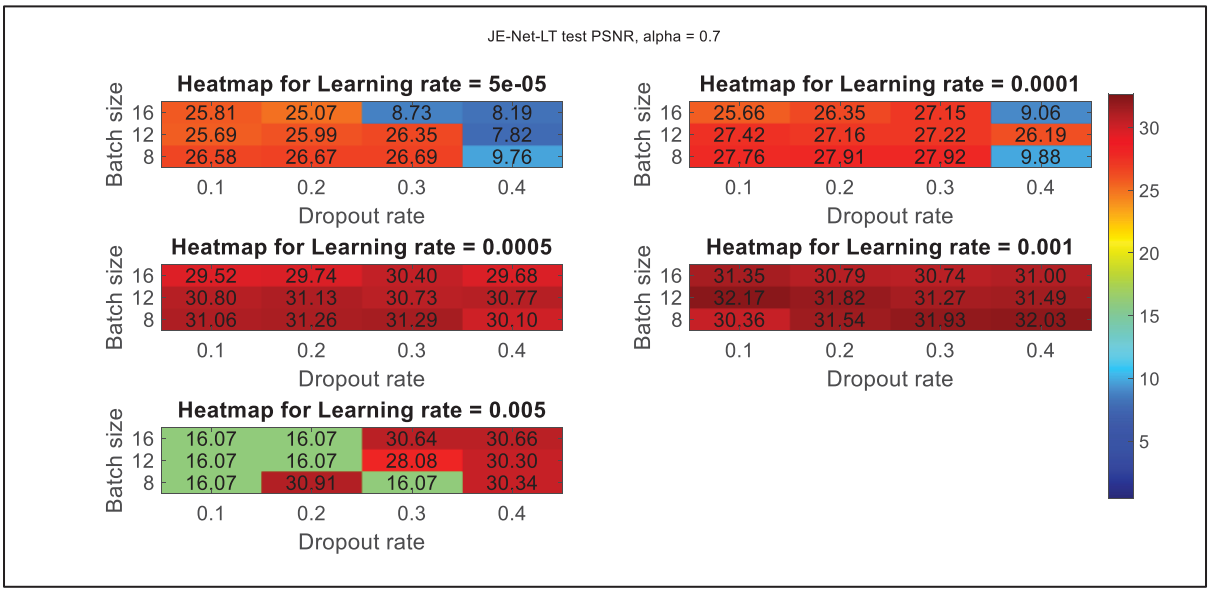

Figure-A III-1: Hyperparameter optimization grid search for JE-Net-LT, Alpha=0.7

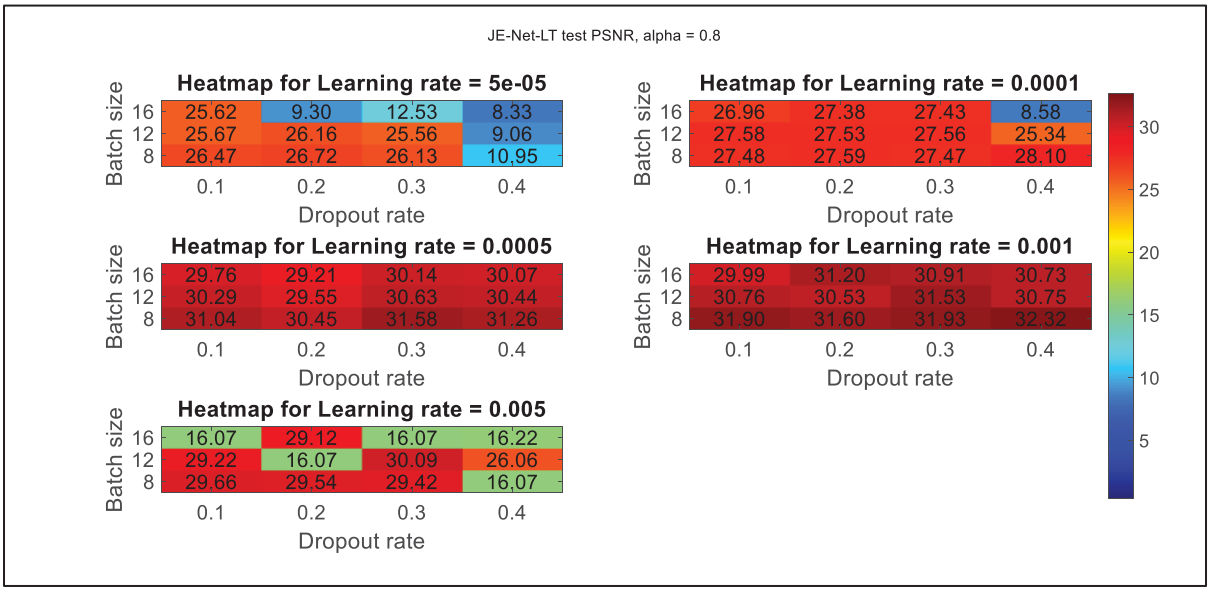

Figure-A III-2: Hyperparameter optimization grid search for JE-Net-LT, Alpha=0.8

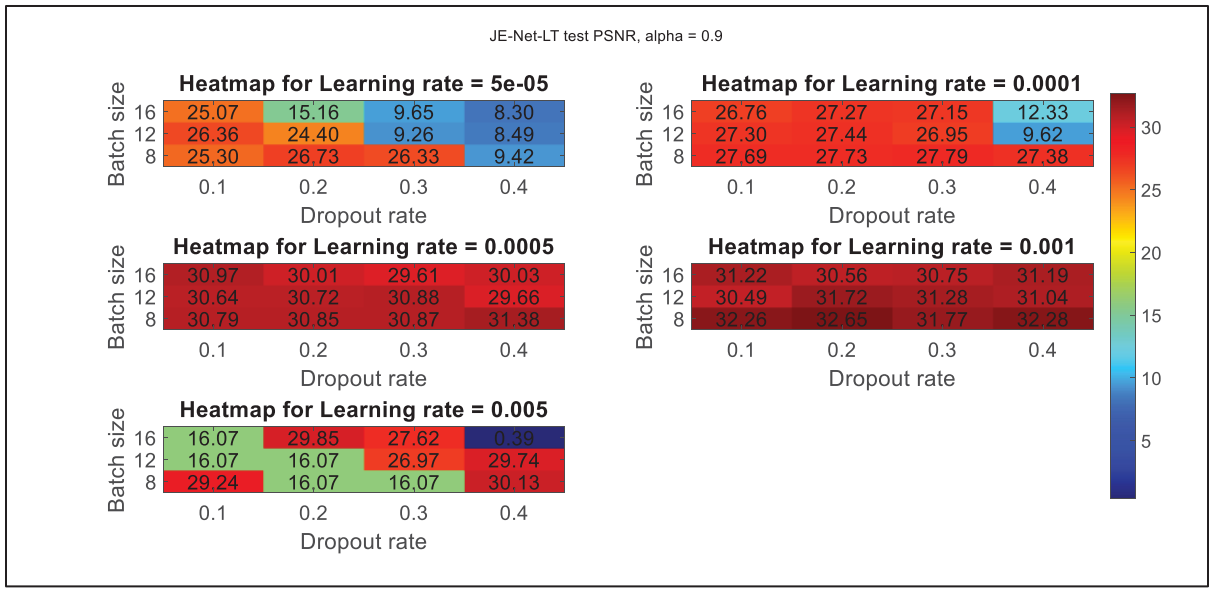

Figure-A III-3: Hyperparameter optimization grid search for JE-Net-LT, Alpha=0.9

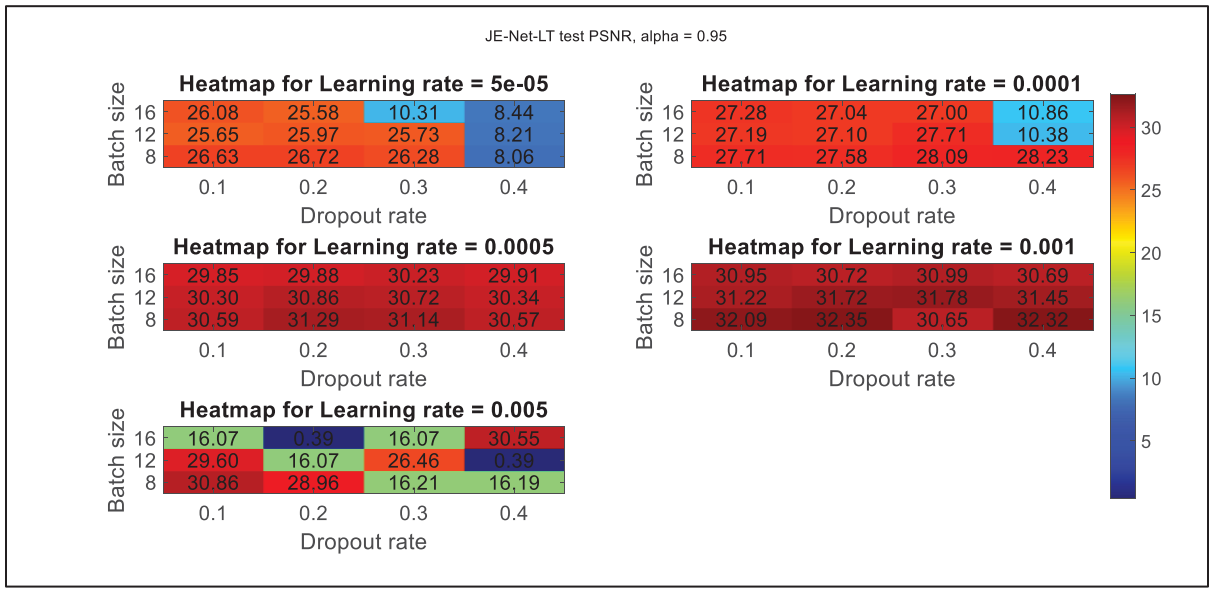

Figure-A III-4: Hyperparameter optimization grid search for JE-Net-LT, Alpha=0.95

LIST OF BIBLIOGRAPHICAL REFERENCES

Agrawal, P., Whitaker, R. T., & Elhabian, S. Y. (2017). Learning deep features for automated placement of correspondence points on ensembles of complex shapes. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10433 LNCS, 185–193. https://doi.org/10.1007/978-3-319-66182-7_22

Akkoul, S., Hafiane, A., Lespessailles, E., & Jennane, R. (2013). 3D femur reconstruction using X-ray stereo pairs. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8157 LNCS (PART 2), 91–100. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41184-7_10

Aldieri, A., Terzini, M., Audenino, A. L., Bignardi, C., & Morbiducci, U. (2020). Combining shape and intensity dxa-based statistical approaches for osteoporotic HIP fracture risk assessment. Computers in Biology and Medicine, 127. https://doi.org/10.1016/j.compbiomed.2020.104093

Ambellan, F., Lamecker, H., von Tycowicz, C., & Zachow, S. (2019). Statistical Shape Models: Understanding and Mastering Variation in Anatomy. Advances in Experimental Medicine and Biology, 1156, 67–84. https://doi.org/10.1007/978-3-030-19385-0_5,

Aouadi, S., & Sarry, L. (2008). Accurate and precise 2D-3D registration based on X-ray intensity. Computer Vision and Image Understanding, 110(1), 134–151. https://doi.org/10.1016/j.cviu.2007.05.006

Arnold, A. S., Blemker, S. S., & Delp, S. L. (2001). Evaluation of a deformable musculoskeletal model for estimating muscle-tendon lengths during crouch gait. Annals of Biomedical Engineering, 29(3), 263–274. https://doi.org/10.1114/1.1355277/METRICS

Aubert, B., Cresson, T., Guise, J. A. De, & Vazquez, C. (2022). X-ray to DRR images translation for efficient multiple objects similarity measures in deformable model 3D / 2D registration. xx(X), 1–14. https://doi.org/10.1109/TMI.2022.3218568

Aubert, B., Vazquez, C., Cresson, T., Parent, S., & De Guise, J. A. (2019). Toward Automated 3D Spine Reconstruction from Biplanar Radiographs Using CNN for Statistical Spine Model Fitting. IEEE Transactions on Medical Imaging, 38(12), 2796–2806. https://doi.org/10.1109/TMI.2019.2914400

Aubert, B., Vidal, P. A., Parent, S., Cresson, T., Vazquez, C., & De Guise, J. (2017). Convolutional neural network and in-painting techniques for the automatic assessment of scoliotic spine surgery from biplanar radiographs. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10434 LNCS, 691–699. Springer Verlag. https://doi.org/10.1007/978-3-319-66185-8_78

Aubin, C. É., Dansereau, J., Parent, F., Labelle, H., & De Guise, J. A. (1997). Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine. Medical & Biological Engineering & Computing, 35(6), 611–618. https://doi.org/10.1007/BF02510968

Audenaert, E. A., Pattyn, C., Steenackers, G., De Roeck, J., Vandermeulen, D., & Claes, P. (2019). Statistical Shape Modeling of Skeletal Anatomy for Sex Discrimination: Their Training Size, Sexual Dimorphism, and Asymmetry. Frontiers in Bioengineering and Biotechnology, 7, 494182. https://doi.org/10.3389/FBIOE.2019.00302/BIBTEX

Bah, M. T., Shi, J., Browne, M., Suchier, Y., Lefebvre, F., Young, P., … Heller, M. O. (2015). Exploring inter-subject anatomic variability using a population of patient-specific femurs and a statistical shape and intensity model. Medical Engineering and Physics, 37(10), 995–1007. https://doi.org/10.1016/j.medengphy.2015.08.004

Baka, N., Kaptein, B. L., de Bruijne, M., van Walsum, T., Giphart, J. E., Niessen, W. J., & Lelieveldt, B. P. F. (2011). 2D-3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. Medical Image Analysis, 15(6), 840–850. https://doi.org/10.1016/j.media.2011.04.001

Baka, N., Niessen, W. J., Kaptein, B. L., van Walsum, T., Ferrarini, L., Reiber, J. H. C., & Lelieveldt, B. P. F. (2010). Correspondence free 3D statistical shape model fitting to sparse X-ray projections. Medical Imaging 2010: Image Processing, 7623(March 2010), 76230D. https://doi.org/10.1117/12.840935

Bakic, P. R., Richard, F. J. P., & Maidment, A. D. A. (2006). 2D-to-3D X-ray breast image registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4057 LNCS, 84–91. Springer Verlag. https://doi.org/10.1007/11784012_11

Balestra, S., Schumann, S., Heverhagen, J., Nolte, L., & Zheng, G. (2014). Articulated statistical shape model-based 2D-3D reconstruction of a hip joint. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8498 LNCS, 128–137. https://doi.org/10.1007/978-3-319-07521-1_14

Baudoin, A., Skalli, W., de Guise, J. A., & Mitton, D. (2008). Parametric subject-specific model for in vivo 3D reconstruction using bi-planar X-rays: Application to the upper femoral extremity. Medical and Biological Engineering and Computing, 46(8), 799–805. https://doi.org/10.1007/s11517-008-0353-8

Bayat, A., Sekuboyina, A., Paetzold, J. C., Payer, C., Stern, D., Urschler, M., … Menze, B. H. (2020). Inferring the 3D Standing Spine Posture from 2D Radiographs. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12266 LNCS, 775–784. https://doi.org/10.1007/978-3-030-59725-2_75

Besl, P. J., & McKay, N. D. (1992). A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2), 239–256. https://doi.org/10.1109/34.121791

Bindernagel, M., Kainmueller, D., Seim, H., Lamecker, H., Zachow, S., & Hege, H. C. (2011). An articulated statistical shape model of the human knee. Informatik Aktuell, 59–63. https://doi.org/10.1007/978-3-642-19335-4_14

Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., & Ayache, N. (2008). Articulated spine models for 3-D reconstruction from partial radiographic data. IEEE Transactions on Biomedical Engineering, 55(11), 2565–2574. https://doi.org/10.1109/TBME.2008.2001125

Brechbuhler, C. H., Gerig, G., & Kubler, O. (1995). Parametrization of closed surfaces for 3-D shape description. Computer Vision and Image Understanding, Vol. 61, pp. 154–170. https://doi.org/10.1006/cviu.1995.1013

Brehler, M., Islam, A., Vogelsang, L. O., Yang, D., Sehnert, W. J., Shakoor, D., … Zbijewski, W. (2019). Coupled active shape models for automated segmentation and landmark localization in high-resolution CT of the foot and ankle. In B. Gimi & A. Krol (Eds.), Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging (p. 23). SPIE. https://doi.org/10.1117/12.2515022

Bushberg, J. T., Seibert, J. Anthony., Leidholdt, E. Marion., & Boone, J. M. (2011). The essential physics of medical imaging. 1048. Retrieved from https://books.google.com/books/about/The_Essential_Physics_of_Medical_Imaging.html?id=RKcTgTqeniwC

Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Xue, Z., … McDonald, C. J. (2014). Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE Transactions on Medical Imaging, 33(2), 577–590. https://doi.org/10.1109/TMI.2013.2290491

Caselles, V., Kimmel, R., & Sapiro, G. (1997). Geodesic Active Contours. International Journal of Computer Vision, 22(1), 61–79. https://doi.org/10.1023/A:1007979827043/METRICS

Cerrolaza, J. J., Picazo, M. L., Humbert, L., Sato, Y., Rueckert, D., Ballester, M. Á. G., & Linguraru, M. G. (2018). Computational Anatomy for Multi-Organ Analysis in Medical Imaging: A Review. Medical Image Analysis, 56, 44–67. https://doi.org/10.1016/j.media.2019.04.002

Cerrolaza, J. J., Villanueva, A., & Cabeza, R. (2012). Hierarchical statistical shape models of multiobject anatomical structures: Application to brain MRI. IEEE Transactions on Medical Imaging, 31(3), 713–724. https://doi.org/10.1109/TMI.2011.2175940

Chaibi, Y., Cresson, T., Aubert, B., Hausselle, J., Neyret, P., Hauger, O., … Skalli, W. (2012). Fast 3D reconstruction of the lower limb using a parametric model and statistical inferences and clinical measurements calculation from biplanar X-rays. Computer Methods in Biomechanics and Biomedical Engineering, 15(5), 457–466. https://doi.org/10.1080/10255842.2010.540758

Chen, J. H., & Shapiro, L. G. (2009). 3D point correspondence by minimum description length with 2DPCA. Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009, 5657–5660. https://doi.org/10.1109/IEMBS.2009.5333769

Chen, J. H., Zheng, K. C., & Shapiro, L. G. (2010). 3D point correspondence by minimum description length in feature space. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6313 LNCS (PART 3), 621–634. https://doi.org/10.1007/978-3-642-15558-1_45

Chen, J., Liu, Y., Wei, S., Bian, Z., Subramanian, S., Carass, A., … Du, Y. (2025, February 1). A survey on deep learning in medical image registration: new technologies, uncertainty, evaluation metrics, and beyond. Medical Image Analysis, Vol. 100. Elsevier B.V. https://doi.org/10.1016/j.media.2024.103385

Chen, X., Varley, M. R., Shark, L. K., Shentall, G. S., & Kirby, M. C. (2007). Automatic 3D-2D image registration using partial digitally reconstructed radiographs along projected anatomic contours. Proceedings - 4th International Conference Medical Information Visualisation: BioMedical Visualisation, MediViz 2007, 3–8. https://doi.org/10.1109/MEDIVIS.2007.7

Chen, X., Varley, M. R., Shark, L. K., Shentall, G. S., & Kirby, M. C. (2008). A computationally efficient method for automatic registration of orthogonal X-ray images with volumetric CT data. Physics in Medicine and Biology, 53(4), 967–983. https://doi.org/10.1088/0031-9155/53/4/010

Chen, Y., Gao, Y., Fu, X., Chen, Y., Wu, J., Guo, C., & Li, X. (2024). Automatic 3D reconstruction of vertebrae from orthogonal bi-planar radiographs. Scientific Reports, 14(1), 1–14. https://doi.org/10.1038/S41598-024-65795-7;SUBJMETA=1421,166,1770,2025,639,692,700,985;KWRD=BIOMEDICAL+ENGINEERING,RADIOGRAPHY,THREE-DIMENSIONAL+IMAGING

Chênes, C., & Schmid, J. (2021). Revisiting Contour-Driven and Knowledge-Based Deformable Models: Application to 2D-3D Proximal Femur Reconstruction from X-ray Images. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12906 LNCS, 451–460. https://doi.org/10.1007/978-3-030-87231-1_44/TABLES/4

Choi, P. T., Lam, K. C., & Lui, L. M. (2015). FLASH: Fast landmark aligned spherical harmonic parameterization for genus-0 closed brain surfaces. SIAM Journal on Imaging Sciences, 8(1), 67–94. https://doi.org/10.1137/130950008

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active Shape Models-Their Training and Application. Computer Vision and Image Understanding, 61(1), 38–59. https://doi.org/10.1006/CVIU.1995.1004

Cootes, Timothy F. (1977). Statistical Models of Appearance for Computer Vision. New England Journal of Medicine, 297(26), 1437–1439. https://doi.org/10.1056/nejm197712292972605

Cresson, T., Branchaud, D., Chav, R., Godbout, B., & de Guise, J. A. (2010). 3D shape reconstruction of bone from two X-ray images using 2D/3D non-rigid registration based on moving least-squares deformation. Medical Imaging 2010: Image Processing, 7623(March 2010), 76230F. https://doi.org/10.1117/12.844098

Cresson, T., Chav, R., Branchaud, D., Humbert, L., Godbout, B., Aubert, B., … De Guise, J. A. (2009). Coupling 2D/3D registration method and statistical model to perform 3D reconstruction from partial X-rays images data. Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009, 1008–1011. https://doi.org/10.1109/IEMBS.2009.5333869

Davies, R. H., Cootes, T. F., & Taylor, C. J. (2001). A minimum description length approach to statistical shape modelling. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2082(5), 50–63. https://doi.org/10.1109/TMI.2002.1009388

Davies, R. H., Twining, C. J., Cootes, T. F., & Taylor, C. J. (2010). Building 3-D Statistical Shape Models by Direct Optimization. IEEE Transactions on Medical Imaging, 29(4), 961–981. https://doi.org/10.1109/TMI.2009.2035048

Dong, X., Ballester, M. A. G., & Zheng, G. (2007). Automatic extraction of femur contours from calibrated X-ray images using statistical information. Journal of Multimedia, 2(5), 46–54. https://doi.org/10.4304/jmm.2.5.46-54

Dong, X., & Zheng, G. (2009). Automatic extraction of proximal femur contours from calibrated X-ray images using 3D statistical models: An in vitro study. International Journal of Medical Robotics and Computer Assisted Surgery, 5(2), 213–222. https://doi.org/10.1002/rcs.253

Dworzak, J., Lamecker, H., Von Berg, J., Klinder, T., Lorenz, C., Kainmüller, D., … Zachow, S. (2010). 3D reconstruction of the human rib cage from 2D projection images using a statistical shape model. International Journal of Computer Assisted Radiology and Surgery, 5(2), 111–124. https://doi.org/10.1007/S11548-009-0390-2,

Eisenberger, M., Lähner, Z., & Cremers, D. (2019). Smooth Shells: Multi-Scale Shape Registration with Functional Maps. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 12262–12271. https://doi.org/10.1109/CVPR42600.2020.01228

Fang, L., Wang, Z., Chen, Z., Jian, F., & He, H. (2016). Reconstruction of 3D lumvar vertebra from two X-ray images based on 2D/3D registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10182 LNCS, 125–134. Springer Verlag. https://doi.org/10.1007/978-3-319-55050-3_12

Fleute, M., & Lavallée, S. (1999). Nonrigid 3-D/2-D registration of images using statistical models. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1679, 138–147. https://doi.org/10.1007/10704282_15

Frangi, A. F., Rueckert, D., Schnabel, J. A., & Niessen, W. J. (2003). Automatic construction of biventricular statistical shape models. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2674, 18–29.

Frangi, Alejandro F., Rueckert, D., Schnabel, J. A., & Niessen, W. J. (2001). Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2082, 78–91.

Fuentes-Jimenez, D., Casillas-Perez, D., Pizarro, D., Collins, T., & Bartoli, A. (2018). Deep Shape-from-Template: Wide-Baseline, Dense and Fast Registration and Deformable Reconstruction from a Single Image. Retrieved from https://arxiv.org/pdf/1811.07791

Galibarov, P. E., Prendergast, P. J., & Lennon, A. B. (2010). A method to reconstruct patient-specific proximal femur surface models from planar pre-operative radiographs. Medical Engineering and Physics, 32(10), 1180–1188. https://doi.org/10.1016/j.medengphy.2010.08.009

Gauvin, C., Dansereau, J., Petit, Y., Guise, J. D. De, & Labelle, H. (1998). Customized 3D radiographic reconstruction of the human pelvis. Annales de Chirurgie.

Gollmer, S. T., Simon, M., Bischof, A., Barkhausen, J., & Buzug, T. M. (2012). Multi-object active shape model construction for abdomen segmentation: Preliminary results. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 3990–3993. https://doi.org/10.1109/EMBC.2012.6346841

Goparaju, A., Iyer, K., Bône, A., Hu, N., Henninger, H. B., Anderson, A. E., … Elhabian, S. Y. (2021). Benchmarking off-the-shelf statistical shape modeling tools in clinical applications. Medical Image Analysis, 76, 102271. https://doi.org/10.1016/J.MEDIA.2021.102271

Goswami, B., & Kr., S. (2015). 3D Modeling of X-Ray Images: A Review. International Journal of Computer Applications, 132(7), 40–46. https://doi.org/10.5120/ijca2015907566

Gower, J. C. (1975). Generalized procrustes analysis. Psychometrika, 40(1), 33–51. https://doi.org/10.1007/BF02291478/METRICS

Groueix, T., Fisher, M., Kim, V. G., Russell, B. C., & Aubry, M. (2018). 3D-CODED: 3D Correspondences by Deep Deformation. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11206 LNCS(c), 235–251. https://doi.org/10.1007/978-3-030-01216-8_15

Gu, W., Gao, C., Grupp, R., Fotouhi, J., & Unberath, M. (2020). Extended Capture Range of Rigid 2D/3D Registration by Estimating Riemannian Pose Gradients. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12436 LNCS, 281–291. https://doi.org/10.1007/978-3-030-59861-7_29/TABLES/2

Guan, S., Meng, C., Xie, Y., Wang, Q., Sun, K., & Wang, T. (2019). Deformable Cardiovascular Image Registration via Multi-Channel Convolutional Neural Network. IEEE Access, 7, 17524–17534. https://doi.org/10.1109/ACCESS.2019.2894943

Guoyan, Z. (2007). Unifying energy minimization and mutual information maximization for robust 2D/3D registration of X-ray and CT images. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4713 LNCS, 547–557. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74936-3_55

Hanaoka, S., Fritscher, K., Schuler, B., Masutani, Y., Hayashi, N., Ohtomo, K., & Schubert, R. (2011). Whole vertebral bone segmentation method with a statistical intensity-shape model-based approach. Medical Imaging 2011: Image Processing, 7962(March 2011), 796242. https://doi.org/10.1117/12.878151

Harmouche, R., Cheriet, F., Labelle, H., & Dansereau, J. (2010). Articulated model registration of MRI/X-ray spine data. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6112 LNCS (PART 2), 20–29. https://doi.org/10.1007/978-3-642-13775-4_3

Hashemibakhtiar, P., Cresson, T., Guise, J. De, & Vázquez, C. (2023). Dense Point-to-Point Correspondences Between Genus-Zero Shapes Using Cubic Mapping and Horn-Schunck Optical Flow. Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 1, 196–205. https://doi.org/10.5220/0011674900003417

Hashemibakhtiar, P., Cresson, T., Nault, M. L., De Guise, J., & Vazquez, C. (2024). 2D/3D Reconstruction of The Distal Tibiofibular Joint from Biplanar Radiographs Using Deep Learning Registration and Statistical Shape and Intensity Model. Proceedings - International Symposium on Biomedical Imaging. https://doi.org/10.1109/ISBI56570.2024.10635619

Heimann, T., & Meinzer, H. P. (2009). Statistical shape models for 3D medical image segmentation: A review. Medical Image Analysis, 13(4), 543–563. https://doi.org/10.1016/j.media.2009.05.004

Heimann, T., Wolf, I., Williams, T., & Meinzer, H. P. (2005). 3D Active Shape Models using gradient descent optimization of description length. Lecture Notes in Computer Science, 3565, 566–577.

Heitz, G., Rohlfing, T., & Maurer, Jr., C. R. (2005). Statistical shape model generation using nonrigid deformation of a template mesh. Medical Imaging 2005: Image Processing, 5747, 1411. SPIE. https://doi.org/10.1117/12.594802

Hosseinian, S., & Arefi, H. (2015). 3D reconstruction from multi-view medical X-ray images - Review and evaluation of existing methods. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 40(1W5), 319–326. https://doi.org/10.5194/isprsarchives-XL-1-W5-319-2015

Houtte, J. Van, Audenaert, E., Zheng, G., & Sijbers, J. (2022). Deep learning-based 2D / 3D registration of an atlas to biplanar X-ray images. International Journal of Computer Assisted Radiology and Surgery. https://doi.org/10.1007/s11548-022-02586-3

Huang, D. X., Zhou, X. H., Xie, X. L., Liu, S. Q., Feng, Z. Q., Hao, J. L., … Yan, L. (2022). A Novel Two-Stage Framework for 2D/3D Registration in Neurological Interventions. 2022 IEEE International Conference on Robotics and Biomimetics, ROBIO 2022, 266–271. https://doi.org/10.1109/ROBIO55434.2022.10011812

Humbert, L., De Guise, J. A., Aubert, B., Godbout, B., & Skalli, W. (2009). 3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences. Medical Engineering and Physics, 31(6), 681–687. https://doi.org/10.1016/j.medengphy.2009.01.003

Hunt, K. J. (2013). Syndesmosis injuries. Current Reviews in Musculoskeletal Medicine, 6(4), 304. https://doi.org/10.1007/S12178-013-9184-9

Hwang, S., Lee, J. J., & Shin, J. (2024). 3D Knee Structure Reconstruction from 2D X-rays Based on Generative Deep Learning Models. 2024 International Technical Conference on Circuits/Systems, Computers, and Communications, ITC-CSCC 2024. https://doi.org/10.1109/ITC-CSCC62988.2024.10628333

Illés, T., & Somoskeöy, S. (2012). The EOSTM imaging system and its uses in daily orthopaedic practice. International Orthopaedics, 36(7), 1325. https://doi.org/10.1007/S00264-012-1512-Y

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065). https://doi.org/10.1098/RSTA.2015.0202

Julien, T., Fotsin, T., Présentée, T., & Technologie, À. L. É. D. E. (2021). Reconstruction 3D du genou ostéophytique à partir de radiographies bi-planes par.

Kadoury, S., Cheriet, F., Kadoury, S., Cheriet, F., & Labelle, H. (2009). Personalized X-Ray 3-D Reconstruction of the Scoliotic Spine from Hybrid Statistical and Image-Based Models. IEEE Transactions on Medical Imaging, 28(9), 1422–1435. https://doi.org/10.1109/TMI.2009.2016756

Kak, A. C., & Slaney, M. (2001). Principles of Computerized Tomographic Imaging. Principles of Computerized Tomographic Imaging. https://doi.org/10.1137/1.9780898719277

Kalisz, S., & Marczyk, M. (2021). Autoencoder-based bone removal algorithm from X-ray images of the lung. BIBE 2021 - 21st IEEE International Conference on BioInformatics and BioEngineering, Proceedings. https://doi.org/10.1109/BIBE52308.2021.9635451

Kasten, Y., Doktofsky, D., & Kovler, I. (2020). End-To-End Convolutional Neural Network for 3D Reconstruction of Knee Bones from Bi-Planar X-Ray Images. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12450 LNCS, 123–133. https://doi.org/10.1007/978-3-030-61598-7_12

Kellett, J. J., Lovell, G. A., Eriksen, D. A., & Sampson, M. J. (2018a). Diagnostic imaging of ankle syndesmosis injuries: A general review. Journal of Medical Imaging and Radiation Oncology, 62(2), 159–168. https://doi.org/10.1111/1754-9485.12708

Kellett, J. J., Lovell, G. A., Eriksen, D. A., & Sampson, M. J. (2018b). Diagnostic imaging of ankle syndesmosis injuries: A general review. Journal of Medical Imaging and Radiation Oncology, 62(2), 159–168. https://doi.org/10.1111/1754-9485.12708

Kirschner, M., & Wesarg, S. (2010). 3D statistical shape model building using consistent parameterization. CEUR Workshop Proceedings, 574, 291–295.

Knaan, D., & Joskowicz, L. (2003). Effective intensity-based 2D/3D rigid registration between fluoroscopic X-ray and CT. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2878, 351–358. https://doi.org/10.1007/978-3-540-39899-8_44

Koehler, C., Wischgoll, T., & Golshani, F. (2019). 3-D Reconstruction of the Human Ribcage Based on Chest X-Ray Images and Geometric Template Models. IEEE Multimedia. https://doi.org/10.1109/MMUL.2009.57

Krähenbühl, N., Weinberg, M. W., Davidson, N. P., Mills, M. K., Hintermann, B., Saltzman, C. L., & Barg, A. (2018). Imaging in syndesmotic injury: a systematic literature review. Skeletal Radiology, 47(5), 631–648. https://doi.org/10.1007/S00256-017-2823-2/FIGURES/6

Kyung, D., Jo, K., Choo, J., Lee, J., & Choi, E. (2023). Perspective Projection-Based 3D CT Reconstruction from Biplanar X-rays. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. https://doi.org/10.1109/ICASSP49357.2023.10096296

Lamecker, H., Wenckebach, T. H., & Hege, H. C. (2006). Atlas-based 3D-shape reconstruction from X-ray images. Proceedings - International Conference on Pattern Recognition, 1, 371–374. https://doi.org/10.1109/ICPR.2006.279

Lamecker, H., & Zachow, S. (2016). Statistical shape modeling of musculoskeletal structures and its applications. Lecture Notes in Computational Vision and Biomechanics, 23, 1–23. https://doi.org/10.1007/978-3-319-23482-3_1/FIGURES/22

Lamer, S., Dube, V., Hebert-Davies, J., Stephane, L., Menard, J., & Nault, M.-L. (2018). Effect of CAM boot immobilization on weightbearing stability in syndesmotic injuries. Https://Doi.Org/10.1177/2473011418S00302, 3(3), 2473011418S0030. https://doi.org/10.1177/2473011418S00302

LaMothe, J. M., Baxter, J. R., Karnovsky, S. C., Murphy, C. I., Gilbert, S., & Drakos, M. C. (2018). Syndesmotic Injury Assessment with Lateral Imaging During Stress Testing in a Cadaveric Model. Foot and Ankle International, 39(4), 479–484. https://doi.org/10.1177/1071100717745660/SUPPL_FILE/FAI745660_-_ICMJE.PDF

Laporte, S., Skalli, W., de Guise, J. A., Lavaste, F., & Mitton, D. (2003). A biplanar reconstruction method based on 2D and 3D contours: Application to the distal femur. Computer Methods in Biomechanics and Biomedical Engineering, 6(1), 1–6. https://doi.org/10.1080/1025584031000065956

Lee, S. C., & Kazhdan, M. (2019). Dense Point-to-Point Correspondences Between Genus-Zero Shapes. 38(5).

Lin, C. F., Gross, M. T., & Weinhold, P. (2006). Ankle syndesmosis injuries: anatomy, biomechanics, mechanism of injury, and clinical guidelines for diagnosis and intervention. The Journal of Orthopaedic and Sports Physical Therapy, 36(6), 372–384. https://doi.org/10.2519/JOSPT.2006.2195

Livyatan, H., Yaniv, Z., & Joskowicz, L. (2003). Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Transactions on Medical Imaging, 22(11), 1395–1406. https://doi.org/10.1109/TMI.2003.819288

Luthi, M., Gerig, T., Jud, C., & Vetter, T. (2018). Gaussian Process Morphable Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(8), 1860–1873. https://doi.org/10.1109/TPAMI.2017.2739743

Mahfouz, M., Badawi, A., Fatah, E. E. A., Kuhn, M., & Merkl, B. (2006). Reconstruction of 3D patient-specific bone models from biplanar X-ray images utilizing morphometric measurements. Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV'06, 2(January), 345–349.

Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. (2012). A review of 3D/2D registration methods for image-guided interventions. Medical Image Analysis, 16(3), 642–661. https://doi.org/10.1016/j.media.2010.03.005

Markelj, Primož, Tomaževič, D., Pernuš, F., & Likar, B. (2008). Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images. IEEE Transactions on Medical Imaging, 27(12), 1704–1714. https://doi.org/10.1109/TMI.2008.923984

Melhem, E., Assi, A., El Rachkidi, R., & Ghanem, I. (2016). EOS® biplanar X-ray imaging: concept, developments, benefits, and limitations. Journal of Children's Orthopaedics, 10(1), 1. https://doi.org/10.1007/S11832-016-0713-0

Mertzanidou, T., Hipwell, J. H., Tanner, C., & Hawkes, D. J. (2010). An intensity-based approach to X-ray mammography: MRI registration. Medical Imaging 2010: Image Processing, 7623(March 2010), 76232Z. https://doi.org/10.1117/12.843988

Mertzanidou, T., Hipwell, J., Johnsen, S., Han, L., Eiben, B., Taylor, Z., … Hawkes, D. (2014). MRI to X-ray mammography intensity-based registration with simultaneous optimisation of pose and biomechanical transformation parameters. Medical Image Analysis, 18(4), 674–683. https://doi.org/10.1016/J.MEDIA.2014.03.003

Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., & Geiger, A. (2018). Occupancy Networks: Learning 3D Reconstruction in Function Space. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June, 4455–4465. https://doi.org/10.1109/CVPR.2019.00459

Miao, S., Wang, Z. J., & Liao, R. (2016). A CNN Regression Approach for Real-Time 2D/3D Registration. IEEE Transactions on Medical Imaging, 35(5), 1352–1363. https://doi.org/10.1109/TMI.2016.2521800

Michael, G. (2001). X-ray computed tomography. Physics Education, 36(6), 442. https://doi.org/10.1088/0031-9120/36/6/301

Mitton, D., DeschêNes, S., Laporte, S., Godbout, B., Bertrand, S., de Guise, J. A., & Skalli, W. (2006). 3D reconstruction of the pelvis from bi-planar radiography. Computer Methods in Biomechanics and Biomedical Engineering, 9(1), 1–5. https://doi.org/10.1080/10255840500521786

Mitton, D., Landry, C., Véron, S., Skalli, W., Lavaste, F., & De Guise, J. A. (2000). 3D reconstruction method from biplanar radiography using non-stereo corresponding points and elastic deformable meshes. Medical & Biological Engineering & Computing, 38(2), 133–139. https://doi.org/10.1007/BF02344767

Mitulescu, A., Semaan, I., De Guise, J. A., Leborgne, P., & Adamsbaum, C. (2001). Validation of the non-stereo corresponding points stereo radiographic 3D reconstruction technique. Medical & Biological Engineering & Computing, 39(2), 152–158. https://doi.org/10.1007/BF02344797

Mitulescu, A., Skalli, W., Mitton, D., & De Guise, J. (2002). Three-dimensional surface rendering reconstruction of scoliotic vertebrae using a non-stereo-corresponding points technique. European Spine Journal, 11(4), 344–352. https://doi.org/10.1007/s00586-002-0432-8

Montillo, A., Song, Q., Liu, X., & Miller, J. V. (2013). Parsing radiographs by integrating landmark set detection and multi-object active appearance models. Medical Imaging 2013: Image Processing, 8669(March 2013), 86690H. https://doi.org/10.1117/12.2007138

Moura, D. C., Boisvert, J., Barbosa, J. G., Labelle, H., & Tavares, J. M. R. S. (2011). Fast 3D reconstruction of the spine from biplanar radiographs using a deformable articulated model. Medical Engineering and Physics, 33(8), 924–933. https://doi.org/10.1016/j.medengphy.2011.03.007

Moura, D. C., Boisvert, J., Barbosa, J. G., & Tavares, J. M. R. S. (2009). Fast 3D reconstruction of the spine using user-defined splines and a statistical articulated model. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5875 LNCS (PART 1), 586–595. https://doi.org/10.1007/978-3-642-10331-5_55

Nault, M. L., Hébert-Davies, J., Laflamme, G. Y., & Leduc, S. (2013). CT scan assessment of the syndesmosis: A new reproducible method. Journal of Orthopaedic Trauma, 27(11), 638–641. https://doi.org/10.1097/BOT.0b013e318284785a

Nguyen, D. C. T., Benameur, S., Mignotte, M., & Lavoie, F. (2023). 3D biplanar reconstruction of lower limbs using nonlinear statistical models. Medical and Biological Engineering and Computing, 61(11), 2877–2894. https://doi.org/10.1007/S11517-023-02882-3,

Nolte, D., Xie, S., & Bull, A. M. J. (2023). 3D shape reconstruction of the femur from planar X-ray images using statistical shape and appearance models. Bio-Medical Engineering Online, 22(1). https://doi.org/10.1186/S12938-023-01093-Z

Oh, J., & Koo, S. (2024). Fast digitally reconstructed radiograph generation using particle-based statistical shape and intensity model. Https://Doi.Org/10.1117/1.JMI.11.3.033503, 11(3), 033503. https://doi.org/10.1117/1.JMI.11.3.033503

Patil, A., Kulkarni, K., Xie, S., Bull, A. M. J., & Jones, G. G. (2023). The accuracy of statistical shape models in predicting bone shape: A systematic review. The International Journal of Medical Robotics + Computer Assisted Surgery: MRCAS, 19(3). https://doi.org/10.1002/RCS.2503

Patil, S., & Ravi, B. (2005). Voxel-based representation, display and thickness analysis of intricate shapes. Ninth International Conference on Computer Aided Design and Computer Graphics (CAD-CG'05), 2005, 415–420. https://doi.org/10.1109/CAD-CG.2005.86

Pearcy, M. J. (1985). Stereo radiography of lumbar spine motion. Acta Orthopaedica Scandinavica. Supplementum, 212(S212), 1–45. https://doi.org/10.3109/17453678509154154

Pomero, V., Mitton, D., Laporte, S., De Guise, J. A., & Skalli, W. (2004a). Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clinical Biomechanics, 19(3), 240–247. https://doi.org/10.1016/j.clinbiomech.2003.11.014

Pomero, V., Mitton, D., Laporte, S., De Guise, J. A., & Skalli, W. (2004b). Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clinical Biomechanics, 19(3), 240–247. https://doi.org/10.1016/J.CLINBIOMECH.2003.11.014

Qiu, H., Hammernik, K., Qin, C., Chen, C., & Rueckert, D. (2021). Embedding Gradient-based Optimization in Image Registration Networks. Retrieved from https://arxiv.org/pdf/2112.03915

Ramlee, M. H., Kadir, M. R. A., & Harun, H. (2013). Three-dimensional modeling and analysis of a human ankle joint. Proceeding - 2013 IEEE Student Conference on Research and Development, SCOReD 2013, (December), 74–78. https://doi.org/10.1109/SCOReD.2013.7002545

Reyneke, C. J. F., Luthi, M., Burdin, V., Douglas, T. S., Vetter, T., & Mutsvangwa, T. E. M. (2019). Review of 2-D/3-D Reconstruction Using Statistical Shape and Intensity Models and X-Ray Image Synthesis: Toward a Unified Framework. IEEE Reviews in Biomedical Engineering, 12, 269–286. https://doi.org/10.1109/RBME.2018.2876450

Reyneke, C., Thusini, X., Douglas, T., Vetter, T., & Mutsvangwa, T. (2018). Construction and validation of image-based statistical shape and intensity models of bone. 2018 3rd Biennial South African Biomedical Engineering Conference, SAIBMEC 2018, 1–4. https://doi.org/10.1109/SAIBMEC.2018.8363176

Rueckert, D., Frangi, A. F., & Schnabel, J. A. (2001). Automatic construction of 3D statistical deformation models using non-rigid registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2208(8), 77–84. https://doi.org/10.1007/3-540-45468-3_10

Sarkalkan, N., Weinans, H., & Zadpoor, A. A. (2014). Statistical shape and appearance models of bones. Bone, 60, 129–140. https://doi.org/10.1016/j.bone.2013.12.006

Sas, A., Tanck, E., Sermon, A., & van Lenthe, G. H. (2020). Finite element models for fracture prevention in patients with metastatic bone disease. A literature review. Bone Reports, 12. https://doi.org/10.1016/J.BONR.2020.100286

Schmid, J., Assassi, L., & Chênes, C. (2023). A novel image augmentation based on statistical shape and intensity models: application to the segmentation of hip bones from CT images. European Radiology Experimental, 7(1), 1–11. https://doi.org/10.1186/S41747-023-00357-6/FIGURES/7

Seitel, A., Rasoulian, A., Rohling, R., & Abolmaesumi, P. (2015). Lumbar and thoracic spine segmentation using a statistical multi-object shape+pose model. Lecture Notes in Computational Vision and Biomechanics, 20, 221–225. https://doi.org/10.1007/978-3-319-14148-0_19

Shalaby, A., Farag, A. A., Ross, A., & Hockenbury, T. (2012). 2D-3D registration of human ankle using X-ray and CT images. 2012 Cairo International Biomedical Engineering Conference, CIBEC 2012, 23–26. https://doi.org/10.1109/CIBEC.2012.6473293

Shiode, R., Kabashima, M., Hiasa, Y., Oka, K., Murase, T., Sato, Y., & Otake, Y. (2021). 2D–3D reconstruction of distal forearm bone from actual X-ray images of the wrist using convolutional neural networks. Scientific Reports, 11(1), 1–12. https://doi.org/10.1038/S41598-021-94634-2;SUBJMETA=166,308,639,692;KWRD=ENGINEERING,MEDICAL+RESEARCH

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. Retrieved from https://arxiv.org/abs/1409.1556v6

Skadłubowicz, P., Król, Z., Wróbel, Z., Hefti, F., & Krieg, A. (2009). USING OF STATISTICAL SHAPE MODELS FOR PELVIS RECONSTRUCTION IN THE ONCOLOGIC SURGERY. Journal of Medical Informatics and Technologies.

Styner, M. A., Rajamani, K. T., Nolte, L. P., Zsemlye, G., Székely, G., Taylor, C. J., & Davies, R. H. (2003). Evaluation of 3D correspondence methods for model building. Information Processing in Medical Imaging: Proceedings of the ... Conference, 18, 63–75.

Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J. J., … Gerig, G. (2006). Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. The Insight Journal, c (1071), 242–250. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21941375%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3062073

Subburaj, K., Ravi, B., & Agarwal, M. (2009). Automated identification of anatomical landmarks on 3D bone models reconstructed from CT scan images. Computerized Medical Imaging and Graphics, 33(5), 359–368. https://doi.org/10.1016/j.compmedimag.2009.03.001

Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. BMC Medical Imaging, 15(1), 1–28. https://doi.org/10.1186/S12880-015-0068-X/TABLES/5

Tchinde Fotsin, T. J., Vazquez, C., Cresson, T., & De Guise, J. (2019). Shape, Pose and Density Statistical Model for 3D Reconstruction of Articulated Structures from X-Ray Images. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2748–2751. https://doi.org/10.1109/EMBC.2019.8857699

Tomazevic, D., Likar, B., & Pernus, F. (2002). Rigid 2D/3D registration of intraoperative digital X-ray images and preoperative CT and MR images. Medical Imaging 2002: Image Processing, 4684(May 2002), 507. https://doi.org/10.1117/12.467193

Unberath, M., Gao, C., Hu, Y., Judish, M., Taylor, R. H., Armand, M., & Grupp, R. (2021). The Impact of Machine Learning on 2D/3D Registration for Image-Guided Interventions: A Systematic Review and Perspective. Frontiers in Robotics and AI, 8, 716007. https://doi.org/10.3389/FROBT.2021.716007/XML/NLM

van Dijk, C. N., Longo, U. G., Loppini, M., Florio, P., Maltese, L., Ciuffreda, M., & Denaro, V. (2016). Classification and diagnosis of acute isolated syndesmotic injuries: ESSKA-AFAS consensus and guidelines. Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA, 24(4), 1200–1216. https://doi.org/10.1007/S00167-015-3942-8

van Houtte, J., Gao, X., Sijbers, J., & Zheng, G. (2021). 2D/3D registration with a statistical deformation model prior using deep learning. BHI 2021 - 2021 IEEE EMBS International Conference on Biomedical and Health Informatics, Proceedings. https://doi.org/10.1109/BHI50953.2021.9508540

Vikas, K., & Bhallamudi, R. (2014). Application of laplacian surface deformation and self-organizing maps to calculate shape correspondence for statistical shape modeling. 2014 IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014, 369–372. https://doi.org/10.1109/ISBI.2014.6867885

Wang, J., & Shi, C. (2017). Automatic construction of statistical shape models using deformable simplex meshes with vector field convolution energy. BioMedical Engineering Online, 16(1), 1–19. https://doi.org/10.1186/s12938-017-0340-0

Weese, J., Buzug, T. M., Lorenz, C., & Fassnacht, C. (1997). An approach to 2D/3D registration of a vertebra in 2d X-ray fluoroscopies with 3D CT images. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1205, 119–128. https://doi.org/10.1007/bfb0029231

Wei, D., Ahmad, S., Guo, Y., Chen, L., Huang, Y., Ma, L., … Wang, Q. (2022). Recurrent Tissue-Aware Network for Deformable Registration of Infant Brain MR Images. IEEE Transactions on Medical Imaging, 41(5), 1219–1229. https://doi.org/10.1109/TMI.2021.3137280

Wei, W., Wang, G., & Chen, H. (2009). 3D reconstruction of a femur shaft using a model and two 2D X-ray images. Proceedings of 2009 4th International Conference on Computer Science and Education, ICCSE 2009, 720–722. https://doi.org/10.1109/ICCSE.2009.5228123

Wilms, M., Handels, H., & Ehrhardt, J. (2017). Multi-resolution multi-object statistical shape models based on the locality assumption. Medical Image Analysis, 38, 17–29. https://doi.org/10.1016/j.media.2017.02.003

Wu, J., Li, G., Lu, H., Kim, H., & Ogunbona, P. O. (2017). Statistical shape model generation using diffeomorphic surface registration. ACM International Conference Proceeding Series, 37–41. https://doi.org/10.1145/3133793.3133796

Xie, Y., Meng, C., Guan, S., & Wang, Q. (2018). Single shot 2D3D image registration. Proceedings - 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017, 2018-Janua, 1–5. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CISP-BMEI.2017.8302202

Xu, H., & Elhabian, S. Y. (2024). Optimization-Driven Statistical Models of Anatomies using Radial Basis Function Shape Representation. Retrieved from https://arxiv.org/pdf/2411.15882

Zhang, B., Faghihroohi, S., Azampour, M. F., Liu, S., Ghotbi, R., Schunkert, H., & Navab, N. (2023). A Patient-Specific Self-supervised Model for Automatic X-Ray/CT Registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 14228 LNCS, 515–524. https://doi.org/10.1007/978-3-031-43996-4_49/TABLES/3

Zhang, J., Lv, L., Shi, X., Wang, Y., Guo, F., Zhang, Y., & Li, H. (2013). 3-D reconstruction of the spine from biplanar radiographs based on contour matching using the hough transform. IEEE Transactions on Biomedical Engineering, 60(7), 1954–1964. https://doi.org/10.1109/TBME.2013.2246788,

Zheng, G. (2006). RECONSTRUCTION OF PATIENT-SPECIFIC 3D BONE MODEL FROM BIPLANAR X-RAY IMAGES AND POINT DISTRIBUTION MODELS. Compute, (2), 1197–1200.

Zheng, G. (2008). Robust intensity-based 3D-2D registration of CT and X-ray images for precise estimation of cup alignment after total hip arthroplasty. Proceedings - International Conference on Pattern Recognition. https://doi.org/10.1109/icpr.2008.4761766

Zheng, G. (2011). Personalized X-ray reconstruction of the proximal femur via intensity-based non-rigid 2D-3D registration. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6892 LNCS (PART 2), 598–606. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-23629-7_73

Zheng, G., Gollmer, S., Schumann, S., Dong, X., Feilkas, T., & González Ballester, M. A. (2009). A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. Medical Image Analysis, 13(6), 883–899. https://doi.org/10.1016/j.media.2008.12.003

Zheng, G., & Nolte, L. P. (2006). Surface reconstruction of bone from X-ray images and point distribution model incorporating a novel method for 2D-3D correspondence. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2, 2237–2244. https://doi.org/10.1109/CVPR.2006.300

Zheng, Guoyan; Yu, W. (2017). Atlas-Based 3D Intensity Volume Reconstruction of Musculoskeletal Structures in the Lower Extremity from 2D Calibrated X-Ray Images. 10549, 11–18. https://doi.org/10.1007/978-3-319-67552-7

Zöllei, L., Grimson, E., Norbash, A., & Wells, W. (2001). 2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. Proceedings of the IEEE Computer Society Conference on Com