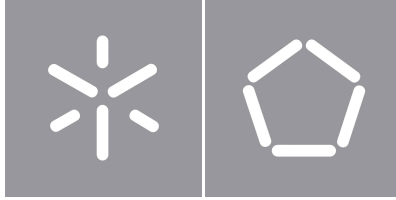


University of Minho
School of Engineering

Francisca Fernandes Silva

Non-invasive bone registration resorting to ultrasound for surgical navigation



University of Minho

Francisca Fernandes Silva

Non-invasive bone registration resorting to ultrasound for surgical navigation

Master's Dissertation in Medical Electronics

Dissertation supervised by

Professor Doctor Cristina P.Santos

Professor Doctor Elena de Momi

Copyright and Terms of Use for Third Party Work

This dissertation reports on academic work that can be used by third parties as long as the internationally accepted standards and good practices are respected concerning copyright and related rights.

This work can thereafter be used under the terms established in the license below.

Readers needing authorization conditions not provided for in the indicated licensing should contact the author through the RepositóriUM of the University of Minho.

License granted to users of this work:



CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Acknowledgements

Esta tese representa o fim de uma etapa marcante na minha vida, durante a qual muito aprendi, vivi e cresci como pessoa. Ao longo destes anos, tive o privilégio de partilhar conhecimento e experiências com um grupo de pessoas, a quem manifesto o meu mais sincero agradecimento.

À Prof. Dr. Cristina P. Santos e à equipa do BirdLab, agradeço a oportunidade de trabalhar neste projeto e o ambiente de partilha. À Diana, deixo um obrigada muito especial. A tua presença constante e o teu apoio prático e emocional em cada etapa foram determinantes.

To Prof. Dr. Elena De Momi, I express my sincere gratitude for letting me join the NearLab team during a significant part of my thesis. This experience was truly enriching and will remain a very special part of this journey.

Gostaria ainda de agradecer ao Dr. Carlos Macedo, pela paciência, dedicação e pela forma generosa como partilhou o seu conhecimento médico, e ao Prof. Dr. Oscar Carvalho, por disponibilizar equipamento essencial à realização desta tese.

Aos meus pais e aos meus avós, o meu mais profundo agradecimento por serem o meu pilar, por sempre se orgulharem de mim e acreditarem que eu seria capaz de alcançar tudo. À minha irmã, por ser um dos meus grandes exemplos, e à minha Lola, por toda a alegria e por me ajudar a aliviar o stress nos momentos mais exigentes. Ao meu namorado, agradeço por me incentivar nas minhas escolhas e por me ajudar no que está ao seu alcance. Esta conquista é tanto minha quanto vossa.

Aos meus amigos de sempre, por continuarem ao meu lado e serem a minha constante, e aos que a universidade me trouxe, em particular ao grupo das Miquelinas, quero agradecer por fazerem desta jornada algo inesquecível. Sem vocês, nada disto teria sido igual. Obrigada pelos ombros amigos, pelos desabafos, pelos momentos de diversão e por todo o amor que me deram ao longo destes anos. Levo-vos comigo para tudo o que vier a seguir.

Statement of Integrity

I hereby declare having conducted this academic work with integrity.

I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, Braga, april 2026

Francisca Fernandes Silva

Abstract

Orthopaedic procedures are vital for alleviating chronic pain and restoring mobility, significantly improving patients' quality of life, but demand extreme surgical precision, as there is a narrow margin for error, and depend heavily on the surgeon's experience. Computer-assisted orthopaedic surgery (CAOS) systems have emerged as an important support tool for surgeons, facilitating planning and intraoperative navigation to achieve better clinical outcomes. Yet, registration, the critical process of aligning virtual models with the patient's anatomy during surgical navigation, remains a major clinical bottleneck. This process is predominantly invasive, requiring bone markers that, while accurate, cause additional surgical trauma and increase the risk of postoperative infections.

To address these clinical limitations, this dissertation proposes a non-invasive, ultrasound based bone registration framework for intraoperative surgical navigation. The computational pipeline seamlessly integrates artificial intelligence with spatial tracking, utilising advanced deep learning architectures, trained on an expertly annotated human US dataset, for the automatic segmentation of bone surfaces. To operationalise the system, custom software modules were developed to manage real-time communication, execute dynamic three-dimensional reconstructions, and perform the registration of the models.

Validation studies of the system have highlighted its intraoperative feasibility in real-world conditions. The iterative closest point algorithm demonstrated superiority over the coherent point drift method, achieving a target registration error of 1.46 mm, within the clinical margin of 1.5 mm, in 87 s. Ultimately, this thesis establishes a highly accurate, non-invasive CAOS system, paving the way for future advances in real-time non-invasive bone tracking and in vivo human validation.

Keywords Computer-assisted orthopaedic surgery, Deep learning, Bone registration, Ultrasound

Resumo

Os procedimentos ortopédicos são vitais para aliviar a dor crónica e restaurar a mobilidade, melhorando significativamente a qualidade de vida dos doentes, mas exigem uma precisão cirúrgica extrema, devido à reduzida margem de erro, e dependem fortemente da experiência do cirurgião. Os sistemas de Cirurgia Ortopédica Assistida por Computador (CAOS) surgiram como importantes ferramentas de apoio aos cirurgiões, facilitando o planeamento e a navegação intraoperatória para alcançar melhores resultados clínicos. No entanto, o registo, o processo crítico de alinhamento de modelos virtuais com a anatomia do paciente durante a navegação cirúrgica, continua a ser um grande obstáculo clínico. Este processo é predominantemente invasivo, exigindo marcadores ósseos que, embora precisos, causam trauma cirúrgico adicional e aumentam o risco de infeções pós-operatórias.

Para dar resposta a estas limitações clínicas, esta dissertação propõe uma metodologia de registo ósseo não invasiva, baseado em ultrassom, para navegação cirúrgica intraoperatória. O fluxo de trabalho computacional integra perfeitamente a inteligência artificial com o rastreamento espacial, utilizando arquiteturas avançadas de aprendizagem profunda treinadas num conjunto de dados de imagens de ultrassom humano anotadas por especialistas para a segmentação automática de superfícies ósseas. Para operacionalizar o sistema, desenvolveram-se módulos de software personalizados para gerir a comunicação em tempo real, realizar reconstruções tridimensionais dinâmicas e efetuar o registo dos modelos.

Estudos de validação do sistema sublinharam a sua viabilidade intraoperatória no mundo real. O algoritmo *iterative closest point* demonstrou superioridade face ao *coherent point drift*, assegurando 1,46 mm de erro de registo do alvo, dentro da margem clínica de 1,5 mm, em 87 s. Em última análise, esta dissertação estabelece um sistema CAOS não invasivo de elevada precisão, abrindo caminho para avanços futuros no rastreio ósseo não invasivo em tempo real e na validação *in vivo*.

Palavras-chave Aprendizagem profunda, Cirurgia ortopédica assistida por computador, Registo do osso, Ultrassom

Contents

1	Introduction	1
1.1	Motivation and Problem Statement	1
1.2	Goals and Research Questions	2
1.3	Contribution to Knowledge	3
1.4	Dissertation Outline	4
2	Background	5
2.1	Anatomy of Knee Joint	5
2.2	Medical Imaging Techniques in Orthopaedics	6
2.2.1	CT	6
2.2.2	MRI	7
2.2.3	Nuclear Imaging	8
2.2.4	US	9
2.3	CAOS systems	10
2.3.1	Surgical Planning	11
2.3.2	Surgical Navigation	11
3	Review of Bone Registration Techniques for Surgical Navigation	14
3.1	Methodology	14
3.2	Results	15
3.2.1	Anatomical Body Area	20
3.2.2	Medical Imaging Modality	20
3.2.3	Hardware	21
3.2.4	Registration Method	22
3.2.5	Tracking	24
3.2.6	Quantitative Outcome	25

3.3	Discussion	26
3.4	Conclusion	28
4	Solution description	30
4.1	System Requirements	30
4.2	Solution Overview	31
4.3	Hardware and Software Infrastructure	32
4.3.1	Hardware	32
4.3.2	Software	34
4.4	Conclusion	34
5	Deep Learning-Based Pipeline for Bone Surface Extraction in Knee Ultrasound	36
5.1	Clinical Dataset Acquisition	36
5.1.1	Dataset Sourcing and Clinical Acquisition Protocol	36
5.1.2	Cohort Demographics and Dataset Characteristics	38
5.2	Ground Truth Generation	40
5.2.1	Development of a Graphical User Interface for Manual Labelling	41
5.2.2	Manual Bone Segmentation and Ground Truth Generation	43
5.3	Deep Learning Framework	44
5.3.1	Data Curation and Subject-Level Partitioning	44
5.3.2	Image Preprocessing and Augmentation	45
5.3.3	Network Architectures	48
5.3.4	Training Configuration and Loss Functions	49
5.3.5	Evaluation Metrics	50
5.4	Results and Discussion	51
5.4.1	Quantitative Analysis	51
5.4.2	Qualitative Analysis	54
5.4.3	Limitations and Final Architecture Selection	55
5.5	Conclusion	56
6	3D Bone Reconstruction and Registration	57
6.1	System Calibration	57
6.1.1	Stylus Calibration and Phantom Registration	58
6.1.2	Temporal Calibration	59

6.1.3	Spatial Calibration	60
6.2	3D Volume Reconstruction	61
6.2.1	Experimental Phantoms and Acquisition	61
6.2.2	Data Communication and Module Architecture	62
6.2.3	Real-Time Volume Reconstruction	63
6.2.4	Point Cloud Extraction and Refinement	64
6.3	3D Surface Registration	66
6.3.1	Cross-Modality Data Harmonisation	66
6.3.2	Coarse Registration	67
6.3.3	Fine Registration	68
6.4	Evaluation Strategy and Metrics	69
6.5	Results and Discussion	71
6.5.1	Calibration	71
6.5.2	3D Reconstruction and Registration	73
6.6	Conclusion	77
7	Conclusion	78
7.1	Future Work	81
	Bibliography	83
	Appendices	96
A	Design of the Probe Reference	96
B	Design of the Calibration Phantom Reference	97
C	Detailed Calibration Measurements	98
C.1	Temporal Calibration	98
C.2	Spatial Calibration	99

List of Figures

2.1	Anatomy of the knee joint.	5
2.2	Ultrasound phased array transducer.	10
2.3	Categorisation of registration approaches.	13
3.1	Flow diagram of search strategy based on PRISMA.	15
3.2	Number of studies per anatomical body area.	20
3.3	Number of studies per preoperative and intraoperative data acquisition modality.	21
3.4	Distribution of studies according to the adopted registration method.	22
4.1	Overall proposed solution.	31
4.2	Equipment used for dataset acquisition: (a) Canon Aplio i800 system; and (b) PLT 1005BT linear probe.	33
4.3	Equipment used for real-time data capture: (a) GE LOGIQ P9 system; (b) L3-12 linear probe; and (c) UGREEN frame grabber.	33
4.4	Optitrack V120 Trio tracking system.	33
5.1	Simulated US scanning within the 3D Slicer, generating a B-mode image (right) from a specific probe position relative to the 3D bone volume extracted from CT (left).	37
5.2	Image acquisition protocol: (1–3) transverse sweeps of the femur (lateral, medial, and anterior), (4–6) transverse sweeps of the tibia, (7–9) longitudinal scans with the knee in full extension, and (10) longitudinal scan of the femoral condyles with the knee at 90° flexion.	38
5.3	Average number of slices for each region.	39
5.4	Distribution of pixel spacing values for the X (left) and Y (right) dimensions.	40
5.5	Customised 3D Slicer module engineered to optimise the manual annotation workflow for large-scale US datasets.	41

5.6	Custom module control panels: (a) dynamic image enhancement settings, and (b) segmentation overlay controls.	42
5.7	Integrated data integrity mechanisms: (a) unsaved progress warning, and (b) export completion verification.	43
5.8	Example of an expert-validated ground truth mask (green overlay) accurately delineating the hyperechoic bone surface.	44
5.9	Resampling and resizing outcomes: (a) normalisation of a larger initial pixel spacing (0.012 mm), and (b) normalisation of a smaller initial pixel spacing (0.008 mm).	46
5.10	Examples of the dynamic data augmentation techniques applied: (a) original slice, (b) scaled zoom, (c) variation in brightness and contrast levels, and (d) synthetic speckle noise injection.	47
5.11	Network architecture of the customised U-Net featuring a pre-trained VGG16 encoder backbone.	48
5.12	Qualitative comparison of the bone surface segmentation results (red lines) of the knee joint obtained by each developed architecture across five distinct anatomical views (columns A–E). Rows display the manual ground truth delineations, followed by the respective model predictions from the nnU-Net, UNet with VGG16, U-Net, and SegFormer models.	54
6.1	Hardware components utilised for system calibration: (a) US probe equipped with the custom-designed 3D-printed tracking reference, (b) pre-existing 3D-printed stylus, (c) fCal-2.1 calibration phantom (PLUS Toolkit standard), and (d) acoustically insulated water container.	58
6.2	Preliminary steps: (a) stylus pivot calibration, and (b) phantom registration.	59
6.3	Temporal calibration procedure: (a) physical setup for oscillatory motion, and (b) software interface for signal correlation.	60
6.4	Coordinate systems of the components involved in the calibration	60
6.5	Spatial calibration procedure: (a) probe manipulation over the N-wire fCal phantom, and (b) real-time algorithmic detection of the wire cross-sections.	61
6.6	Experimental setup showing the stationary physical femur phantom submerged in: (a) water, and (b) tissue-mimicking.	62
6.7	Workflow linking the data acquisition, the PLUS server, and the 3D Slicer client through the OpenIGTLink protocol.	62
6.8	'SlicerFreehandUS' module performing real-time automated segmentation.	63

6.9	'SlicerFreehandUS' rendering the reconstructed bone surface in real-time.	63
6.10	Ground truth volumetric reconstruction from preoperative CT data.	64
6.11	Sequential data refinement pipeline: transitioning from raw 3D volumetric reconstruction to a conditioned point cloud.	64
6.12	Visual comparison demonstrating the raw point cloud (green) and the resulting anatomically filtered structure (red).	65
6.13	Data refinement and normalisation: (a) geometric adaptation of the CT model to extract the anterior acoustic shell, and (b) spatial resolution equalisation for both CT and US point clouds.	67
6.14	Visualisation of the spatial alignment achieved by the two-stage automatic coarse registration algorithm.	68
6.15	Spatial partitioning of the femur point cloud into five distinct anatomical clusters for hold-out validation.	70
6.16	Statistical distribution of the system calibration metrics: (a) temporal latency (ms) for US-tracking synchronisation, and (b) spatial registration error (mm) evaluated under dynamic freehand motion. Boxplots illustrate the median, interquartile ranges, and outliers across the 15 independent acquisitions.	72
6.17	Temporal calibration results illustrating the fixed (blue line) and moving (red and green lines) tracking signals (a) before, and (b) after applying the calculated latency value. . .	72
6.18	Visual comparison of the spatial registration outcomes: (a) anatomically faithful alignment achieved by ICP algorithm, and (b) degenerate mathematical convergence produced by CPD algorithm.	76
A.1	Detailed technical drawing of the custom-designed US probe reference. The schematic includes the primary top-down view and the corresponding structural cross-section engineered for the integration of OptiTrack markers. All dimensional units are provided in millimetres.	96
B.1	Detailed technical drawing of the custom-designed reference frame for the calibration phantom. The schematic includes multiple structural views, such as the cross-section, explicitly engineered to accommodate the OptiTrack markers. All dimensional units are provided in millimetres.	97

List of Tables

2.1	Hounsfield unit (HU) for different tissues	7
3.1	Characteristics of the included studies	16
5.1	Demographic and clinical characteristics of the participants	39
5.2	Data distribution of training, validation and test data sets	45
5.3	Comparison of different architectures across performance metrics. Results are presented as mean $\pm$ standard deviation	52
6.1	Breakdown of average computational execution times for the 3D surface reconstruction pipeline	74
6.2	Comparative analysis of spatial registration metrics (ICP vs. CPD) across different acoustic propagation media	74
C.1	Individual latency measurements (ms) recorded during the temporal calibration procedure	98
C.2	Individual spatial error measurements (mm) obtained during the dynamic calibration process	99
C.3	Baseline spatial error measurements (mm) acquired during the static calibration test . .	99

List of Acronyms

2D: Two-dimensional

3D: Three-dimensional

BMI: Body Mass Index

CAOS: Computer-Assisted Orthopaedic Surgery

CNN: Convolutional Neural Network

CPD: Coherent Point Drift

CT: Computed Tomography

DL: Deep Learning

DSC: Dice Similarity Coefficient

FP: False Positive

FN: False Negative

GMM: Gaussian Mixture Model

GPU: Graphics Processing Unit

ICP: Iterative Closest Point

IoU: Intersection over Union

PC: Point Cloud

PET: Positron Emission Tomography

RGB: Red, Green, and Blue

RMS: Root Mean Square

RMSE: Root Mean Square Error

RQ: Research Question

SPECT: Single Photon Emission Computed Tomography

SR: Success Rate

TN: True Negative

TP: True Positive

TRE: Target Registration Error

US: Ultrasound

Chapter 1

Introduction

This dissertation presents the work carried out during the second year of the Master's in Medical Electronics at the University of Minho, during the 2024-2025 academic year. Half of the academic year was spent at the Laboratory of Neuroengineering and Medical Robotics (NearLab) located at the Politecnico di Milano, Milan, Italy. The other half was passed at the Biomedical Robotic Devices Laboratory (BiRDLab), included in the Center of MicroElectroMechanical Systems (CMEMS) Research Center, at the University of Minho, Guimarães, Portugal. During this period, a non-invasive bone registration solution for surgical navigation using ultrasound (US) was developed. This system was developed with the aim of reducing the invasiveness and complexity of computer-assisted orthopaedic surgery (CAOS) procedures.

1.1 Motivation and Problem Statement

Orthopaedic surgeries, such as knee arthroplasty and osteotomy, are heavily dependent on the surgeon's experience to manually identify anatomical landmarks, which serve as essential reference points to guide bone cuts and implant positioning [1; 2]. These landmarks are often occluded by soft tissue and blood, which restricts the intraoperative field of view and leaves a potential margin for error. Inaccurate bone cuts or slight mechanical misalignments in the knee can severely change the load distribution, which, alongside other clinical factors, significantly increase the risk of irreversible complications, premature implant failure and chronic joint instability [1]. To mitigate these risks, CAOS systems have emerged as a critical technological advancement. They provide intraoperative navigation, integrating preoperative planning directly into the surgical workflow, leading to better outcomes for patients [3–6].

However, to enable surgical navigation, preoperative models must be accurately aligned with the patient's intraoperative anatomy [7], a process known as registration [8; 9]. Currently, the clinical gold standard for bone registration relies on invasive techniques. This approach typically requires the surgeon to expose and drill into healthy bone segments to place fiducial markers directly onto the patient's skele-

ton to establish a rigid coordinate system [10; 11]. Although highly accurate, this invasive methodology presents serious clinical drawbacks. It inherently causes additional surgical trauma, increasing the risk of intraoperative haemorrhage and postoperative infections [12].

Alternative non-invasive registration methods, based primarily on medical imaging for intraoperative use, have been explored. Currently, volumetric imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI), form the cornerstones of preoperative imaging [7]. Applying these modalities to the intraoperative environment, however, presents significant challenges. MRI systems are fundamentally incompatible with standard orthopaedic operating rooms due to their bulk and strict magnetic safety constraints regarding metallic surgical instruments. On the other hand, CT exposes both the patient and the surgical team to harmful ionising radiation [13]. Consequently, there is a need to reduce intraoperative radiation exposure by transitioning to radiation-free imaging alternatives [1; 14].

US is an ideal imaging modality, as it is non-ionising and, more importantly, the high acoustic impedance contrast between soft tissues and bone generates a strong echo, making US highly suitable for obtaining bone surface profiles [14]. To fully exploit this physical advantage intraoperatively and overcome inherent US limitations like speckle noise and artefacts, it is crucial to integrate automatic segmentation, capable of robustly isolating the bone surface from these complex images. By successfully extracting these reliable anatomical profiles, we can enable an accurate US-based registration process. Consequently, there is a pressing need to develop robust, non-invasive registration solutions based on intraoperative US.

1.2 Goals and Research Questions

The main goal of this dissertation is to design, develop and validate a robust, non-invasive method of bone registration for surgical navigation using US. This approach provides two major clinical benefits, as it avoids the harmful intraoperative radiation exposure and significantly reduces surgical trauma by removing the reliance on invasive fiducial markers, providing better post-operative outcomes for the patient.

To achieve the main objective, the following objectives have been set:

Objective 1: To conduct a comprehensive literature review of the existing registration approaches. This is addressed in Chapter 3.

Objective 2: To define key clinical and technical specifications for the development of an intraoperative US registration system. This is addressed in Chapter 4.

Objective 3: To design and implement a robust segmentation pipeline for the automatic extraction of bone surfaces from US images. This is addressed in Chapter 5.

Objective 4: To develop and validate the proposed non-invasive registration workflow, assessing if its suitable for real-time surgical navigation. This is addressed in Chapter 6.

These objectives will provide answers to the following research questions (RQs):

RQ1: What are the current state-of-the-art techniques for bone registration in surgical navigation, and how do they compare? The answer is included in Chapter 3.

RQ2: What specific clinical and technical requirements must be met to develop a viable intraoperative US-based registration system? The answer is included in Chapter 4.

RQ3: Which deep learning (DL) algorithm is best suited for extracting bone contours from US images, to reliably reconstruct point clouds for registration? The answer is included in Chapter 5.

RQ4: How do the spatial accuracy and computational performance of the proposed non-invasive registration pipeline align with the clinical demands of surgical navigation? The answer is included in Chapter 6.

1.3 Contribution to Knowledge

The contributions of the present dissertation are as follows:

- A review of the state-of-the-art bone registration techniques for surgical navigation;
- A custom-built software module for the efficient and accurate annotation of medical US images;
- A new, comprehensively annotated, knee US dataset, an essential resource to address the scarcity of knee bone US data in the literature;
- Optimised DL algorithms tailored for the automatic extraction of bone contours from US images;
- An integrated non-invasive bone registration workflow capable of accurately aligning intraoperative reconstructed 3D US data with preoperative bone models;
- Empirical evidence validating the clinical feasibility, spatial accuracy, and computational performance of the proposed registration pipeline.

The work carried out during this dissertation has led to the writing of the following journal paper:

- **F. Silva**, D. Rito, C. Macedo, C. P. Santos, "Real-time automatic knee bone segmentation from 2D ultrasound images using deep learning methods" [submitting]

1.4 Dissertation Outline

This dissertation is organised into 7 chapters, as follows.

Chapter 2 provides the theoretical foundation for this dissertation. It begins with an overview of knee joint anatomy, followed by an exploration of non-invasive orthopaedic imaging modalities. Furthermore, the chapter explores CAOS systems, with a specific focus on surgical planning, navigation systems, tracking technologies, and registration methods.

Chapter 3 presents a comprehensive literature review of existing bone registration workflows. It critically analyses the state-of-the-art by comparing target anatomies, imaging modalities, hardware setups, tracking systems, registration algorithms, and their respective quantitative outcomes.

Chapter 4 outlines the proposed solution, defining the fundamental clinical and technical requirements for the registration system. It also details the hardware components and software architecture employed throughout the project.

Chapter 5 describes the development of the US segmentation pipeline. It details the entire workflow, from dataset construction and the development of a custom labelling tool, to data preprocessing, DL model architectures, training protocols, and evaluation metrics. The chapter concludes with a comparative and critical discussion of the segmentation results for different automatic bone contour extraction.

Chapter 6 details the implementation and evaluation of the 3D registration workflow. It begins with the calibration procedures essential for synchronising the coordinate systems of the optical tracker, the US probe, and the imaging data. Subsequent sections describe the acquisition protocol, freehand 3D reconstruction, and point cloud post-processing. Finally, the chapter presents the results of the non-invasive registration method, evaluating its spatial accuracy and performance using the established metrics.

Chapter 7 provides the final conclusions of this work. It summarises the main findings, directly answers the proposed RQs, and outlines directions for future work.

Chapter 2

Background

This chapter establishes the theoretical foundation necessary to contextualise the work developed in this dissertation. It introduces the fundamental concepts of bone registration for surgical navigation, with a specific focus on the anatomy of the knee joint, relevant medical imaging modalities, and the core principles of CAOS systems.

2.1 Anatomy of Knee Joint

The knee is the largest and one of the most heavily loaded joints in the human body, enduring constant and substantial biomechanical stress. Throughout human evolution, it has adapted to withstand significant loads and forces, functioning as a modified articulation that permits flexion and extension, alongside slight medial and lateral rotation, all while ensuring stability and motor control. While individual knee anatomy may exhibit slight natural variations, its complex biomechanical function remains remarkably consistent. This functionality relies on the intricate interplay between bony structures, ligaments, muscles, tendons, and the joint capsule [15]. The overall anatomy of the knee is illustrated in Figure 2.1.

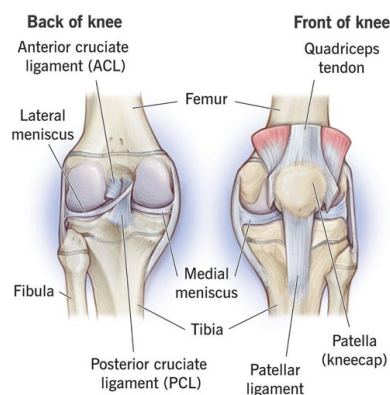


Figure 2.1: Anatomy of the knee joint.

The osseous part of the knee primarily consists of the distal end of the femur, the proximal end of the tibia, and the patella. The distal femur contains two prominent convex articular surfaces, the lateral and medial condyles. Anteriorly, these condyles merge to form a shallow depression known as the trochlear groove, which serves as the articulation site for the patella. The proximal tibia presents a relatively flat superior region known as the tibial plateau, which is divided into medial and lateral articular surfaces by the intercondylar eminence. Although situated laterally to the tibia, the fibula does not directly participate in the knee joint articulation [16].

The knee comprises two main articulations, the tibiofemoral and the patellofemoral joints. The tibiofemoral joint articulates the femur with the tibia, acting primarily as a hinge, while allowing slight rotation. Due to the inherent incongruity between the femoral and tibial articular surfaces, the C-shaped menisci sit between them to enhance joint congruency, prevent direct bone-to-bone contact, and function as crucial shock absorbers by evenly distributing compressive loads. Furthermore, the patellofemoral joint involves the posterior articular surface of the patella sliding along the femoral trochlear groove. In synergy with the surrounding musculature, this joint plays a vital role in absorbing and dissipating biomechanical forces during the gait cycle [17].

2.2 Medical Imaging Techniques in Orthopaedics

To assess the complex internal anatomy of the knee, medical imaging relies on a variety of modalities to accurately diagnose, monitor, and facilitate surgical planning. Each imaging technique exploits different physical principles to extract specific anatomical or functional information from the human body. This section explores the fundamental principles of the most widely used modalities: CT, MRI, Nuclear Imaging and US.

2.2.1 CT

Widely regarded as the primary modality for capturing detailed osseous structures, CT relies on the attenuation of ionising X-rays passing through the body. The transmitted X-ray intensity data is sent to a computer, where it is processed and displayed as two-dimensional (2D) cross-sectional images [18].

The technique involves the use of a rotating x-ray tube coupled with a set of detectors on the opposite side capable of collecting transmission measurements [19]. The resulting absorption or attenuation coefficients are mapped to a grayscale to construct the image. During the acquisition, the table slides with high precision so that the patient remains correctly positioned relative to the incident beam. The

resulting volumetric dataset comprises tens to hundreds of slices, each composed of three-dimensional (3D) volumetric elements called voxels.

While a portion of the X-rays passes directly through the tissue, some radiation is scattered. Scattering generates attenuation that can be measured using the Hounsfield unit. Table 2.1 shows typical values for differentiate tissues in the human body [18]. Consequently, CT provides high-resolution images that allow for the clear differentiation of structures such as bones, muscles, organs, fat and even blood vessels.

Table 2.1: Hounsfield unit for different tissues [18].

Tissue	Typical values
Bone	+1000
Blood	40
Muscle	10 – 40
Water	0
Fat	-50 – -100
Air	-1000

However, since CT relies on ionising radiation, patient exposure must be strictly optimised. Protocols must adhere to the As Low As Reasonably Achievable (ALARA) principle [20], and efficient hardware, such as Ultra-Fast Ceramic detectors with rapid decay and intense light output.

2.2.2 MRI

MRI offers a completely different physical approach to avoid ionising radiation. The scanner is typically a large, cylindrical device that generates a powerful static magnetic field and emits precise pulses of radiofrequency waves [21]. When placed in a magnetic field, particles with an intrinsic spin absorb radiofrequency photons at a specific resonance frequency (ω), defined by the Larmor equation (Equation 2.1), where B is the magnetic field strength and γ is the gyromagnetic ratio [22]. Consequently, stronger magnetic fields yield higher resonance frequencies, which generally translates to an improved signal-to-noise ratio and spatial resolution [18].

$$\omega = \gamma B \quad (2.1)$$

During the imaging process, three types of magnetic fields are combined. The first magnetic field polarises and orients the particles along the z-axis, aligning the spins. The second magnetic field performs transverse magnetisation, which makes it possible to assess the behaviour of the spins. The third magnetic field provides gradients in the x, y and z axes, allowing acquisition in different planes [18].

Thus, the image is created based on the response and relaxation times of the spin [23]:

- T1: The time taken for the longitudinal magnetisation to reach equilibrium. With this relaxation, energy is released into the extracellular matrix.
- T2: The relaxation time of the transverse magnetisation to zero, which is shorter than T1. It shows magnetic susceptibility and depends on the chemical composition of the material.

As a direct consequence of intense magnetic fields, strict safety protocols are imperative. Ferromagnetic metallic objects pose severe physical hazards, and even MRI-safe orthopaedic implants can generate significant susceptibility artefacts that severely distort the surrounding anatomical structures in the images. Despite these constraints, MRI offers a major safety advantage by avoiding ionising radiation, thereby eliminating the risk of radiation-induced cellular damage. Furthermore, this modality is highly regarded for its superior soft-tissue contrast and discrimination capabilities.

2.2.3 Nuclear Imaging

While modalities like CT and MRI excel at delineating anatomical structures, they provide limited insight into the underlying cellular and metabolic activity. When the clinical focus shifts from morphology to physiology, nuclear imaging becomes essential. This modality relies on the detection of radioactive labelled biochemical compounds (biological tracers), to map physiological and metabolic processes within the body. Upon administration, these unstable tracers undergo radioactive decay, emitting high-energy photons that are captured by external detectors to reconstruct their spatial distribution [24]. The two predominant techniques in this domain are positron emission tomography (PET) and single-photon emission computed tomography (SPECT).

PET utilises positron-emitting radionuclides. When an emitted positron encounters an electron, an annihilation event occurs, generating two gamma photons traveling in roughly opposite directions (180° apart). A surrounding ring of detectors registers these simultaneous hits (coincidences), establishing lines of response that allow computers to reconstruct highly accurate 3D maps of tracer concentration [25].

Conversely, SPECT relies on tracers that emit single gamma photons. The spatial origin of these photons is determined using heavy lead collimators that only allow perpendicularly traveling rays to reach

the detectors. The detectors—typically sodium iodide scintillation crystals—convert the gamma energy into visible light, which is then translated into electrical signals to form the final image [26].

In modern clinical practice, PET and SPECT are frequently integrated with CT or MRI hardware (e.g., PET/CT, SPECT/CT). This multimodal approach provides fused images that combine exquisite metabolic data with precise anatomical localisation [27]. In orthopaedics, these hybrid techniques are invaluable for evaluating complex trauma, detecting periprosthetic joint infections in total arthroplasties, and staging musculoskeletal oncology [28].

The primary advantage of nuclear imaging is its exceptional functional sensitivity, enabling the detection of subtle metabolic abnormalities long before structural changes become visible on CT or MRI [24]. Nevertheless, the modality faces logistical limitations, including high operational costs, reliance on isotopes with short half-lives, lower spatial resolution, the inherent necessity of administering radioactive materials to the patient, and the occurrence of artefacts caused by photon scattering and attenuation [25].

2.2.4 US

In stark contrast to the ionising radiation of CT and nuclear imaging, or the powerful magnetic fields utilised in MRI, US imaging employs mechanical sound waves, typically in the frequency range of 1 to 20 megahertz (MHz), to probe the body's internal structures [18]. Because it does not use ionising radiation, it is considered highly safe. However, prolonged exposure at high intensities can induce thermal and mechanical bio-effects. To mitigate this, ALARA principles also apply to US, dictating minimal scan times and optimal output power. Clinically, US is highly valued for being cost-effective, portable, and capable of 2D real-time imaging, unlike the modalities mentioned above [29].

Images are generated using a transducer that emits short bursts of pressure waves and records the returning echoes. These longitudinal mechanical waves propagate by sequentially compressing and rarefying the medium's particles. In addition, the transducer also receives the echoes and sends them to the computer, which creates the sonogram. This is possible due to the presence of several piezoelectric crystals, which, when an electric current is applied to them, generate sound, a phenomenon known as the piezoelectric effect. The piezoelectric effect also works in reverse, when sound waves hit the crystals, they vibrate and produce a voltage, which can be measured as a signal [18]. The internal structure of a typical US transducer is schematised in Figure 2.2.

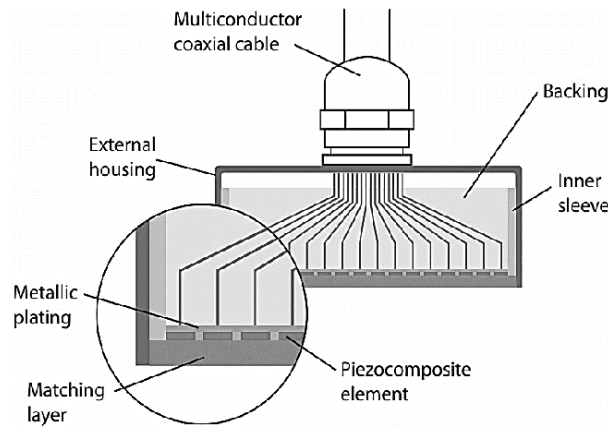


Figure 2.2: Ultrasound phased array transducer.

During the interaction of US waves with human tissue, transmission, dispersion, heat transformation, and reflection may happen. It is through reflection that images can be constructed. This happens because the echoes received by the transducer depend on the acoustic impedance of the tissue, in specific, the greater the contrasting impedance the stronger are the echoes [18].

To accurately map depth, the system calculates the time elapsed between pulse emission and echo reception [18]. This relies on the established propagation velocity of sound, which averages 1540 m/s in soft tissue, but varies significantly (e.g., 1450 m/s in fat and up to 4000 m/s in cortical bone) [30]. While US excels in providing dynamic, real-time feedback, it is fundamentally limited by user dependency, a restricted field of view, and susceptibility to imaging artefacts (such as acoustic shadowing, reverberation, and speckle noise) that can degrade image clarity.

2.3 CAOS systems

The concept of CAOS was first introduced in 1992 [31], leveraging the unique mechanical properties of the skeletal system to enhance surgical precision. Unlike soft tissues, bones are non-deformable, rigid and clearly distinguishable in medical images, making them the most suitable tissue for computer-assisted surgery. Consequently, CAOS provides better surgical results, lower learning curves for novice surgeons and less invasiveness of the procedure than traditional techniques [32]. Depending on the level of interaction between the system and the surgeon, these architectures are classified into three distinct categories [33]:

- **Passive systems:** The computer and navigation devices help the surgeon during planning and provide information during surgery, but do not act directly. The surgeon controls all phases of the surgery.

- Active systems: Surgical robots autonomously perform some steps during surgery.
- Semi-active systems: The surgeon and the robot have combined interactions, such as robotic arms that provide a guiding hand, but do not perform the surgery.

Regardless of the system's level of autonomy, the clinical execution of CAOS is consistently structured into two primary stages: preoperative and intraoperative. In the preoperative phase, surgical planning is obtained, while in the intraoperative phase, the surgical navigation is performed [31].

2.3.1 Surgical Planning

During the preoperative phase, computerised surgical plans are developed for a specific patient, based on images captured prior to surgery and using a dedicated software [34]. The result is a patient-specific surgical strategy that may involve the generation of a precise virtual model. This model is useful for defining the surgical steps, with the patient's anatomy virtually reconstructed and analysed, allowing the surgeon to simulate the procedure, predict outcomes, and identify potential complications [35]. Moreover, the incorporation of patient-specific customisations allows the surgery to be tailored to each patient's individual anatomical characteristics, contributing to a more efficient post-operative recovery [36].

However, the planning approach has some drawbacks that span both the acquisition and processing phases. Its reliability is constrained by the initial quality of the medical images, which can be compromised by acquisition artefacts and limited spatial resolution. Furthermore, during the subsequent image processing pipeline, inaccuracies in anatomical segmentation can introduce critical errors into the virtual model. Although the images themselves are typically already acquired during the routine diagnostic phase, the subsequent processing and 3D analysis require dedicated time, adding to the preoperative workload. The reliance on adequate surgical training may pose challenges for its clinical implementation [36].

2.3.2 Surgical Navigation

In the intraoperative phase, the patient's physical anatomy is registered to the preoperative virtual model. Following this registration, real-time tracking systems enable continuous surgical navigation, providing guidance for the procedure. Thus, the surgeon can reach the planned areas without direct visual contact, in order to delimitate the safe zones, preventing the device from approaching critical areas [37].

Tracking

The tracking system makes it possible to localise and orientate surgical devices in relation to the anatomical structures. The accuracy, robustness, and responsiveness of tracking are vital for ensuring a correspondence between the preoperative plan and intraoperative execution. The main tracking methods used in surgical navigation include [38]:

- Mechanical tracking: Robot kinematics combined with joint sensing are used to compute an end-effector position. However, accuracy depends on the correct kinematics and the physical equipment can limit the workspace, being obtrusive.
- Optical tracking: Benefits from camera rigs to track fiducial markers that are attached to the instrument or anatomical structure of interest. Active markers emit typically infrared light and are activated by an electrical signal. In contrast, passive markers are spherical reflectors, that can either bounce back infrared light emitted by the position sensor, or distinct visual patterns recognized by red, green and blue (RGB) cameras. Another approach includes radiographic markers, also passive, which are visible only under X-ray or CT imaging.
- Electromagnetic tracking: A magnetic field generator transmitter induces a current in sensor coils, that can be embedded into the tracked objects.

Registration

Image registration is a crucial step in CAOS, as it performs a spatial transformation to align a set of images that have a common reference coordinate system, in order to enable real-time navigation [10]. Registration, according to Figure 2.3, can be classified based on modality-driven approaches, such as unimodal or multimodal, at the patient-level, as interpatient or inpatient, and types of transformation, namely rigid or deformable [39].

The unimodal approach involves registering images obtained by the same modalities. Its main advantage lies in the ability to correct spatial discrepancies and easily align the anatomical structures or features present in the images. Furthermore, this registration is the simplest, since images from the same modality share intensity characteristics and feature representations [40]. Meanwhile, the multimodal approach involves aligning images obtained by different modalities, perspectives or different parameters that cause significant differences in the image [41]. The fact that it comes from different modalities makes it possible to combine different anatomical features visible in each modality.

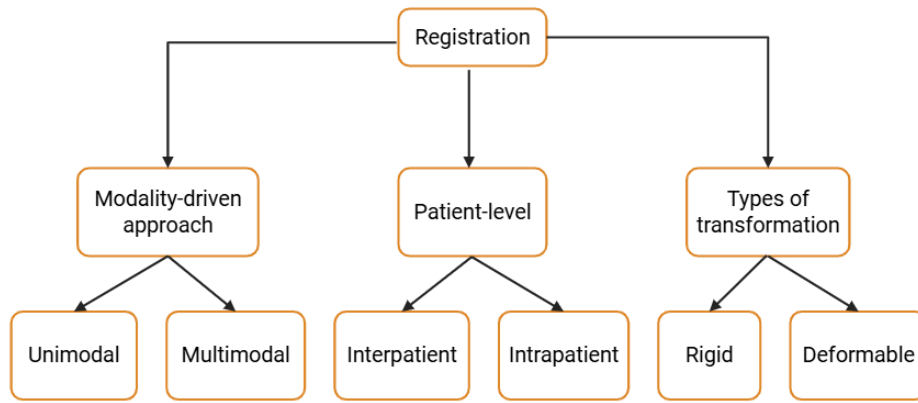


Figure 2.3: Categorisation of registration approaches.

Interpatient registration occurs between different patients with a standard frame of reference. This can be challenging due to differences in the overlap of anatomical structures between individuals. Intrapatient registration, by contrast, aligns images obtained from the same patient on the same or different days. The challenges of this registration reside in changes to the images due to metabolic processes or weight fluctuations. This approach is important for monitoring the evolution of a pathology in a patient [42].

In terms of the type of transformation, rigid registration uses a simple transformation applied uniformly, which is why it is useful for non-deforming structures, such as bones, or for intrapatient registration, as there are fewer discrepancies. Deformable registration is applied non-uniformly, taking discrepancies into account, as the image is deformed to match the other, making it useful for interpatient registration [3].

Chapter 3

Review of Bone Registration Techniques for Surgical Navigation

As an efficient CAOS system relies on accurate registration, so a systematic review of the state-of-the-art in bone registration was conducted. While previous reviews have extensively mapped the general use of computer-guided technology in orthopaedic surgery [5; 7; 43–45] or focused on its application in specific anatomical regions such as the knee [46; 47], there is a lack of systematic reviews dedicated exclusively to the underlying registration methodologies. To bridge this gap, this chapter aims to answer the following **RQ**: *What are the current state-of-the-art techniques for bone registration in surgical navigation, and how do they compare?*

3.1 Methodology

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The literature search concluded on October 10, 2024, and was performed across four databases: Pubmed, Scopus, Web of Science and IEEE. The search used the following combination of keywords: (bone or orthopedic or orthopaedic) AND registration AND (navigation or navigated), and limited to the title, abstract and keywords. Only studies published after 2017 were considered.

Studies were excluded if: 1) The access was not possible; 2) The article was not written in English; 3) The publication was a review article; 4) It addresses non-orthopaedic procedures; 5) It was not developed in a surgical context; 6) Focuses on extended reality; 7) Lack of information on the registration method.

3.2 Results

The initial literature search yielded 1163 records across the four databases. Following the removal of duplicates, 700 unique studies remained. During the initial screening of titles and abstracts, 574 articles were excluded as they were not relevant to the scope of this review. The remaining articles underwent a full-text assessment, leading to the exclusion of a further 89 studies based on the predefined exclusion criteria. Ultimately, a total of 37 articles met all criteria and were included in this review. The complete selection process is illustrated in the PRISMA flow diagram (Figure 3.1).

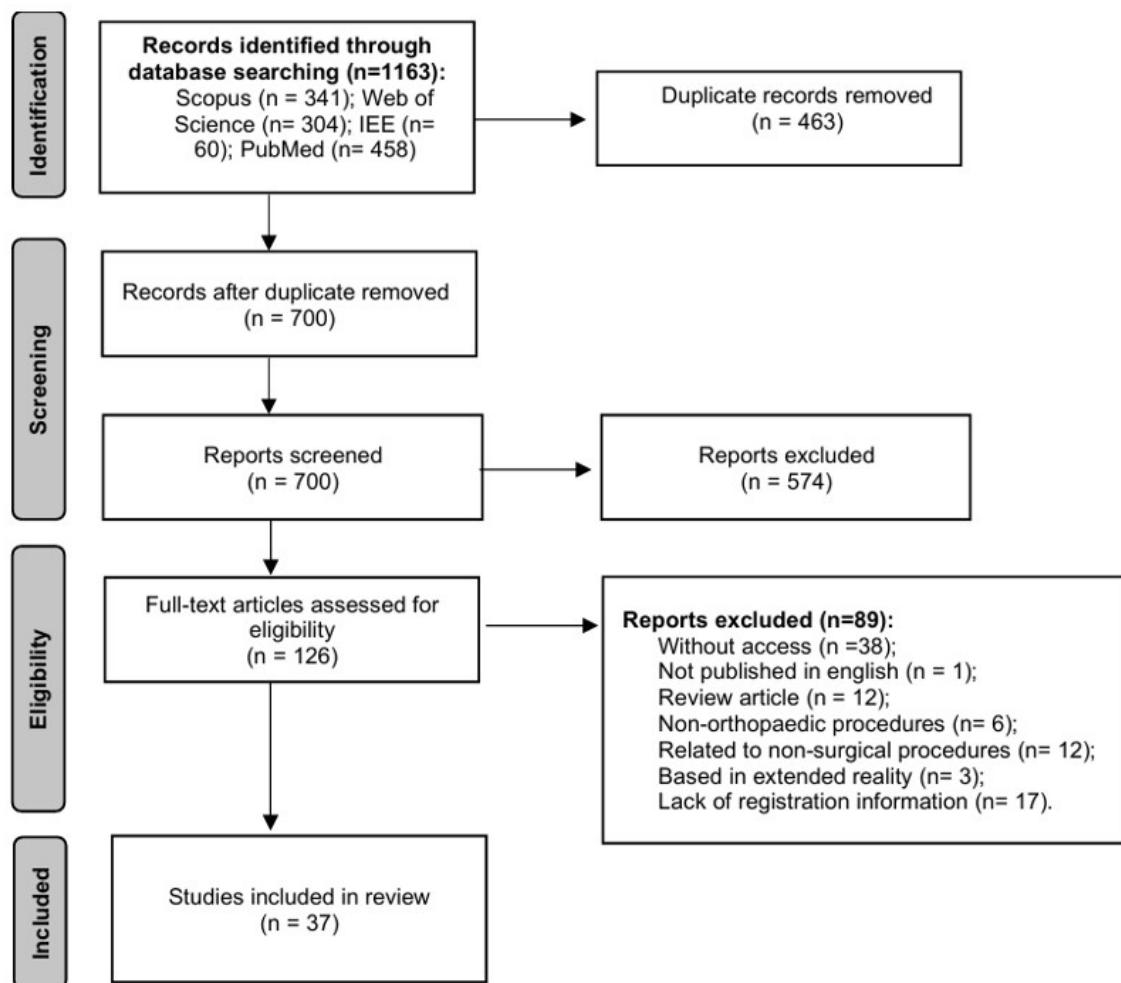


Figure 3.1: Flow diagram of search strategy based on PRISMA.

Data from the included studies were extracted and analysed based on key characteristics. These included the target anatomical body area, preoperative and intraoperative imaging modalities, hardware setup, registration method, tracking system, and quantitative outcomes, as detailed in Table 3.1.

Table 3.1: Characteristics of the included studies

Study	Anatomical body area	Modality-driven for preoperative/intra-operative	Hardware	Registration method	Tracking	Quantitative outcome
Barbosa et al. (2017) [48]	Orthopaedic, Un-specified Region (UR)	Computed Tomography (CT)/Not Applied (NA)	Polaris Vicra (NDI)	Point-to-point completed with Iterative Closest Point (ICP)	Optical	Average of distances between the points collected and the nearest point = 0.2719 mm; Standard deviation = 0.1130 mm; Root Mean Square Error (RMSE) = 0.2923 mm
Goerres et al. (2017a) [49]	Pelvis	CT/Fluoroscopy	Not Mentioned (NM)	Deformable Known-Component (dKC)	NA	RMSE (surface) = 2.3 ± 0.6 mm; RMSE (after 40 iterations) = 4 mm; RMSE (after 130 iterations) = 2 mm
Goerres et al. (2017b) [50]	Spine	CT/Fluoroscopy	NM	Deformable Coherent Point Drift (CPD)	NA	Minimum (min) surface error = 0.89 ± 0.10 mm; Maximum (max) surface error = 1.29 ± 0.15 mm
Vlachopoulos et al.(2017) [51]	Humerus	CT/CT	NM	ICP	NA	Mean (m) hit count ratio = $2.3 \pm 1.1^\circ$
Fanti et al.(2018) [52]	Tibia	CT/Ultrasound (US)	Aloka 1000 (Hitachi); Polaris Spectra (NDI)	ICP	Optical	max Target Registration Error (TRE) < 1.55 mm; mTRE = 0.64 mm
Niu et al.(2018) [53]	Femur	CT/US	Aquilion ONE (Toshiba); Imasonic SAS probe; Visualeyex VZ4000v (PTI)	ICP-Perturbation Search (PS) , ICP	Optical	RMSE (ICP farthest case) = 5.88 mm; RMSE (ICP-PS farthest case) = 3.96 mm; RMSE (ICP closest case) = 5.84 mm; RMSE (ICP-PS closest case) = 2.81 mm; Unknown Point Localisation Error (UPL) (farthest case) ≤ 4.25 mm; UPL (closest case) ≥ 3.68 mm
Pandey et al.(2018) [54]	Pelvis	CT/US	CT750HD Scanner (GE Healthcare); SonixTouch Q+ with 4DL14-5/38 probe (Ultrasonix); Polaris (NDI)	Gaussian Mixture Model (GMM), CPD, Normalised Cross-Correlation (NCC)	Optical	Median (med) TRE (CPD) = 3.19 mm; mTime (CPD) = 4.2 s; medTRE (NCC) = 2.44 mm; mTime (NCC) = 27.3 s
Raposo et al.(2018)[55]	Orthopaedic UR	CT or MRI/Volumetric images	Optotrak Certus (NDI)	Curve and surface	Optical	Rotation error, translation error and time represented graphically
Silva et al.(2018) [56]	Spine	MRI/US	E-Cube 12R with 4DSCA01 probe (Alpinion)	NCC	NA	TRE (< 8 mm), registration error and accuracy represented graphically (< 3 mm)

Continued on next page

Table 3.1: (continued)

Study	Anatomical body area	Modality-driven for preoperative/intra-operative	Hardware	Registration method	Tracking	Quantitative outcome
Wang et al.(2018) [57]	Orthopaedic UR	CT/US	Vivid Portable US (GE Healthcare); Polaris (NDI)	ICP, CPD	Optical	Average translation error of the centre point position for each cavity (CPD): 1.5, 1.31, 1.19 mm; Average rotational error of the centre point position for each cavity(CPD): 0.85, 0.61, 0.99°
Gueziri et al.(2019a) [58]	Spine	CT/US	Optical Tracking Camera (NDI); HDI 5000 with P4-7 probe (Philips)	Gradient alignment	Optical	Time = 11 s; medAccuracy = 1.48 mm; Success Rate (SR) = 84.42%
Gueziri et al.(2019b) [59]	Spine	CT/US	Optical Tracking Camera (NDI); HDI 5000 with P4-7 probe (Philips)	Gradient alignment	Optical	medTRE = 1.48 mm; SR = 84.42%; Fiducial Registration Error (FRE) = 0.328 mm; Time = 10.79 ± 1.27 s
Guha et al.(2019) [60]	Spine	CT/Optical Topographic Imaging (OTI)	NM	ICP	NA	Registration error (cilindrical symmetry) = 0.315 ± 0.078 mm; Registration error (spherical symmetry) = 0.375 ± 0.082 mm; Registration error (planar symmetry) = 0.292 ± 0.054 mm
Gueziri et al.(2020) [61]	Spine	CT/US	FusionTrack 500 (Atracsys); P4-7 probe (Philips)	Gradient alignment and/or intensity	Optical	TRE (Gradient) = 1.27 mm; SR (Gradient) = 72.43%; Time (Gradient) = 5.57 s; TRE (Intensity) = 1.70 mm; SR (Intensity) = 53%; Time (Intensity) = 5.25 s; TRE (Combined) = 1.52 mm; SR (combined) = 83.14%; Time (Combined) = 6.62 s
Han et al.(2020) [62]	Pelvis	CT/Fluoroscopy	NM	Gradient alignment	NA	TRE (phantom) = 2.5 mm; Positional Deviation Error (PDE) (clinical study) = 3.1 mm; Time = [1,2] min
Liu et al.(2020) [63]	Femur	CT/Camera with Red, Green, and Blue detection (RGB)	RealSense D415 (Intel); FusionTrack 500 (Atracsys)	Point-to-plane ICP	Optical	Average translation error = 2.74 mm; Average rotational error = 6.66°
Min et al.(2020) [64]	Femur	CT /NA	NM	Expectation-maximisation	Optical	Translation RMSE < 1 mm; Rotational RMSE < 0.05°

Continued on next page

Table 3.1: (continued)

Study	Anatomical body area	Modality-driven for preoperative/intra-operative	Hardware	Registration method	Tracking	Quantitative outcome
Scharffet et al.(2020a) [65]	Spine	Volumetric images/Fluoroscopy	NM	Point-to-plane ICP	NA	min mTRE = 0.32 ± 0.13 mm; max mTRE = 0.49 ± 0.32 mm; maxSR = 87.3%; minTime = 7 ± 1.4 s
Scharffet et al.(2020b) [66]	Spine	CT/X-ray	NM	Point-to- plane ICP	Optical	min mTRE = 0.46 ± 0.19 mm; maxSR = 99.7%
Zhu et al.(2020) [67]	Femur	CT/NA	NM	ICP	NA	Average registration error = 2.9 voxels; Time = 8.64s
Chan et al.(2021) [68]	Spine	CT/US	OptiTrack Prime 13W (NaturalPoint); SonixTablet with L14-5/38 probe (BK Ultrasound)	Gaussian pyramid-based, ICP	Optical	Translational accuracy = 1.2 ± 0.5 mm; Rotational accuracy = $0.2 \pm 0.4^\circ$; SR = 89.3%
Han et al.(2021) [69]	Pelvis	CT/Fluoroscopy	Somatom Definition Flash (Siemens)	Shape-based	NA	Rotational error = $3.3 \pm 2.1^\circ$; Translational error = 2.8 ± 1.5 mm
Hu et al.(2021) [70]	Femur	NA	HDI Compact 3D Scanner (LMI Technologies)	Bounded ICP	Optical	Rotational error = $0.81 \pm 0.44^\circ$; Translational error = 4.75 ± 0.27 mm
Min et al.(2021) [71]	Femur; Pelvis	Volumetric images/ NA	Optical Tracking System (NDI)	Rigid point set registration with normal and tangent vectors	Optical	mTRE < 0.3 mm; mRotational error < 0.1° ; mTranslation error < 0.3 mm
Wei et al.(2021)[72]	Femur	NM	NM	Simple Consensus Initial Alignment followed by ICP	Optical	Registration error = 2.12 mm
Abe et al.(2022) [73]	Femur	CT/ Arthroscopy camera	Sensation 16 Scanner (Siemens)	ICP	NM	Resection error > 131 mm
Chan et al.(2022) [74]	Spine	CT/US	OptiTrack Prime 13W (NaturalPoint); SonixTablet with high-frequency linear transducer (BK Ultrasound)	Intensity-based followed by ICP	Optical	Average accuracy = 0.4 ± 0.4 mm and $2.1 \pm 0.9^\circ$; Time = 8.8 s
Min et al.(2022) [75]	Femur; Pelvis	NM	NM	Expectation-maximisation	NA	Rotational error = 0.5747° ; Translational error = 1.293 mm

Continued on next page

Table 3.1: (continued)

Study	Anatomical body area	Modality-driven for preoperative/intra-operative	Hardware	Registration method	Tracking	Quantitative outcome
Naik et al.(2022) [76]	Spine	CT/Fluoroscopy	BV Endura (Philips); Ziehm 8000 C-arm (Ziehm)	ICP with digitally reconstructed radiographs	Optical	$0.5 \text{ mm} \geq \text{mEuclidean distance} \leq 1.6 \text{ mm}$; mTrajectory length error = 2.67 mm; Time $\approx 3 \text{ min}$ per vertebra
Bröbner et al.(2023) [4]	Scaphoid	CT/US	FusionTrack 500 (Atracsys); Sonix-TOUCH with 4DL14-5/38 probe (Ultrasonix)	Point Registration net and Robust Point Matching (RPM) net with ICP, ICP, GMM	Optical	Translational error (ICP) = $5.9 \pm 3.2 \text{ mm}$; Rotational error (ICP) = $48.5 \pm 31.5^\circ$; Rotational error (RPMNET + ICP) = $1.4 \pm 0.7^\circ$; Translational error (RPMNET + ICP) = $0.3 \pm 0.1 \text{ mm}$
Liu et al.(2023) [77]	Tibia	CT/NA	Polaris Vega (NDI)	Steepest Perturbation Search-ICP	Optical	FRE = $0.80 \pm 0.30 \text{ mm}$; Rotational error $< 1^\circ$; Translational error $< 1.5 \text{ mm}$; Average time $< 40 \text{ s}$
Shrestha et al.(2023) [78]	Pelvis	CT/X-ray	NM	Convolutional Neural Network (CNN)	NA	mTRE = $9.65 \pm 4.07 \text{ mm}$; Time = 1.75 s; Good Feature Ratio (GFR) $> 20\%$
Xie et al.(2023) [79]	Pelvis	CT/NA	MV-CA050-10GM and MVL-KF1628M-12MPE (Hikrobot)	Fiducial-based	Optical	mRMSE = 0.79 mm
Liu et al.(2024) [80]	Knee	CT/US	Polaris Vega (NDI); DP-20 with linear probe (Mindray)	ICP Sound speed of the soft tissue	Optical	Time (femur) = 120.40 s; Time (tibia/fibula) = 254.57 s; mTRE = [0.58, 1.63] mm; FRE (femur) = $0.98 \pm 0.88 \text{ mm}$; FRE (tibia/ fibula) = $1.29 \pm 0.19 \text{ mm}$
Li et al.(2024) [81]	Spine	CT/X-ray	NM	Discrete Cosine Transform net and Convolutional Recurrent (CR) net	NM	Percentage of Landmark Correspondence (PLC) = 99.997%; RMSE = 0.1707%
Pandey et al.(2024) [82]	Pelvis	CT/US	Vega (NDI); Telemed US system	NCC	Optical	medTRE = 9.7 mm; medFRE = 0.7 mm
Zhang et al.(2024) [83]	Femur; Hip	CT/NA	Polaris P4 (NDI)	Generative Hybrid Mixture Model	Optical	$2.78 \pm 1.65^\circ \leq \text{Rotational error} \leq 7.28 \pm 3.55^\circ$; $2.64 \pm 1.73 \text{ mm} \leq \text{Translational error} \leq 5.62 \pm 2.66 \text{ mm}$; $3.10 \pm 1.71 \text{ mm} \leq \text{RMSE} \leq 6.95 \pm 3.86 \text{ mm}$; $84.88\% \leq \text{Recall} \leq 99.66\%$

3.2.1 Anatomical Body Area

In this review, only bone-related registration methodologies were considered. Within this scope, the distribution across anatomical regions is illustrated in Figure 3.2. Of the 37 studies, the spine was the most studied area with 12 studies [50; 56; 58–61; 65; 66; 68; 74; 76; 81]. The femur was covered in 10 studies [53; 63; 64; 67; 70–73; 75; 83], and the pelvis in 9 studies [49; 54; 62; 69; 71; 75; 78; 79; 82]. Few studies referred to the tibia [52; 77], humerus [51], scaphoid [4], knee [80] and hip [83].

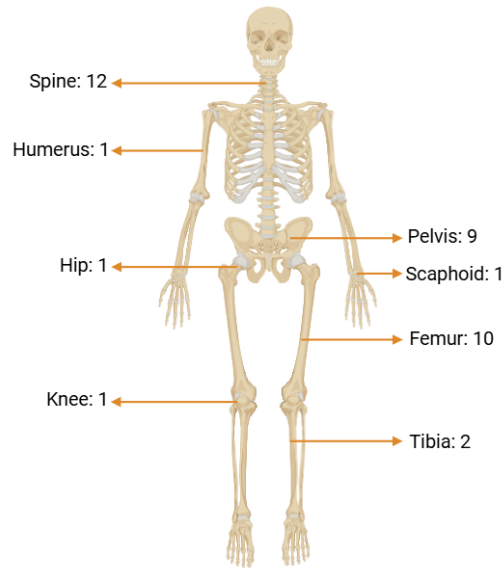


Figure 3.2: Number of studies per anatomical body area.

3.2.2 Medical Imaging Modality

This subsection details the imaging modalities utilised across the included studies during both the preoperative and intraoperative phases. Because the specific characteristics of the chosen imaging technology inherently influence the registration process, these modalities were systematically extracted and analysed. A comprehensive visual distribution of the adopted preoperative and intraoperative modalities is presented in Figure 3.3.

The vast majority of the included studies utilised a preoperative imaging modality. The exceptions were Wei et al. [72] and Min et al. [75], which did not report the modality used, and Hu et al. [70], where preoperative imaging was not performed. CT was the predominant choice, employed by 29 studies to acquire preoperative anatomical data [4; 48–54; 57–64; 66–69; 73; 74; 76–83]. Conversely, Silva et al. [56] relied on MRI. Finally, Schaffert et al. [65], Min et al. [71], and Raposo et al. [55] did not explicitly state the exact imaging technique, noting only the use of generic volumetric data.

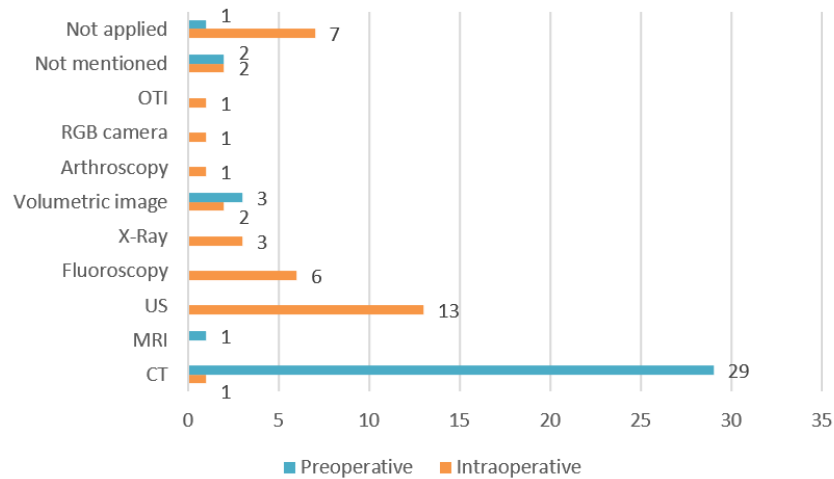


Figure 3.3: Number of studies per preoperative and intraoperative data acquisition modality.

Regarding intraoperative data acquisition, US emerged as the most widely used modality, featured in 13 published studies [4; 52–54; 56–59; 61; 68; 74; 80; 82]. Fluoroscopy was the second most common, utilised in 6 articles [49; 50; 62; 65; 69; 76]. Additionally, 3 articles utilised standard X-ray imaging [66; 78; 81]. Beyond conventional medical imaging such as CT [51], various optical and computer vision-based sensors were explored for intraoperative data acquisition. This category includes Optical Topographic Imaging (OTI) [60], as well as images from RGB cameras [63] and arthroscopic video [73]. Alternatively, rather than relying on intraoperative imaging, 7 studies performed surface digitization via physical point acquisition using a tracked pointer [48; 64; 67; 70; 71; 77; 79; 83]. Similar to the preoperative phase, 1 study [55], reported using generic volumetric images without specifying the exact technology, while Wei et al. [72] and Min et al. [75] omitted this information.

3.2.3 Hardware

The selection of appropriate hardware is a fundamental determinant of the accuracy, feasibility, and overall success of any surgical registration pipeline. Across the reviewed literature, hardware configurations can be broadly divided into two primary categories: imaging acquisition systems (responsible for capturing the anatomical data) and spatial tracking systems (essential for localising instruments and sensors in 3D space). A comprehensive, study-by-study compilation of all specific hardware models is provided in Table 3.1.

Regarding imaging acquisition, preoperative data was predominantly acquired using standard, clinical-grade CT scanners from major medical manufacturers (Siemens, GE Healthcare, Toshiba) [53; 54; 69; 73]. In contrast, intraoperative acquisition exhibited a much wider diversity of specialised hardware, reflecting

the varying experimental approaches. US imaging frequently relied on the Sonix line (BK Ultrasound/Ultrasoundix) [4; 68; 74], Aloka [52], Philips [58] and Alpinion devices [56]. The performance of these systems is intrinsically linked to the connected transducers, for instance, the L14-5/38 linear probe (BK Ultrasound) was widely adopted for its high-frequency surface resolution [4; 54; 68]. Other configurations utilised the 4DSCA01 volumetric probe (Alpinion) for 3D data capture [56] or the DP-20 (Mindray) [80]. Other intraoperative imaging setups incorporated mobile C-arms for fluoroscopy (Philips BV Endura, Ziehm 8000) [76], RGB depth cameras (Intel RealSense) [63; 79], and structured light 3D point cloud scanners [70].

In addition to imaging devices, dedicated hardware is required for spatial tracking. The NDI Polaris family (Northern Digital Inc.) was the most widely adopted optical tracking hardware [52; 54; 56; 57; 59; 64; 77; 80; 83]. Other high-end systems included the Atracsys FusionTrack 500 [4; 61; 63] and the OptiTrack Prime 13W [68; 74]. The specific methodological application of these tracking systems is detailed in the subsequent section.

3.2.4 Registration Method

Registration consists of calculating the optimal spatial transformation that aligns preoperative with intraoperative data [80]. Because the accuracy of this alignment directly dictates the precision of the surgical guidance, the registration process is widely considered the most critical computational step in any surgical navigation pipeline [71].

All the reviewed studies employed rigid registration, with the notable exception of Goerres et al. [49; 50]. The primary objective of rigid transformation is to compute a precise rotation matrix and translation vector that perfectly aligns the corresponding points without deforming them. Based on their underlying mathematical architectures, the registration algorithms identified in the literature were classified into four main categories: classical, probabilistic, learning-based, and combined methods. The distribution of these methodologies is illustrated in Figure 3.4.

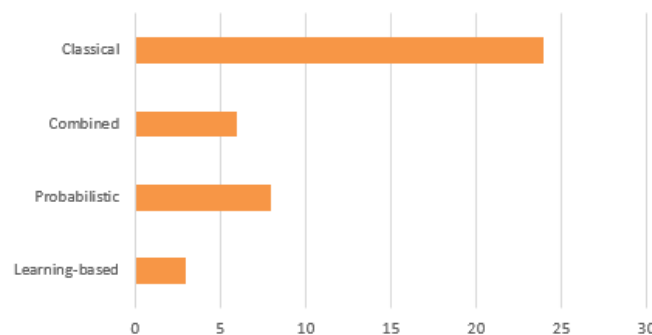


Figure 3.4: Distribution of studies according to the adopted registration method.

Classical Methods

Classical registration methods rely on traditional geometric and mathematical approaches, typically aiming to minimise a predefined error function to achieve data alignment.

Point-based algorithms were the most prevalent, with the iterative closest point (ICP) algorithm serving as the foundation in 9 studies [4; 51–53; 57; 60; 67; 68; 73]. ICP operates iteratively: it first establishes correspondences by finding the closest point pairs (minimising the root mean square (RMS) distance), then calculates and applies the rigid transformation matrix, repeating this cycle until convergence criteria are met. While more iterations generally increase accuracy, Zhu et al. [67] demonstrated that exceeding an optimal threshold can cause instability and metric fluctuation. Several studies employed customised ICP variants. For instance, search-based optimisations like perturbation search (PS) and steepest perturbation search were integrated into ICP by Niu et al. [53] and Liu et al. [77], respectively. Other researchers focused on geometric constraints: Liu et al. [63] and Schaffert et al. [65; 66] implemented point-to-plane ICP, which aligns points to tangent surfaces, while Min et al. [71] incorporated normal and tangent vectors. Alternatively, Hu et al. [70] utilised bounded ICP, which forces alignment using distant point pairs to restrict distal errors and reduce convergence degrees of freedom. Most recently, Liu et al. [80] proposed an ICP variant specifically calibrated for the variable sound speed of soft tissues.

Feature-based methods, such as Gradient Alignment, were explored in 4 studies [58; 59; 61; 62]. It operates by maximising the coincidence between intensity gradients found in preoperative and intraoperative images. Additionally, Xie et al. [79] addressed fiducial-based registration, which relies on tracking artificial markers or distinct anatomical landmarks with known spatial correlations across coordinate systems.

Correlation and intensity-based algorithms were adopted by 3 studies using normalised cross-correlation (NCC) [54; 56; 82]. It calculates the cross-correlation between volumetric points, maximising intensity overlap. Its primary strength lies in its geometric simplicity and feature-agnostic nature.

Contour and shape-based methods were less common. Raposo et al. [55] developed curve and surface registration to align specific geometric features, such as curves, rather than the entire point cloud. An earlier approach by Han et al. [69] utilised Statistical Shape Models (SSM) to register bone fragments.

Probabilistic Methods

Probabilistic methods, identified in 7 studies, frame registration as a maximum likelihood estimation problem, utilising probability density functions to represent point clouds. This approach is highly robust to noise and outliers. Pandey et al. [54] and Wang et al. [57] applied coherent point drift (CPD), which

conceptualizes the alignment as a gaussian mixture model (GMM) fitting process. Min et al. [64; 75] expanded on this statistical framework using expectation-maximisation to iteratively update transformation parameters. Regarding structural variations, Chan et al. [68] utilised Gaussian pyramids, while Pandey et al. [54] and Bröbner et al. [4] implemented GMM trees. In these tree models, point clouds are represented as a combination of Gaussian distributions, allowing the algorithm to find the most probable point correspondences even in highly noisy datasets.

Learning-based Methods

Learning-based methods leverage artificial intelligence to automatically extract high-dimensional geometric features and detect keypoints for alignment, bypassing traditional iterative search processes. Shrestha et al. [78] employed a fully convolutional neural network (CNN) for spatial registration. Li et al. [81] utilised a combination of a discrete cosine transform net and a Convolutional Recurrent (CR) net. Lastly, Zhang et al. [83] developed a generative hybrid mixture model network, innovatively combining GMMs with Fisher mixture models.

Combined Methods

Combined (or hybrid) methods integrate multiple computational approaches to overcome individual algorithmic limitations, maximising both accuracy and robustness. 6 notable studies adopted this strategy, all of which utilised ICP as the core refinement step following an initial coarse alignment. Barbosa et al. [48] performed initial point-to-point matching before executing ICP, whereas Wei et al. [72] achieved initial alignment using the sample consensus initial alignment algorithm. Other frameworks integrated intensity data to guide the geometric ICP: Gueziri et al. [61] combined gradient alignment with intensity registration, while both Bröbner et al. [4] and Chan et al. [74] built hybrid intensity-based pipelines. Furthermore, Naik et al. [76] paired ICP with digitally reconstructed radiographs, iteratively aligning a 3D CT model against synthetic intraoperative projections. Finally, Bröbner et al. [4] combined DL and classical mechanics by integrating a point registration network with a robust point matching (RPM) algorithm.

3.2.5 Tracking

To establish the dynamic spatial relationship between the patient, the imaging devices, and the surgical environment, spatial tracking is essential. Optical tracking emerged as the undisputed gold standard, explicitly adopted by the vast majority of the literature (24 studies) [4; 48; 52–54; 57–59; 61; 64; 66; 68; 70–72; 74; 77; 79; 80; 82–84]. This technique relies on a system of infrared cameras to locate and track

visible markers attached to surgical instruments or patient anatomy with sub-millimetre precision. Within this category, the NDI Polaris family was the most prevalent technological choice. However, several studies achieved comparable spatial localisation using alternative high-end optical trackers, such as the Atracsys FusionTrack 500 [4; 61; 63] and the OptiTrack Prime 13W [68; 74]. Other optical systems, including the Visualeyex VZ4000v and Optotrak Certus, were also noted in earlier studies [53; 55].

Conversely, 12 studies did not apply real-time hardware tracking, choosing instead to focus exclusively on the computational and algorithmic aspects of registration in an offline environment [49–51; 56; 60; 62; 65; 67; 69; 75; 78; 81]. Finally, a single study (Abe et al. [73]) did not disclose its spatial tracking methodology.

3.2.6 Quantitative Outcome

Quantitative evaluation is critical for assessing the clinical viability and technical performance of any registration methodology. By analysing the metrics reported across the literature (summarised in Table 3.1), the evaluation of these systems can be broadly categorised into spatial accuracy, robustness, and computational efficiency.

The primary objective of registration is spatial accuracy, most commonly evaluated through explicit error measurements. Rotational and translational errors were the most frequently reported metrics [4; 55; 57; 63; 64; 68–71; 75; 77; 83], with minimum value of 0.1° and 0.3 mm reported in Min et al. [71]. Rotational error quantifies the angular discrepancy between the coordinate axes, whereas translational error measures the absolute linear distance between corresponding target points. Clinically focused studies heavily relied on target registration error (TRE), which measures the exact distance between a clinically relevant anatomical target and its computed position post-registration [52; 54; 56; 59; 61; 62; 65; 66; 71; 78; 80; 82], the minimum value of 0.3 mm was also obtained by Min et al. [71], and the maximum of 9.65 mm in Shrestha et al. [78]. Similarly, fiducial registration error (FRE) was utilised to calculate the alignment error based specifically on tracked fiducial markers [59; 77; 80; 82], with the lowest value of 0.3 mm [59]. To evaluate overall surface alignment, several studies employed the root mean square error (RMSE) [48; 49; 53; 79; 81; 83] with the lowest value of 0.29 mm [48], alongside general registration errors [56; 60; 72; 84] and specific metrics like positional deviation error (PDE) [62].

Beyond simple point-to-point distance, evaluating the stability of the algorithm is essential. Overall reliability was frequently measured using the registration success rate (SR), which defines the percentage of trials where the algorithm successfully converged within an acceptable clinical threshold [61; 65; 66; 68], the maximum value obtained was 99.7% [66]. Other specialised metrics were used to quantify specific

aspects of reliability, such as surface error and hit count ratio [48; 51], unknown point localisation error (UPLE) [53], good feature ratio (GFR) [78], recall [83], and percentage of landmark correspondence (PLC) [81].

Given the intraoperative nature of surgical navigation, temporal performance is a fundamental constraint. Execution time was rigorously evaluated across numerous studies to determine if the proposed algorithms could provide real-time or near-real-time feedback without disrupting the surgical workflow [54; 55; 58; 59; 61; 62; 65; 67; 74; 76–78; 80]. The highest computational efficiency was reported by Shrestha et al. [78], achieving a registration time of only 1.75 s.

3.3 Discussion

This chapter presents a critical analysis of bone registration methodologies currently established in the literature. The analysis synthesises trends across anatomical target areas, preoperative and intraoperative imaging modalities, hardware, registration algorithms, spatial tracking technologies, and their corresponding quantitative metrics.

Spinal procedures heavily dominate the CAOS literature due to the severe risks of neurological injury and the inherent complexity of the interventions, which require high precision and a comprehensive field of view. The femur and pelvis have also been extensively studied, primarily driven by the demands of pelvic screw fixation and hip arthroplasties. In contrast, fewer studies have addressed the tibia, scaphoid, humerus, and specifically the knee joint. Given the rising clinical demand for complex knee joint interventions, ranging from ligament reconstructions to total arthroplasties, this indicates a significant gap in the literature where CAOS systems remain underdeveloped. Consequently, there is a clear imperative to develop registration frameworks that can adapt to the specific morphological challenges of the knee joint without increasing surgical invasiveness.

Regarding preoperative data acquisition, volumetric imaging remains the standard. CT is explicitly favoured as it provides exceptional contrast for cortical bone and well-defined Hounsfield units, significantly facilitating 3D model reconstruction. The adoption of MRI for preoperative 3D models is restricted not only by operational factors (cost and time) but by the complexity of its bone segmentation. Due to the absence of a standardised intensity scale, such as the Hounsfield units used in CT, MRI signals are relative and fluctuate significantly across different protocols and scanners. This inherent instability hinders the automation of 3D reconstruction. This reliance on CT-based workflows, however, maintains a dependency on ionising radiation, suggesting that the next logical evolution in CAOS must focus on maximising the

utility of safer, intraoperative data streams.

In the intraoperative stage, US emerges as the modality with the greatest potential for surgical navigation. Its increasing adoption is driven by its radiation-free nature, portability, cost-effectiveness, and real-time 2D imaging capabilities. By eliminating the need for bone-implanted fiducial markers, US provides a truly non-invasive alternative that is highly appealing to both surgeons and patients. The reviewed studies demonstrate its clinical feasibility, achieving acceptable error margins (e.g., TRE ranging from 1.27 mm [61] to 9.7 mm [78]) and SR between 72.4% [61] and 99.7% [66]. Conversely, the reliance on intraoperative fluoroscopy and X-ray introduces the significant drawback of cumulative ionising radiation exposure. Additionally, the use of tracked physical pointers to collect points, while avoiding intraoperative imaging altogether, severely limits the field of view, increases operator dependency, and typically requires invasive surgical exposure of the bone surface. Therefore, the primary challenge lies in overcoming the signal-to-noise ratio limitations of US through superior algorithmic processing to compare with invasive methods.

For spatial tracking, while imaging hardware varies widely, optical tracking systems, specifically the NDI Polaris family, remain the undisputed predominant choice. Optical tracking is preferred due to its capacity to provide precise, consistent, and contactless spatial localisation. The literature reflects the system's capacity for high precision, with translational errors reaching as low as 0.3 mm [4] and rotational error recording 0.1° [71]. However, accuracy can vary significantly depending on the experimental setup and environmental factors, with upper limits reported at 5.6 mm [83] for translation and 7.28° [83] for rotation.

In terms of algorithmic approaches, the vast majority of studies employ rigid registration, as bones are non-deformable structures. The ICP algorithm remains the classical, most widely used foundational method, achieving TRE values as low as 1.5 mm [61]. However, it presents notable disadvantages: it is highly sensitive to outliers, requires significant computational iterations, and is prone to converging at local minima if the initial alignment is poor. To address these limitations, researchers are increasingly adopting combined methodologies that pair ICP with robust coarse-alignment techniques.

Probabilistic methods have demonstrated exceptional flexibility and robustness in challenging scenarios. By framing alignment as a statistical probability, these methods achieve high precision (e.g., translational error < 1 mm and rotational RMS $< 0.05^\circ$ [64]). However, this robustness often comes at the cost of requiring numerous iterations to achieve convergence, which can impact real-time intraoperative performance.

Recently, learning-based methods have shown promising, although highly variable, outcomes. The

reported accuracies (e.g., translational errors of $[2.64 \pm 1.73, 5.62 \pm 2.66]$ mm, rotational errors of $[2.78 \pm 1.65, 7.28 \pm 3.55]^\circ$ and RMSE of $[3.10 \pm 1.71, 6.95 \pm 3.86]$ mm) [83] fluctuate significantly depending on the complexity of the neural network architectures and the quality of the training datasets. Despite the inherently high computational cost of model training, inference times are becoming highly competitive. Notably, Shrestha et al. [78] achieved a remarkable runtime of 1.75 s for the registration, demonstrating the immense potential of DL for real-time surgical applications. Nevertheless, a major persisting drawback in the current learning-based literature is the heavy reliance on massive annotated datasets, which are exceptionally scarce. Therefore these DL models often lack generalisation across different US transducers and patient-specific anatomies, limiting their widespread clinical translation.

Ultimately, the literature demonstrates a continuous evolution in bone registration methodologies. While traditional and probabilistic algorithms have established a high benchmark for accuracy, often achieving sub-millimetre and sub-degree precision, the future of CAOS systems relies on balancing this precision with computational efficiency and clinical practicality. Overcoming these intersecting limitations remains an open challenge. Moving forward, the integration of non-invasive US imaging with knee registration algorithms represents the most promising frontier for next-generation surgical navigation.

3.4 Conclusion

This chapter presented a comprehensive review and critical analysis of existing bone registration systems for CAOS systems. By systematically evaluating the technological methodologies that drive modern surgical navigation, this review highlighted both the current state-of-the-art and the prevailing limitations within the field.

A primary conclusion drawn from this analysis is the inherent trade-off present in traditional CAOS systems. While high spatial accuracy, often reaching sub-millimetre and sub-degree values, is routinely achievable, it traditionally relies heavily on ionising radiation (CT, fluoroscopy) or invasive physical tracking (implanted markers or digitising pointers). Consequently, the literature points toward a necessary paradigm shift. Non-invasive, radiation-free modalities, particularly US, are no longer merely experimental alternatives, they represent the future standard of care. To overcome the inherent challenges of US imaging, such as acoustic shadowing and speckle noise, the field is decisively moving away from standalone classical algorithms, like standard ICP, toward more resilient hybrid and learning-based registration frameworks.

Furthermore, the systematic mapping of targeted anatomies revealed a clear discrepancy in research

focus. While spinal and pelvic procedures benefit from highly mature navigation systems, the knee joint remains notably underrepresented. Given the biomechanical complexity of the knee, this represents a major unresolved clinical challenge.

Ultimately, the findings from this review highlight a clear trajectory for the evolution of surgical navigation. Addressing the identified gaps requires the development of fully non-invasive, US-based registration pipelines paired with reliable optical tracking. By integrating advanced computational algorithms capable of overcoming imaging artefacts and achieving high spatial accuracy, future CAOS systems can successfully expand into complex articulations like the knee, meeting both the safety and precision demands of modern orthopaedics.

Chapter 4

Solution description

Establishing a well-structured architectural strategy is crucial prior to the technical development of any surgical navigation pipeline. This chapter outlines the proposed methodology, details the overarching workflow, and specifies the underlying hardware and software infrastructure. Fundamentally, it aims to answer the following **RQ**: *What specific clinical and technical requirements must be met to develop a viable intraoperative US-based registration system?*

4.1 System Requirements

Before initiating the implementation phase, a strict set of system requirements was defined. These prerequisites dictate what the proposed system must deliver to ensure it is safe, highly accurate, and practically deployable within a fast-paced surgical environment.

Clinical Requirements

From a clinical standpoint, spatial accuracy is paramount. In accordance with established literature, such as Gueziri et al. [61], a viable US to CT bone registration system must achieve a TRE of approximately 1.5 mm to guarantee safe surgical navigation. Exceeding this error margin severely compromises critical tasks, such as precise screw placement or accurate bone resection. Furthermore, the proposed method must be entirely non-invasive, definitively eliminating the need for bone drilling to insert fiducial markers, while also minimising reliance on intraoperative ionising radiation. Finally, the system must not disrupt the surgical workflow. According to Sciberras et al. [85], a clinically acceptable intraoperative registration workflow should ideally be completed in under 3 min and 20 s.

Technical Requirements

From an engineering perspective, achieving the aforementioned clinical goals relies on several technical prerequisites. First, rigorous spatial and temporal calibration is mandatory. The hardware infrastructure must flawlessly synchronise the continuously captured 2D US frames with the 3D coordinate data provided

by the spatial tracking system in real-time. Second, a highly robust and automated method for bone surface extraction is required. Without accurately isolating the bone contour from the noisy US images, the subsequent surface registration will inevitably fail. Lastly, the software architecture must be capable of establishing seamless, low-latency communication via standard protocols to integrate the imaging output, the tracking data, and the computational processing nodes.

4.2 Solution Overview

To successfully meet the established requirements, the proposed solution was designed as a modular, multi-phase pipeline. The data flow and architectural workflow are visually summarised in Figure 4.1.

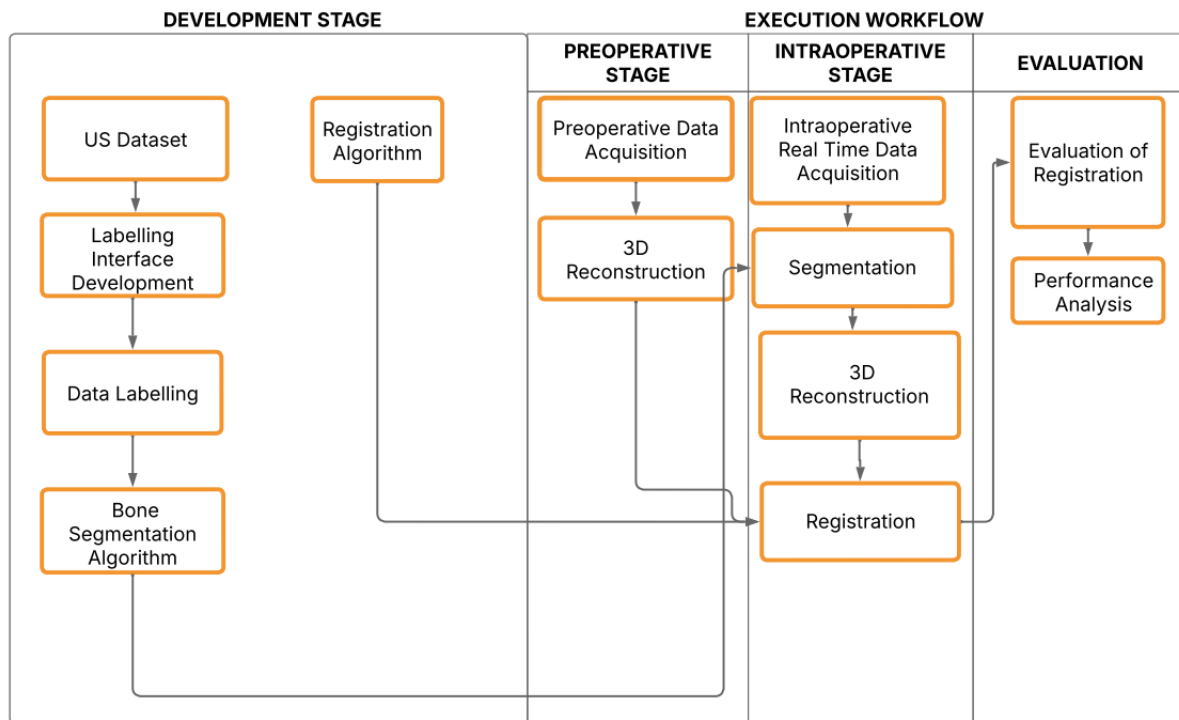


Figure 4.1: Overall proposed solution.

The proposed architectural pipeline is systematically structured into four interdependent phases:

- **Development Stage:** This initial stage establishes the core algorithmic foundation of the system. To satisfy the technical requirement for automated bone surface extraction, the architecture necessitates the acquisition and meticulous annotation of a custom US dataset via a dedicated labelling interface. The resulting ground truth data serves to train and validate a DL segmentation model. Concurrently, the mathematical framework governing the rigid registration algorithm is parameterised and configured.

- **Preoperative Stage:** This stage dictates the acquisition of preoperative volumetric data, specifically standard CT scans. The acquired data is subjected to 3D reconstruction algorithms to generate a high-fidelity anatomical model of the target bone. This volumetric model subsequently serves as the definitive static ground truth reference for the navigation coordinate system.
- **Intraoperative Stage:** Representing the active surgical execution, this phase initiates with the acquisition of real-time, spatially tracked US imaging. The captured 2D B-mode images are processed by the previously trained DL segmentation model to output binary masks defining the bone profile. By integrating these segmented masks with their corresponding spatial tracking data, a dynamic 3D reconstruction is performed to yield an intraoperative point cloud. Ultimately, the registration algorithm aligns this US-derived point cloud against the preoperative CT model, thereby establishing the continuous spatial guidance required for surgical navigation.
- **Evaluation Stage:** To ascertain the clinical viability and technical robustness of the proposed architecture, the pipeline’s final outputs are subjected to rigorous quantitative testing. This phase systematically evaluates the spatial accuracy of the anatomical alignment, measured via TRE, alongside the overall computational latency, thereby verifying compliance with the stringent operational thresholds defined in Section 4.1.

4.3 Hardware and Software Infrastructure

To physically construct and execute the architecture described above, the computational load and data acquisition were strategically partitioned across specialised hardware and software stacks.

4.3.1 Hardware

Depending on the specific phase of the workflow, distinct hardware setups were employed to optimise data integrity and processing speed. Regarding imaging acquisition, the US dataset required for training the segmentation algorithms was acquired using a clinical-grade Canon Aplio i800 system, equipped with a high-frequency PLT 1005BT linear probe (Canon Medical Systems), as presented in Figure 4.2. For the real-time intraoperative phases, images were captured using a GE LOGIQ P9 system paired with an L3-12 linear transducer (GE Healthcare), which is highly suitable for musculoskeletal imaging. To ensure real-time processing, the video output was streamed to the local workstation via a UGREEN frame grabber at a resolution of 720x480 pixels and a frame rate of 60 frames per second (Figure 4.3).



Figure 4.2: Equipment used for dataset acquisition: (a) Canon Aplio i800 system; and (b) PLT 1005BT linear probe.

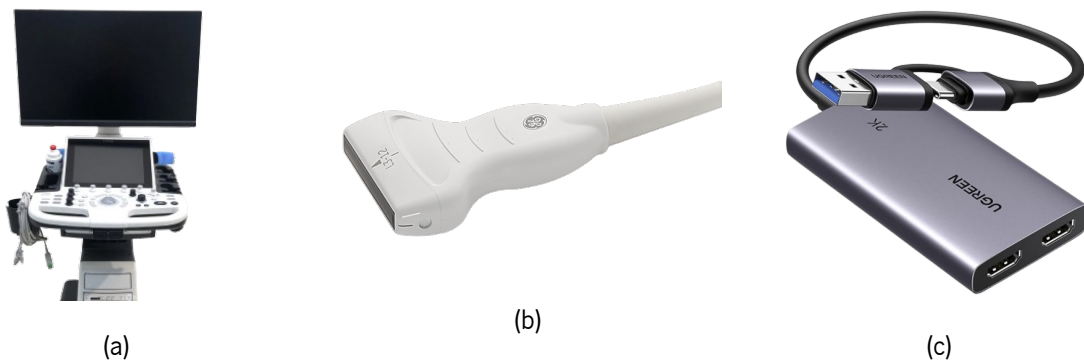


Figure 4.3: Equipment used for real-time data capture: (a) GE LOGIQ P9 system; (b) L3-12 linear probe; and (c) UGREEN frame grabber.

For spatial tracking and dynamic pose data acquisition, the OptiTrack V120 Trio (NaturalPoint Inc.) was employed (Figure 4.4). This industry-standard device measures the 3D position of tracked tools with an accuracy of 0.5 mm RMS, an acquisition rate of 120 Hz and latency of 8.33 ms. Spatial localisation is achieved through the detection of passive, spherical infrared markers attached to custom-designed reference rigid bodies. In consequence, references with specific configurations were designed for this purpose.



Figure 4.4: Optitrack V120 Trio tracking system.

Finally, due to the vast disparity in processing requirements, the computational infrastructure was divided between a local workstation and a high-performance computing system. The highly intensive segmentation training phase was executed on the Deucalion supercomputer, utilising a compute node equipped with two AMD EPYC 7742 64-core processors (2.25 GHz), 512 GB DDR4 RAM, and four NVIDIA A100 Graphics Processing Unit (GPU) (80 GB VRAM each). Conversely, the real-time intra-operative pipeline and registration algorithms were executed on a local workstation containing an Intel Core i7-10750H processor, 16 GB of RAM, and an NVIDIA GeForce GTX 1650 Ti GPU.

4.3.2 Software

This subsection details the specific software stack driving the hardware infrastructure. For the computer-aided design (CAD) of custom physical components, such as the rigid bodies for optical tracking, SOLIDWORKS (Dassault Systèmes) was employed. Medical image visualisation, manual annotation, and 3D reconstruction were conducted within 3D Slicer [86] (The Slicer Community, version 5.6.2).

The core computational pipeline was implemented in Python (version 3.11.4). The neural network architectures and training routines were built using the PyTorch framework (version 2.7.1). Crucially, this framework was extended with the Medical Open Network for AI (MONAI, version 1.5.2) [87], which provided specialised transformations and loss functions tailored for medical imaging. To leverage GPU acceleration during model training, the environment was configured with the CUDA toolkit (version 11.8). Finally, hardware communication and spatial calibration were managed using the standardised algorithms provided by the Public Software Library for Ultrasound (PLUS) Toolkit [88], ensuring robust integration between the US feed and the optical tracker.

4.4 Conclusion

This chapter established the comprehensive architectural framework and system requirements for the proposed non-invasive, multimodal (US to CT) rigid registration pipeline. From the initial definition of clinical constraints to the final evaluation strategy, every component of the system was purposefully structured. Fundamentally, the architecture is explicitly designed to meet the strict clinical imperative of achieving a TRE of approximately 1.5 mm, while guaranteeing a highly efficient intraoperative workflow completely devoid of invasive fiducial markers.

To satisfy these rigorous demands, the underlying hardware and software infrastructure was strategically partitioned. High-fidelity US imaging and sub-millimetre optical tracking provide the requisite sensory input, while the computational load is systematically distributed. Specifically, High-Performance Computing resources are allocated for the intensive training of DL models, whereas a dedicated local workstation governs the real-time intraoperative tasks.

Furthermore, the chapter delineated the overarching workflow, systematically dividing it into four inter-dependent phases. A preparatory phase establishes the algorithmic core for segmentation and registration, seamlessly integrating with the preoperative generation of 3D anatomical models. Subsequently, the intraoperative phase executes dynamic, real-time US segmentation and point cloud registration, culminating in a rigorous evaluation phase to validate the system's clinical feasibility.

With the conceptual architecture, operational constraints, and technological stack now clearly defined, the theoretical blueprint of the solution is complete. This establishes a robust and well-defined foundation for the practical development phase of this dissertation, detailed in the following chapters, covering both the technical implementation and algorithmic execution of the proposed navigation system.

Chapter 5

Deep Learning-Based Pipeline for Bone Surface Extraction in Knee Ultrasound

This chapter details the methodological framework developed to automatically extract the bone surface from raw, 2D US data. Transitioning from the theoretical architecture established in the previous chapter, this section comprehensively describes the practical segmentation engineering pipeline. The workflow is systematically presented, beginning with the clinical acquisition and curation of a custom dataset, followed by the development of a dedicated annotation interface for ground truth generation. Subsequently, the chapter explores the implementation, training, and evaluation of diverse DL architectures. Ultimately, by analysing the spatial accuracy and inference latency of the resulting segmentation masks, this chapter aims to address the following **RQ**: *Which DL algorithm is best suited for extracting bone contours from US images, to reliably reconstruct point clouds for registration?*

5.1 Clinical Dataset Acquisition

Achieving highly accurate automated bone surface extraction depends fundamentally on the quality of the underlying training data. Consequently, the systematic construction of a robust and representative dataset was an essential prerequisite. This section details the initial exploratory approaches for data sourcing, culminating in the establishment of a rigorous clinical acquisition protocol and the subsequent curation of a custom US dataset.

5.1.1 Dataset Sourcing and Clinical Acquisition Protocol

An initial investigation was conducted to identify publicly accessible datasets relevant to US bone segmentation of the knee joint. Datasets such as ‘Ultrasound Cases’ [89], the ‘AI/ML-ready knee ultrasound dataset in a population-based cohort’ [90], and the ‘Open-Access Dataset for the Low-limb-Leg multilabel

dataset' [91] were evaluated. However, despite their public availability, these datasets presented substantial limitations for this specific application. Issues such as restricted anatomical representativeness (failing to cover both healthy and pathological conditions), low spatial resolution, severe compression artefacts, and a lack of multi-planar views severely compromised their utility for training a robust surgical navigation model.

To mitigate the logistical resources required for clinical data collection, simulated US generation was initially explored. Utilising 3D Slicer and the PLUS Toolkit, synthetic US images were generated from a preoperative CT scan. Following threshold-based bone segmentation, simulation parameters such as transducer type, acoustic noise, frame rate, and tissue acoustic properties were configured to perform a volumetric sweep. However, while the resulting synthetic images exhibited high structural fidelity and were free of unexpected artefacts, they severely lacked the realistic acoustic noise and characteristic speckle typical of authentic US imaging, as illustrated in Figure 5.1. Consistent with previous literature, such as Patel et al. [92], it was concluded that simulated US data is generally ineffective for training DL models intended for accurate, real-world clinical segmentation.

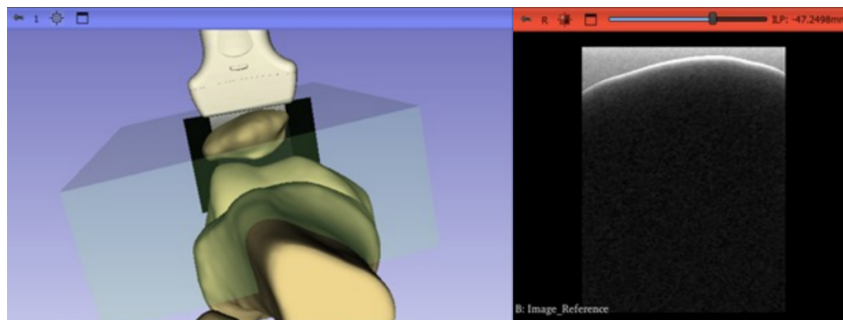


Figure 5.1: Simulated US scanning within the 3D Slicer, generating a B-mode image (right) from a specific probe position relative to the 3D bone volume extracted from CT (left).

Consequently, a custom, prospective dataset was acquired. Axial and sagittal knee US scans were collected at a single clinical centre within the Lusíadas Saúde healthcare group. To ensure protocol consistency, all acquisitions were performed by a single, highly experienced musculoskeletal radiologist between June and October 2025. All data were fully anonymised prior to analysis. The cohort consisted of 7 subjects, encompassing both healthy individuals and patients with confirmed pathological knee conditions. Inclusion criteria required participants to possess confirmed skeletal maturity. Exclusion criteria included pregnancy, recent traumatic injuries to the knee or adjacent soft tissues, and dermatological conditions precluding the use of standard US conductive gel.

All imaging sequences were recorded in dynamic 2D B-Mode, utilising a spatial resolution of 960

$\times 1280$ pixels, and exported in DICOM format. The acquisitions strictly adhered to a standardised protocol specifically designed to capture the entire osseous topography of the knee joint. Each participant contributed scans from one knee, comprising 10 distinct sweeps per subject, yielding a comprehensive dataset of 15,480 individual US frames.

The acquisition protocol initiated with the subject in a supine position with the knee in full extension. In this configuration, six transverse (short-axis) sweeps were collected by positioning the linear probe perpendicular to the bone and sliding it from a proximal to a distal position. Specifically, three transverse sweeps were performed over the femur in the lateral (1), medial (2), and anterior (3) regions, and three over the tibia in the lateral (4), medial (5), and anterior (6) regions. Maintaining the extended supine position, the probe was rotated 90° to obtain longitudinal (long-axis) scans, in this orientation, three targeted sweeps were performed over the femur (7), the tibia (8), and the tibial plateau (9). Subsequently, the subject was instructed to flex the knee to 90° while remaining supine to allow for a final longitudinal scan of the femoral condyles (10). The complete 10-step scanning protocol is illustrated in Figure 5.2.

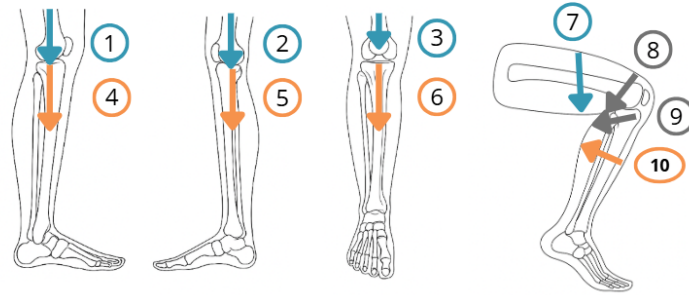


Figure 5.2: Image acquisition protocol: (1–3) transverse sweeps of the femur (lateral, medial, and anterior), (4–6) transverse sweeps of the tibia, (7–9) longitudinal scans with the knee in full extension, and (10) longitudinal scan of the femoral condyles with the knee at 90° flexion.

5.1.2 Cohort Demographics and Dataset Characteristics

Following the established acquisition protocol, a comprehensive descriptive analysis was performed to evaluate the global cohort demographics, data distribution, and inherent image properties.

Baseline demographic and clinical data, including biological sex, knee laterality, age, and body mass index (BMI), were recorded for all 7 subjects. The individual profiles of each participant are detailed in Table 5.1. Overall, the acquired dataset presents a balanced sex distribution and covers a representative range of adult ages and body compositions. Additionally, the overall laterality of the scanned knees demonstrates an approximate symmetry, ensuring that the global dataset is anatomically diverse.

Table 5.1: Demographic and clinical characteristics of the participants

Participant	Sex	Knee laterality	Age (years)	BMI (kg/m ²)
1	Female	Right	24	21.9
2	Female	Left	40	28.7
3	Female	Left	44	32.4
4	Male	Right	26	25.7
5	Male	Right	34	24.3
6	Male	Left	39	23.6
7	Male	Left	21	24.2

Regarding anatomical coverage, the distribution of the 15,480 acquired US frames reflects the morphological characteristics of the knee joint. As shown in Figure 5.3, the data acquisition remained balanced per subject, with a representative average of slices for each anatomical area. Notably, the average number of slices is generally higher for the femur compared to the tibia, a direct correlation with the femur's significantly larger surface area. Across all regions, the sagittal views yielded the highest volume of data, peaking at approximately 353 slices for the sagittal femur and 310 for the sagittal tibia. Within the axial orientation, the anterior sweeps provided the most extensive coverage. Furthermore, robust data acquisition of the critical joint-specific structures was achieved, with the tibial plateau and femoral condyles averaging approximately 275 and 259 slices, respectively.

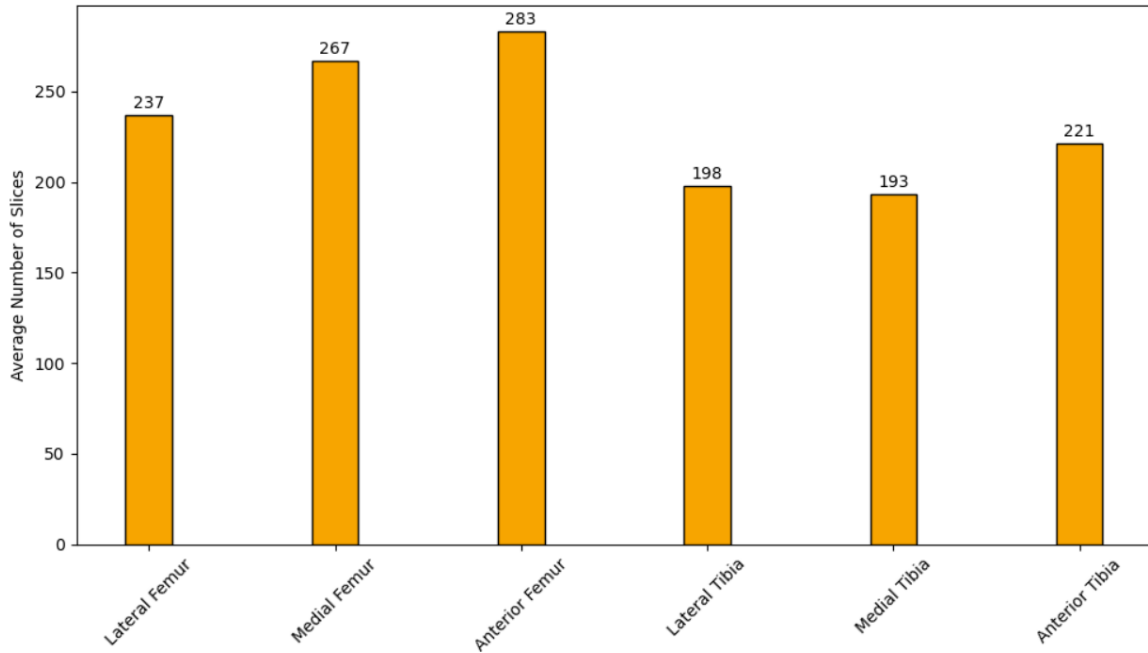


Figure 5.3: Average number of slices for each region.

Analysing the raw image metadata is critical to define the appropriate input parameters for the subsequent DL architectures. Confirming the established acquisition protocol, the dataset exhibits perfect spatial uniformity, with all 15,480 US frames maintaining the native rectangular resolution of 960×1280 pixels. Additionally, as illustrated in Figure 5.4, the pixel spacing values for both the X and Y dimensions share an identical distribution, ranging from 0.008 to 0.012 mm/pixel. This confirms that the 2D image voxels are completely isotropic, ensuring that no directional distortion exists within the anatomical planes.

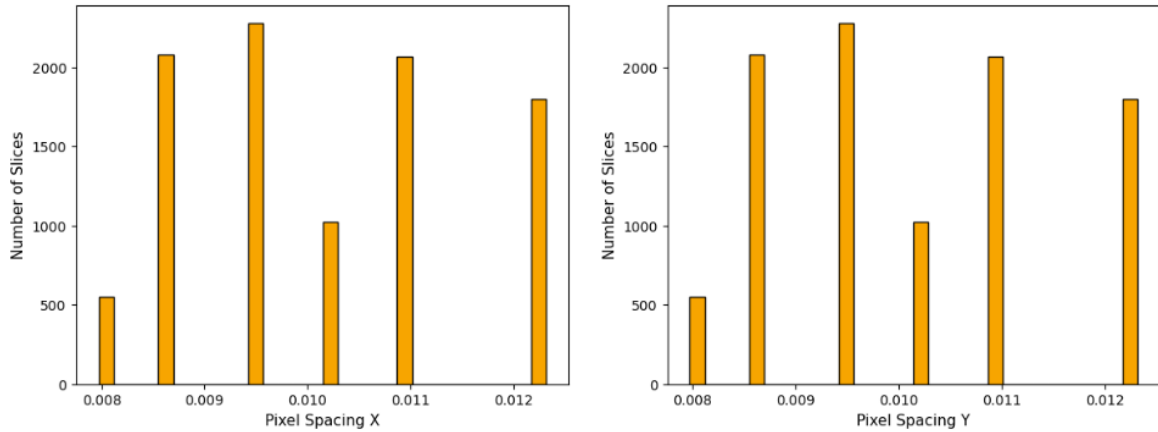


Figure 5.4: Distribution of pixel spacing values for the X (left) and Y (right) dimensions.

Finally, it is crucial to acknowledge an inherent characteristic of freehand 2D US acquisitions. Although the imaging data is exported as continuous multiframe sequences, the physical distance between consecutive frames (Z-axis translation) is naturally variable, as it strictly depends on the manual sweep speed of the operator. Consequently, the DICOM metadata inherently lacks inter-slice spatial relationships, meaning there is no intrinsic volumetric consistency between the 2D images. From an engineering perspective, this physical reality explicitly justifies the architectural requirement of employing an external optical tracking system (as established in Chapter 4). By synchronising the imaging stream with the continuous tracking data, it becomes possible to overcome this modality's inherent constraint and accurately map the segmented 2D frames into a cohesive 3D Point Cloud during the intraoperative phase.

5.2 Ground Truth Generation

Supervised DL architectures require vast amounts of highly accurate, annotated data to converge effectively. However, due to the inherent speckle noise, acoustic artefacts, and low signal-to-noise ratio characteristic of US imaging, fully automated ground truth extraction is unfeasible. Consequently, a rigorous manual annotation phase was imperative. This section details the design of a dedicated annotation

module and the expert-supervised clinical protocol implemented to ensure the anatomical precision of the generated ground truth masks.

5.2.1 Development of a Graphical User Interface for Manual Labelling

Manually annotating 15,480 individual US frames using standard medical imaging software presents significant logistical challenges, primarily regarding workflow inefficiency and data mismanagement. To overcome these bottlenecks, a customised Graphical User Interface (GUI) was engineered directly within 3D Slicer, utilising Qt Designer for the frontend architecture and Python for the backend logic (Figure 5.5). While the foundational pixel-wise drawing mechanics leveraged 3D Slicer's native Segment Editor toolkit, this dedicated module was explicitly designed to orchestrate and optimise large-scale dataset labelling.

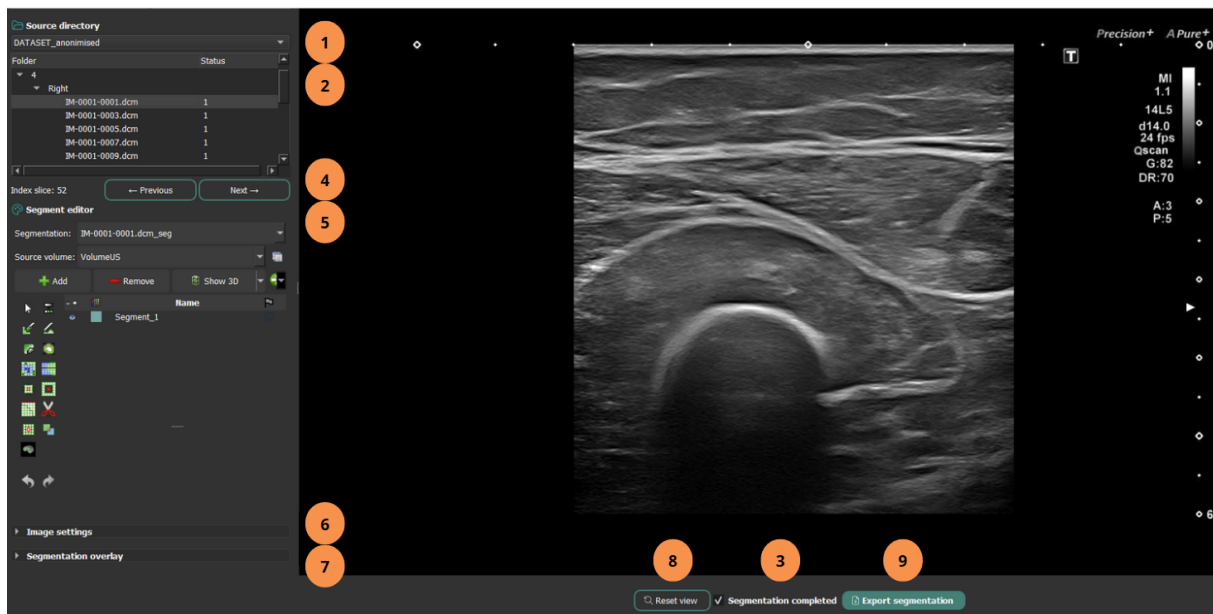


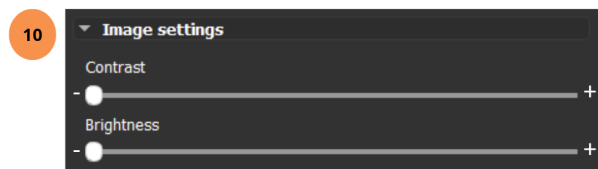
Figure 5.5: Customised 3D Slicer module engineered to optimise the manual annotation workflow for large-scale US datasets.

The software architecture was driven by three primary functional requirements:

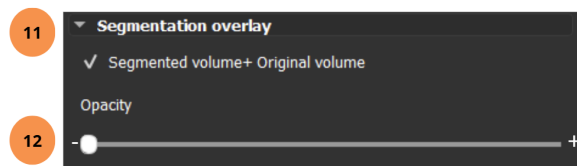
- Workflow and State Management:** To manage thousands of frames, the system incorporates a source directory panel (1) with an integrated QTreeView (2) for dataset navigation. The module dynamically tracks the segmentation status of each volume (0 - Not Started, OS - Ongoing Segmentation, or 1 - Segmentation Completed) and synchronises these states via a JSON file. Once the annotator considers a segmentation definitive, a specific checkbox (3) marks the volume as complete. To facilitate rapid navigation, the interface includes sequential navigation controls, 'Previous' and 'Next' buttons (4), and supports mouse-wheel indexing, allowing for the inspection of

high-volume datasets. Furthermore, the environment provides direct access to the native segment editor (5) for pixel-wise drawing.

- **Visual Enhancement and Real-time Verification:** Identifying precise bone boundaries in noisy US images is a visually demanding task. Consequently, the module integrates an image settings panel (6) containing interactive sliders for the dynamic adjustment of contrast and brightness (10) (Figure 5.6(a)). These non-destructive operations assist in distinguishing anatomical boundaries in hypoechoic or hyperechoic regions. Additionally, a segmentation overlay panel (7) (Figure 5.6(b)) allows the user to toggle the display (11) and adjust the mask's opacity (12) in real-time. Additional features, such as a localized reset view button (8), ensure a standardised visualisation throughout the process.



(a)



(b)

Figure 5.6: Custom module control panels: (a) dynamic image enhancement settings, and (b) segmentation overlay controls.

- **Data Integrity and Export Failsafes:** To eliminate the risk of data loss, two critical protection mechanisms were hardcoded into the backend. A primary export routine (9) standardises the output into Neuroimaging Informatics Technology Initiative (NIfTI) format. If a user attempts to navigate away from an unsaved volume, a forced confirmation dialogue (13) intercepts the action (Figure 5.7(a)). Upon a successful export, the application triggers a system notification (14) confirming the exact file path and metadata storage (Figure 5.7(b)).

Ultimately, this custom module successfully abstracted the complexity of the default 3D Slicer environment. By minimising graphical interaction time and mitigating human error, it established a highly optimised, task-specific pipeline for ground truth generation.

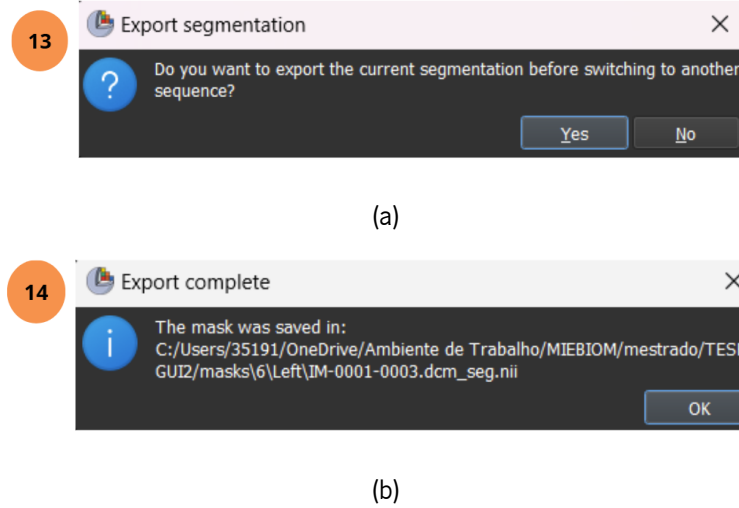


Figure 5.7: Integrated data integrity mechanisms: (a) unsaved progress warning, and (b) export completion verification.

5.2.2 Manual Bone Segmentation and Ground Truth Generation

In B-mode US, the bone interface produces a highly specific acoustic response. The significant impedance mismatch between soft tissue and cortical bone causes massive reflection of the US waves, creating a bright, hyperechoic ridge. Consequently, the acoustic energy fails to penetrate deeper, resulting in a pronounced posterior acoustic shadow [93]. While this shadow is crucial for distinguishing bone from other highly reflective connective tissues, it often obscures the precise anatomical margins, making pixel-perfect boundary delineation highly subjective.

To guarantee clinical accuracy and mitigate observer bias during annotation, a strict, expert-supervised protocol was executed. The initial pixel-wise segmentation was meticulously performed by two trained researchers utilising the custom GUI (Figure 5.8). Crucially, this process was conducted under the direct, real-time supervision of a senior musculoskeletal radiologist. The clinical expert provided continuous directives for resolving ambiguous boundaries, particularly in regions corrupted by reverberation artefacts or signal dropouts.

Upon completion of the initial draft, all 15,480 generated masks underwent a rigorous slice-by-slice review. The final approval and validation of each binary mask were strictly granted by the same expert radiologist. This meticulous validation pipeline ensured that the resulting ground truth dataset reliably and objectively reflected the true osseous anatomy, providing a gold-standard foundation for the subsequent training of the DL models.

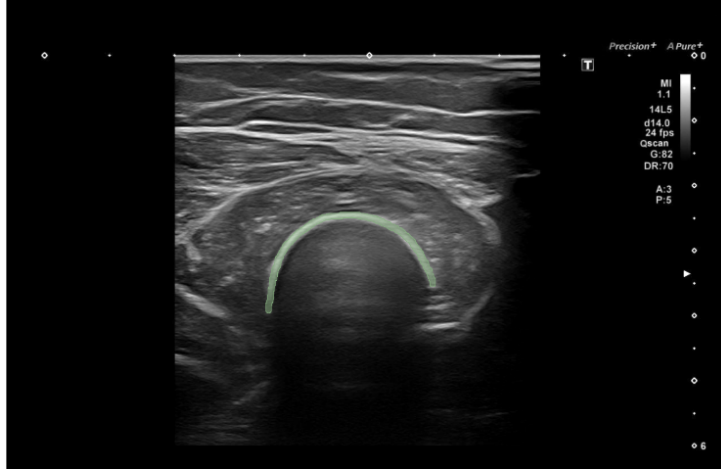


Figure 5.8: Example of an expert-validated ground truth mask (green overlay) accurately delineating the hyperechoic bone surface.

5.3 Deep Learning Framework

Following the acquisition of clinical data and the generation of validated ground truth masks, the subsequent phase focused on the computational extraction of the bone surfaces. This section details the complete DL pipeline, encompassing the data preparation strategies, the selected neural network architectures, the training configurations, and the quantitative evaluation metrics.

5.3.1 Data Curation and Subject-Level Partitioning

To ensure the integrity and robustness of the learning process, a final quality assurance filtering was performed on the raw dataset. This curation step aimed to exclude frames exhibiting severe acoustic degradation, probe detachment, or a complete absence of discernible osseous interfaces, which could introduce noise into the model's convergence. Consequently, 646 non-viable scans were systematically removed, resulting in a refined dataset of 14,834 high-quality US frames and their corresponding expert-validated ground truth masks.

To strictly prevent data leakage and ensure unbiased evaluation, the dataset was partitioned at the subject level. By isolating entire subjects rather than individual frames, we guarantee that the model is evaluated on previously unseen anatomical morphologies. Two subjects (approximately 30% of the data) were reserved exclusively as an independent hold-out test set for final performance reporting. The remaining five subjects (approximately 70%) were allocated for model development, utilising a subject-based 5-fold cross-validation strategy. Specifically, in each of the five folds, the imaging data from four

subjects was utilised to train the model, while the data from the remaining one subject was held out exclusively for validation. Because this rotation was strictly subject-based, the neural network was never exposed to the validation subject’s anatomy during the training phase of any given fold. This approach fundamentally prevents data leakage, maximising the utility of the available data while ensuring a rigorous and realistic assessment of the model’s generalisation capabilities.

Following this subject-level partition, it is crucial to analyse the clinical and anatomical balance across the resulting subsets. The demographic characteristics of the training and testing datasets are summarised in Table 5.2.

Table 5.2: Data distribution of training, validation and test data sets

Variable	Training and validation	Test
No. of images	11081	4399
Age (mean $\pm$ STD)	30.19 $\pm$ 8.12	38.91 $\pm$ 5.00
BMI (mean $\pm$ STD)	24.56 $\pm$ 2.53	28.43 $\pm$ 4.05
Sex (Male / Female, %)	5193 (47) / 5888 (53)	2239 (51) / 2160 (49)
Side (Left / Right, %)	6694 (61) / 4387 (39)	2160 (49) / 2239 (51)

As detailed in the table, the mean age of subjects was 30.19 ± 8.12 years (range: 21–39 years) in the training and validation sets, and 38.91 ± 5.00 years (range: 34–44 years) in the test set, with a corresponding mean BMI of 24.56 ± 2.53 and 28.43 ± 4.05 , respectively. Furthermore, the dataset featured a balanced distribution of male and female subjects. Regarding anatomical laterality, the training and validation sets presented a slight predominance of left knees (61%), while the test set maintained an equal distribution.

5.3.2 Image Preprocessing and Augmentation

Prior to network ingestion, the partitioned data was subjected to a standardised preprocessing pipeline to ensure spatial and intensity consistency across the disparate clinical scans.

Resampling and Resizing

To maintain the true physical scale of the anatomical structures, all images and masks were resampled to an isotropic target pixel spacing of 0.02 mm/pixel. The new image sizes were determined according to Equation 5.1:

$$\frac{\text{Original size}}{\text{New size}} = \frac{\text{New spacing}}{\text{Original spacing}} \quad (5.1)$$

For the US data, B-spline interpolation was employed to preserve the smooth intensity gradients characteristic of acoustic imaging. Conversely, nearest-neighbour interpolation was strictly applied to the ground truth masks to maintain their binary nature. Following resampling, a uniform input resolution was enforced through centered cropping or zero-padding, resulting in a fixed tensor resolution of 512×512 pixels. As illustrated in Figure 5.9, this process ensures that the anatomical proportions remain consistent regardless of the original scanner configuration.

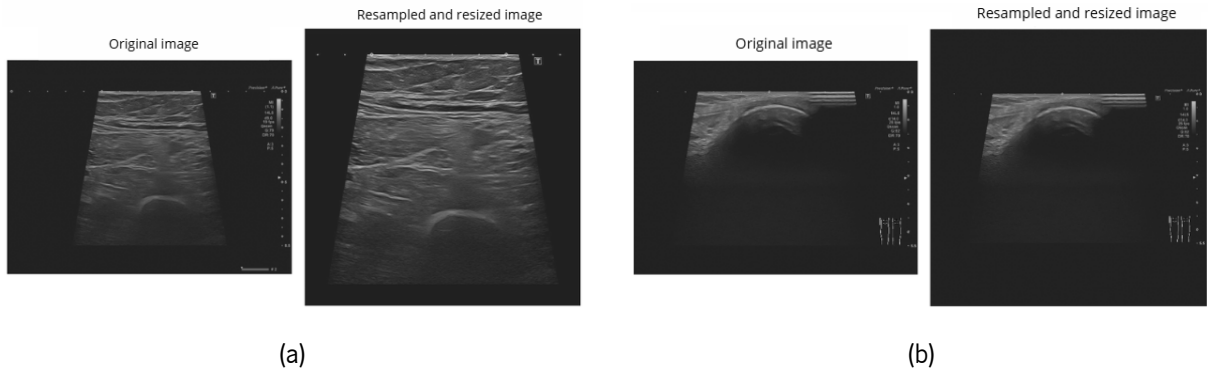


Figure 5.9: Resampling and resizing outcomes: (a) normalisation of a larger initial pixel spacing (0.012 mm), and (b) normalisation of a smaller initial pixel spacing (0.008 mm).

Crucially, the complete set of transformation metadata, including the original and target dimensions, pixel spacings, and applied interpolation methods, was persistently stored for each sequence. This data retention is a fundamental architectural requirement, as it enables the exact inversion of these spatial transformations, allowing the resulting 2D segmentation masks to be accurately mapped back into the 3D physical coordinate system during the intraoperative reconstruction phase.

Greyscale Conversion

Although the US scans were originally exported in an RGB format, B-mode US inherently relies on acoustic echo intensity rather than colour. Consequently, the RGB channels introduce unnecessary computational redundancy without providing additional clinical information. Therefore, all images were first natively converted to a single-channel greyscale format.

Intensity Normalisation

Subsequently, to ensure consistent intensity distributions across different patients and varying scanner gains, Z-score normalisation was applied. This method was selected because it standardizes the pixel intensities of each slice to a zero mean and a unit variance ($\mu = 0, \sigma = 1$). Importantly, since this is a linear transformation, it accelerates network convergence while strictly preserving the underlying anatomical contrast and visual integrity. The pixel-wise transformation is defined by Equation 5.2:

$$I_{norm}(x, y) = \frac{I(x, y) - \mu}{\sigma} \quad (5.2)$$

where $I_{norm}(x, y)$ represents the normalised pixel intensity, of the image at pixel position (x, y) , and μ and σ denote the global mean and standard deviation of the image, respectively.

Data Augmentation

DL architectures are inherently prone to overfitting, particularly in medical imaging where target structures are sparse. To enhance the models' generalisation to unseen clinical conditions, a dynamic data augmentation strategy was implemented during the training loop. As illustrated in Figure 5.10, the applied augmentations were carefully selected to mimic real-world US acquisition variability. Augmentations included random spatial scaling ($\pm 20\%$) to simulate depth variations, intensity modulation of brightness and contrast ($\pm 15\%$) to replicate disparate gain settings, and multiplicative speckle noise injection (sampled from $[0.9, 1.1]$) to model the stochastic nature of US wave propagation.

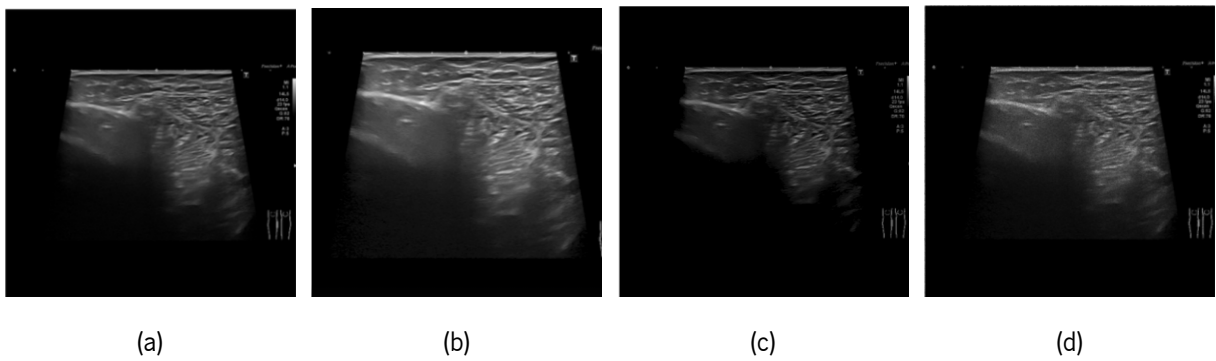


Figure 5.10: Examples of the dynamic data augmentation techniques applied: (a) original slice, (b) scaled zoom, (c) variation in brightness and contrast levels, and (d) synthetic speckle noise injection.

5.3.3 Network Architectures

Traditional image segmentation methods rely heavily on manual feature engineering and are frequently inadequate for addressing the inherent acoustic shadowing, speckle noise, and anatomical variability of medical US. Consequently, DL has emerged as the standard approach for osseous segmentation [94–101]. To rigorously evaluate the optimal solution for real-time surgical navigation, four distinct DL architectures were implemented and compared: a standard U-Net, a customised U-Net with a VGG16 backbone, an auto-configuring framework (nnU-Net), and an attention-based Vision Transformer (SegFormer).

Standard U-Net and Customised U-Net with VGG16

The standard U-Net [102] was selected as the primary baseline due to its established efficacy in medical image segmentation. Its symmetric encoder-decoder structure, fortified by skip connections, effectively propagates spatial information across network depths, a critical requirement for recovering fine osseous boundaries obscured by acoustic noise [103].

Building upon this baseline, a customised U-Net variant was implemented to leverage the advantages of transfer learning. As illustrated in Figure 5.11, the standard contraction path was replaced with a VGG16 encoder [104], pre-trained on the large-scale ImageNet database [105]. Given the relatively constrained size of the clinical US dataset, initialising the network with robust, pre-learned feature weights significantly accelerates training convergence and mitigates the risk of overfitting.

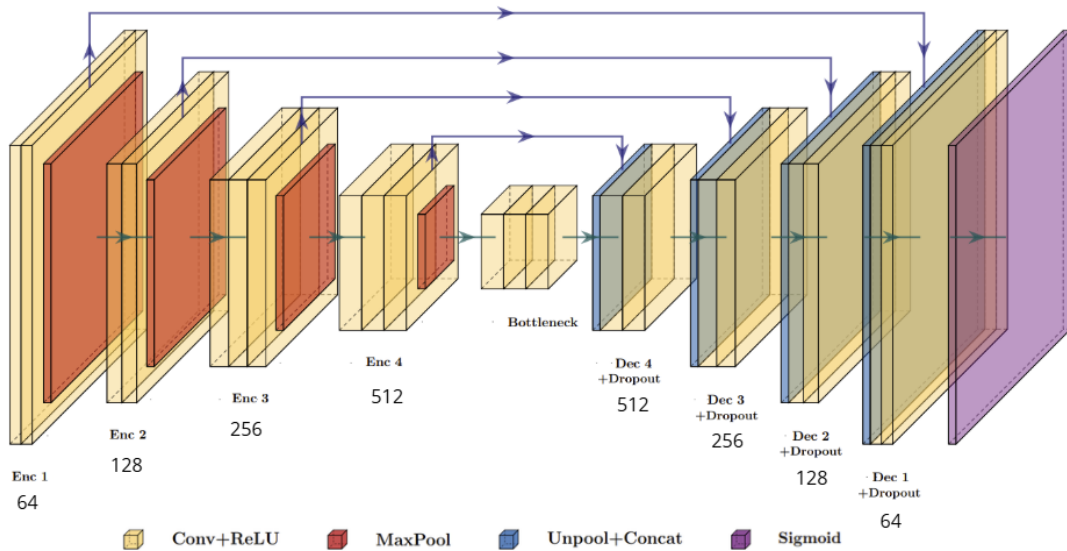


Figure 5.11: Network architecture of the customised U-Net featuring a pre-trained VGG16 encoder backbone.

Architecturally, the VGG16 encoder extracts hierarchical features through successive blocks of 3×3 convolutions, followed by Rectified Linear Unit (ReLU) activations and Batch Normalisation. Spatial down-sampling is achieved via 2×2 Max Pooling operations. Following the bottleneck, the expansion path recovers the original spatial resolution through a sequence of 2×2 transposed convolutions. The definitive feature of this topology lies in its skip connections, which concatenate the high-resolution, low-level feature maps from the pre-trained VGG16 encoder directly into the decoder. This fusion is essential to compensate for the spatial degradation inherent to the pooling layers. Furthermore, to enhance generalisation, a Dropout rate of 0.5 was integrated into the expansion blocks. The pipeline culminates in a 1×1 convolutional layer with a Sigmoid activation function, outputting a pixel-wise probability map that distinguishes the bone structure from the background.

Self-Configuring and Transformer-Based Benchmarks

To establish a state-of-the-art benchmark, the 2D full-resolution configuration of the nnU-Net framework [106] was implemented. Unlike standard static architectures, nnU-Net utilises a self-configuring, data-driven pipeline that dynamically adapts preprocessing protocols, network topology, and training hyperparameters directly to the dataset’s fingerprint, thereby eliminating empirical manual tuning.

Finally, to evaluate the efficacy of self-attention mechanisms against traditional CNNs, SegFormer [107] was evaluated. CNNs typically struggle to capture long-range contextual dependencies due to their localized receptive fields. SegFormer addresses this by employing a multi-level hierarchical Transformer encoder paired with a lightweight multi-layer perceptron decoder. This design efficiently aggregates both local textural features (e.g., bone surface hyperechogenicity) and broad global anatomical context (e.g., the posterior acoustic shadow), presenting a theoretically powerful alternative for overcoming complex US artefacts.

5.3.4 Training Configuration and Loss Functions

In US imaging, the osseous interface represents a minimal fraction of the total pixel area, leading to a severe foreground-background class imbalance. To effectively address this, a combined loss function (L_{total}) was designed, integrating a weighted Binary Cross-Entropy (L_{wBCE}) with a Dice loss (L_{Dice}):

$$L_{total} = L_{wBCE} + L_{Dice} \quad (5.3)$$

To heavily penalise false negatives (FN) and force the network to prioritize the sparse bone contours, a foreground weight of $w = 3$ was applied within the L_{wBCE} formulation:

$$L_{wBCE} = -\frac{1}{N} \sum_{i=1}^N [3 \cdot y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)], \quad (5.4)$$

where N is the total number of pixels, $y_i \in \{0, 1\}$ represents the ground truth label, and $\hat{y}_i \in [0, 1]$ is the predicted probability for pixel i . Concurrently, the Dice loss component was employed to directly optimise spatial overlap:

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N y_i \hat{y}_i + \epsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i + \epsilon}, \quad (5.5)$$

with ϵ acting as a small smoothing constant to prevent division by zero and ensure numerical stability.

Regarding hyperparameter configuration, the standard U-Net, U-Net with VGG16, and SegFormer models were optimised using the Adam algorithm. Training was initiated with a learning rate of 1×10^{-4} , coupled with a cosine annealing scheduler that gradually decayed the learning rate to a minimum of 1×10^{-6} to ensure smooth local convergence. To maximise computational throughput, a batch size of 64 was utilised for the CNN-based U-Net variants. However, to accommodate the intensive GPU memory constraints of the Transformer-based architecture, the batch size for the SegFormer was necessarily reduced to 16.

The models were configured to train for a maximum of 200 epochs. To actively prevent overfitting and systematically capture the optimal network weights, an early stopping mechanism was implemented, halting training if the validation performance failed to improve for 12 consecutive epochs. Finally, consistent with its self-configuring architectural design, the nnU-Net framework autonomously governed its own hyperparameter optimisation and epoch management, bypassing these manual configurations.

5.3.5 Evaluation Metrics

To objectively quantify the performance of the trained DL architectures, a comprehensive evaluation framework was established. The predicted segmentation masks were assessed across three primary domains: (1) spatial overlap, evaluated via the dice similarity coefficient (DSC) and intersection over union (IoU); (2) pixel-wise classification reliability, measured through precision and recall [108; 109]; and (3) clinical feasibility for real-time intraoperative navigation, determined by computing the inference latency per US slice.

While DSC serves as the established standard for evaluating general spatial agreement in medical imaging, IoU was concurrently analysed due to its stricter penalisation of isolated false positives (FP) and FP. Furthermore, precision and recall were included to provide a granular assessment of model behaviour.

In the specific context of surgical navigation, maximising precision, thereby minimising false positive bone surface predictions, is exceptionally critical to prevent surgical instruments from being erroneously guided into surrounding soft tissues [110].

The spatial overlap and classification metrics are mathematically defined in Equations 5.6 to 5.9:

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (5.6)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (5.7)$$

$$Precision = \frac{TP}{TP + FP} \quad (5.8)$$

$$Recall = \frac{TP}{TP + FN}, \quad (5.9)$$

where TP, FP, and FN denote the total number of true positive, false positive, and false negative pixels, respectively. Finally, to rigorously assess operational viability within the strict time constraints of a surgical workflow, the inference time was recorded as the average computational latency (in ms) required for each model to process and segment a single 2D US slice.

5.4 Results and Discussion

This section presents a comprehensive evaluation of the developed DL architectures for US bone segmentation. First, model performance is analysed quantitatively, focusing on spatial overlap metrics and processing latency. Following this, a qualitative assessment visually inspects morphological consistency and addresses specific geometric failures dictated by acoustic physics.

5.4.1 Quantitative Analysis

The quantitative performance metrics of the evaluated DL architectures are summarised in Table 5.3. The self-configuring nnU-Net framework achieved the highest overall performance across both spatial overlap and pixel-wise classification metrics. It recorded a maximum DSC of 0.821 ± 0.008 , an IoU of 0.696 ± 0.011 , a precision of 0.856 ± 0.016 , and a recall of 0.7985 ± 0.027 . Following closely behind, the customised U-Net with VGG16 architecture demonstrated highly competitive performance, emerging as the second most accurate model (DSC of 0.811 ± 0.005 , IoU of 0.682 ± 0.007 , recall of $0.789 \pm$

0.005, precision of 0.777 ± 0.023). Crucially, from the perspective of clinical feasibility, the U-Net with VGG16 exhibited an average inference time of 28.70 ± 0.79 ms per slice. This indicates that while it exhibits a marginal spatial accuracy penalty compared to the nnU-Net, it operates approximately 7 times faster (compared to the nnU-Net latency of 193.85 ± 0.64 ms).

Table 5.3: Comparison of different architectures across performance metrics. Results are presented as mean $\pm$ standard deviation

Metric	nnU-Net	U-Net with VGG16	U-Net	SegFormer
DSC	0.821 ± 0.008	0.811 ± 0.005	0.745 ± 0.008	0.695 ± 0.007
IoU	0.696 ± 0.011	0.682 ± 0.007	0.594 ± 0.010	0.533 ± 0.009
Recall	0.799 ± 0.027	0.789 ± 0.011	0.717 ± 0.005	0.631 ± 0.011
Precision	0.856 ± 0.016	0.834 ± 0.002	0.777 ± 0.023	0.774 ± 0.002
Inf. Time (ms/slice)	193.85 ± 0.64	28.70 ± 0.79	19.95 ± 0.78	64.24 ± 0.62

Conversely, while the standard U-Net achieved the absolute fastest inference time (19.95 ± 0.78 ms), it suffered a substantial drop in spatial accuracy, registering DSC values of 0.745 ± 0.008 , and IoU of 0.594 ± 0.010 . Consequently, the inclusion of the pre-trained VGG16 encoder proved crucial for extracting complex textural features, effectively bridging the accuracy gap between the lightweight U-Net and the computationally heavy nnU-Net, without incurring clinically prohibitive delays.

When contextualising these results within the literature, the achievements of the proposed pipeline become particularly evident. While recent DL architectures have achieved strong spatial metrics (DSC between 0.65 [96] and 0.92 [94]), many competitive approaches still rely heavily on computationally demanding hand-crafted feature engineering, such as fusing raw B-mode images with local phase tensors [111] or multi-feature guided filters prior to CNN extraction [14; 103]. In contrast, the proposed approach extracts features directly from raw B-mode images, entirely bypassing pre-processing overheads.

Furthermore, the target anatomy plays a definitive role in model success. Extremely high DCS scores are frequently documented for morphologically superficial or flat structures, such as the hand (DSC = 0.90 [112]), finger bones (DSC = 0.88 [95]), wrist (DSC = 0.87 [113]), or long bones diaphysis like femur or radius (DSC of 0.92 [14; 103]). Conversely, when applying DL to more complex geometric structures with steep curvatures, performance drops significantly, as demonstrated by Yuan et al. [98] (DSC of 0.76 on facial bones) and Bröbner et al. [4] (DSC of 0.82 on the scaphoid). In this study, the severe acoustic shadowing and steep curves of the femoral condyles and tibial plateau inherently limit the maximum

achievable overlap, making a DSC of 0.821 on such a highly complex joint a remarkable extraction.

In addition, dataset origins heavily dictate achievable metrics. Existing literature demonstrates that models trained and tested on synthetic phantoms, devoid of clinical noise, blood flow, or tissue artefacts, exhibit inflated segmentation performance. Salehi et al. [14] reported a DSC drop from 0.92 on phantoms to 0.87 on human data, while Li et al. [96] observed a similar degradation from 0.73 on phantoms to 0.65 on human spine. In stark contrast, this study’s metrics are derived exclusively from *in vivo* clinical data, representing an honest and highly robust evaluation for real-world application.

Beyond the nature of the training data, the validation strategies employed across different studies must be critically considered, as these significantly influence the reported metrics. Notably, exceptionally high DSC values in the literature may be partially attributed to experimental designs lacking strict patient-level cross-validation. When dataset splitting is performed randomly at the image level rather than the patient level, an approach observed in studies by Zeng et al. [112], Happila et al. [95], and Yuan et al. [98], there is a severe risk of data leakage. In such scenarios, networks inadvertently memorise patient-specific anatomical features present in both training and testing sets, rather than learning to generalise true bone boundaries.

Shifting the focus to specific boundary detection, an analysis of precision and recall provides critical clinical insight. Both the nnU-Net and the U-Net with VGG16 exhibited noticeably higher precision (0.856 and 0.834) compared to recall (0.798 and 0.789). This imbalance indicates a slightly conservative, under-segmenting behaviour. Clinically, this is the ideal trade-off: for a surgical navigation system, it is vastly safer to strictly identify definitive bone boundaries while minimising FP in surrounding echogenic fascia, rather than over-segmenting and guiding surgical tools into soft tissue. The danger of prioritising recall is evident in fuzzy-based algorithms [114], which achieve recalls of 0.98 but suffer from devastatingly low precision (0.10).

Ultimately, from a clinical deployment perspective, architectural efficiency is paramount for intraoperative surgical navigation. A model exhibiting near-perfect overlap is clinically unviable if it cannot process frames in real-time. The proposed U-Net with VGG16 achieved an average inference latency of 28.70 ms per slice, which translates to a processing speed of nearly 35 frames per second (FPS). This significantly outperforms the 57.14 ms/slice reported by Bröbner et al. [4] and competes directly with the fastest optimised CNNs in recent literature, which report inference times of approximately 30 ms per scan [111]. This sub-30 ms performance satisfies the strict latency constraints of real-time optical tracking systems.

Interestingly, the evaluated Vision Transformer (SegFormer) yielded the lowest spatial metrics (DSC = 0.695). This behaviour corroborates recent observations [115] that pure transformer encoders struggle

to isolate bone boundaries from surrounding echogenic artefacts without explicit spatial constraints. Furthermore, unlike CNNs which possess strong inductive biases for local feature extraction, transformers are notoriously data-hungry, making them less suited for the relatively constrained dataset sizes typical of specialised medical imaging domains.

5.4.2 Qualitative Analysis

To provide a visual comparison of the predictions, representative qualitative results are presented in Figure 5.12, illustrating the segmentation performance across five distinct anatomical views (A to E).

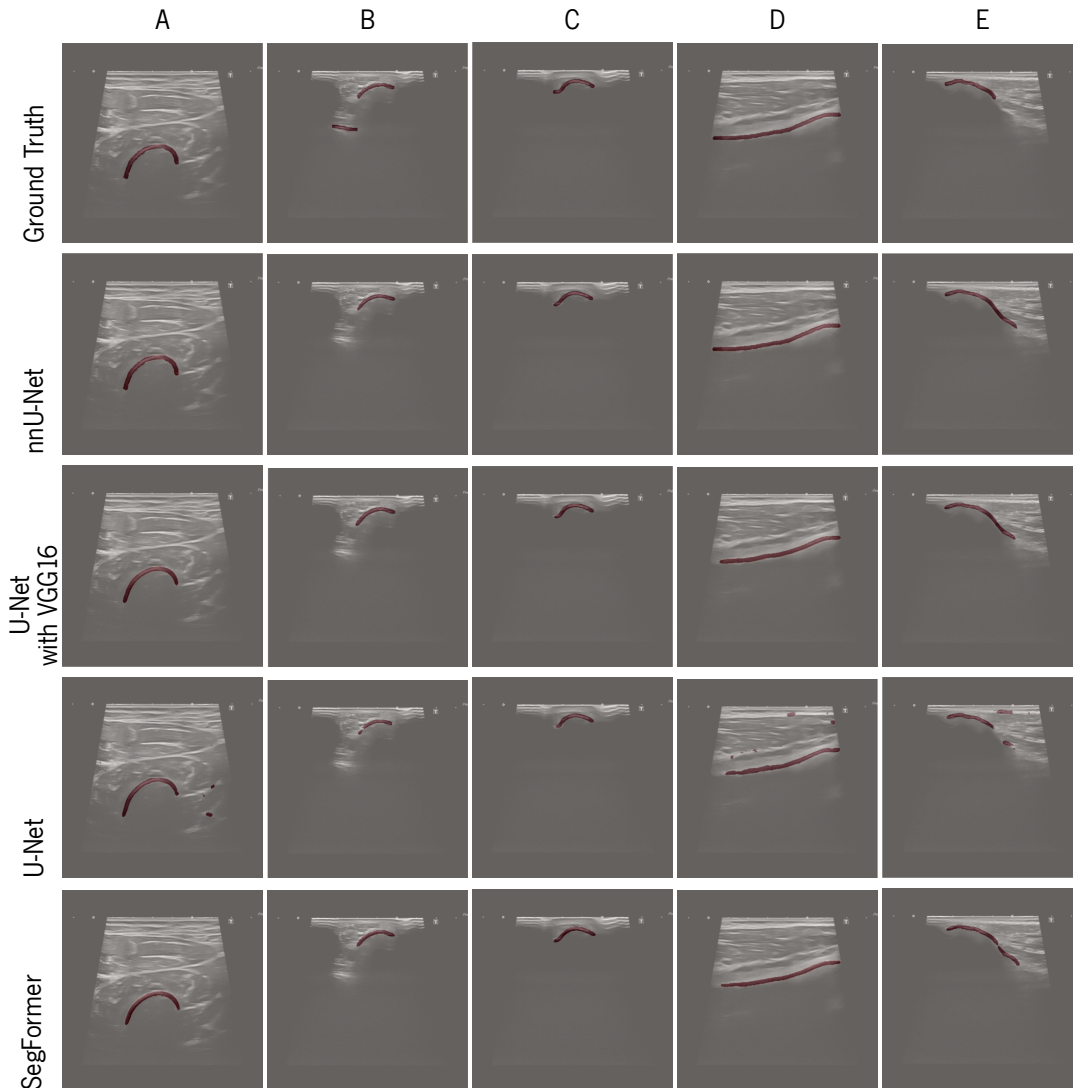


Figure 5.12: Qualitative comparison of the bone surface segmentation results (red lines) of the knee joint obtained by each developed architecture across five distinct anatomical views (columns A–E). Rows display the manual ground truth delineations, followed by the respective model predictions from the nnU-Net, UNet with VGG16, U-Net, and SegFormer models.

Visual inspection strictly aligns with the quantitative findings, revealing that the nnU-Net and U-Net with VGG16 provided the most accurate boundary delineations. Furthermore, the extracted contours from the U-Net with VGG16 appeared morphologically smoother and more continuous compared to the fragmented outputs of the standard U-Net and SegFormer.

Despite the high accuracy of the leading models, specific geometric failures dictated by US physics were observed across all architectures. In View B, all implemented models failed to segment the bone contour on the extreme left side of the image. Because cortical bone acts as a specular reflector, steep incident angles cause the US beam to reflect away from the transducer. This phenomenon creates severe signal dropout and a lack of visible hyperechogenic boundaries, prompting the models to avoid segmenting regions devoid of acoustic evidence.

The baseline U-Net frequently misclassified peripheral clusters of soft tissue as bone. Dense connective tissues, such as ligaments or deep fascia, possess high acoustic impedance, generating strong echogenic responses that visually mimic bone contours. Notably, in View D, the standard U-Net erroneously predicted disjointed bone structures overlapping at different depths on the same vertical axis. Due to total acoustic shadowing, only the first reflective boundary can be captured by the transducer, making overlapping bone surfaces a physical impossibility in US imaging. The capacity of the VGG16 and nnU-Net models to avoid this error suggests they successfully learned implicit spatial priors regarding bone continuity. Finally, in View E, all networks demonstrated a collective tendency to over-segment, predicting the continuous bone surface significantly beyond its actual limits as defined by the ground truth.

5.4.3 Limitations and Final Architecture Selection

Despite the promising results, this study presents certain limitations that must be addressed in future work. Although the dataset comprised a robust number of frames, the cohort was limited to 7 subjects acquired at a single institution using a single US machine. This localised sampling may restrict the models' ability to generalise to images acquired from different scanner vendors, varying transducer frequencies, or patients presenting with severe osseous deformities, such as advanced osteoarthritis and osteophytes.

Nevertheless, the top-performing models (nnU-Net and U-Net with VGG16) consistently delineated bone boundaries regardless of the varying soft tissue thicknesses introduced by the diverse BMI of the cohort, confirming their robustness to natural anatomical variability. Overall, the U-Net with VGG16 emerged as the undeniably superior solution for this specific surgical application. By offering an optimal balance between high spatial accuracy (DSC of 0.811) and video-rate processing latency (28.70 ms), it successfully enables the continuous, automatic, and reliable digitisation of the complex knee bone surface in real time.

5.5 Conclusion

This chapter successfully established and validated a robust DL pipeline for the automated extraction of knee bone surfaces from *in vivo* US images. Diverging from the prevalent reliance on simulated phantoms or oversimplified public datasets, this framework was fundamentally anchored in the rigorous curation and expert validation of a custom clinical dataset. This strategic foundation ensured that the developed architectures were explicitly exposed to real-world acoustic artefacts and natural anatomical variability, guaranteeing a high degree of clinical translatability.

The comparative architectural analysis yielded a critical engineering paradigm for surgical navigation: spatial accuracy cannot be pursued at the absolute expense of computational latency. While the self-configuring nnU-Net established the theoretical performance ceiling for spatial overlap (DSC of 0.821), its heavy inference burden renders it incompatible with the strict temporal demands of intraoperative tracking. Consequently, the customised U-Net with a VGG16 backbone emerged as the optimal practical solution. By maintaining highly competitive structural fidelity (DSC of 0.811) while processing at an inference speed of 28.70 ms per slice (equivalent to ~ 35 frames per second), this network successfully bridges the gap between state-of-the-art spatial extraction and real-time operational viability.

Ultimately, this chapter validates the U-Net with VGG16 as a robust, high-speed extractor of osseous morphology. The reliable 2D anatomical boundaries generated autonomously by this network now provide the essential geometric foundation required for the subsequent 3D spatial reconstruction and surface registration pipeline detailed in Chapter 6.

Chapter 6

3D Bone Reconstruction and Registration

This chapter details the methodological framework and the engineering architecture for 3D volumetric reconstruction and spatial registration of the knee joint. Based on the 2D automated segmentations achieved in Chapter 5, this phase details the transformation of these isolated frames into cohesive 3D point clouds using optical tracking, followed by their spatial alignment with a preoperative target anatomical reference (CT scan). Ultimately, the primary objective of this chapter is to address the final **(RQ)**: *How do the spatial accuracy and computational performance of the proposed non-invasive registration pipeline align with the clinical demands of surgical navigation?*

6.1 System Calibration

The implementation of the system calibration was conducted using the **PLUS (Public Software Library for Ultrasound) Toolkit** [88], an open-source framework specifically designed for US research. Transforming tracked 2D US data into a spatially coherent 3D volume fundamentally requires flawless spatial and temporal alignment between the imaging modality and the optical tracking system. Any residual error in calibration propagates directly to the 3D reconstruction, leading to anatomical distortions. To guarantee the geometric fidelity of the reconstruction, the rigid body attached to the US probe must provide a stable, non-obstructive, and mathematically rigid interface for the reflective markers. To fulfill these mechanical constraints, a custom tracking reference was engineered using SolidWorks CAD software (Figure 6.1a and Appendix A). Crucially, to ensure strict geometric tolerances and enable rapid prototyping, this bespoke mount was manufactured in-house using a Bambu Lab X1 Carbon 3D printer with Polylactic Acid filament. This ensures that the spatial relationship between the optical markers and the probe's acoustic footprint remains strictly invariant during freehand manipulation.

The spatial calibration protocol requires the interaction of a rigid tracking stylus with an N-wire fCal-2.1 calibration phantom. For this purpose, a pre-existing 3D-printed stylus (Figure 6.1b) was integrated into

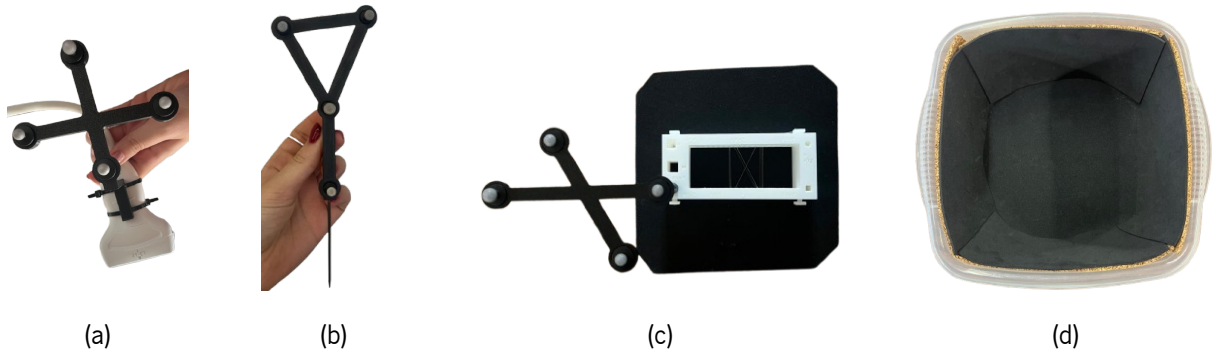


Figure 6.1: Hardware components utilised for system calibration: (a) US probe equipped with the custom-designed 3D-printed tracking reference, (b) pre-existing 3D-printed stylus, (c) fCal-2.1 calibration phantom (PLUS Toolkit standard), and (d) acoustically insulated water container.

the current pipeline. Regarding the calibration phantom, while its standardised CAD geometry is provided as an open-source standard by the PLUS Toolkit community, its chassis and corresponding tracking reference (Figure 6.1c and Appendix B) were 3D-printed in the laboratory using the same 3D printer and filament. This approach guaranteed uniform dimensional accuracy and seamless structural integration with the optical tracker. To enable optimal US wave propagation during the scans, while avoiding severe acoustic reverberation from the container boundaries, the phantom was immersed in a controlled aqueous environment (Figure 6.1d). The inner surfaces of this water-filled container were lined with highly attenuating materials, specifically cork and Ethylene-Vinyl Acetate (EVA) foam, to maximise acoustic wave absorption and scattering, thereby significantly reducing image artefacts.

To achieve full system synchronisation, the calibration workflow was divided into three sequential stages: preliminary physical registration, temporal synchronisation, and final spatial alignment.

6.1.1 Stylus Calibration and Phantom Registration

Establishing the global coordinate framework requires two preliminary alignments. First, the rigid displacement between the tracked marker and the stylus tip is determined via a pivot calibration (Figure 6.2a). By maintaining the stylus tip at a fixed spatial point while rotating its body across multiple orientations, the software computes the precise transformation matrix and rigid offset through a least-squares optimisation of the sampled data.

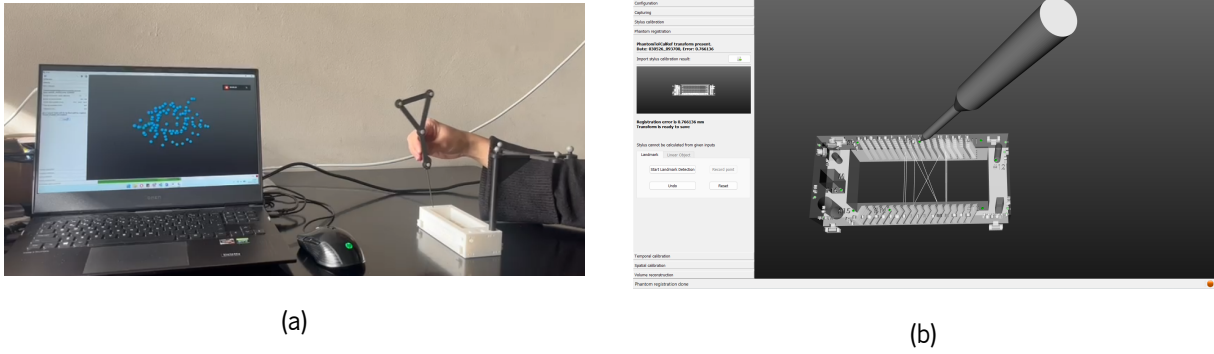


Figure 6.2: Preliminary steps: (a) stylus pivot calibration, and (b) phantom registration.

Subsequently, this calibrated stylus is utilised to register the fCal phantom by digitising 16 defined reference landmarks (Figure 6.2b). By comparing these physically acquired spatial coordinates with the known theoretical geometry defined in the system’s XML configuration, a transformation matrix is calculated, seamlessly aligning the phantom’s virtual coordinate space with the optical marker attached to its chassis.

6.1.2 Temporal Calibration

The optical tracking system transmits spatial data at a significantly higher frequency (120 Hz) than the image frame grabber (60 fps), inherently creating a temporal mismatch. Furthermore, data transmission latency is susceptible to hardware fluctuations, such as central processing unit load. Consequently, to isolate the true system lag, temporal calibration was strictly executed with minimal background processes running on the workstation.

The protocol involved inducing a periodic vertical oscillatory motion of the probe inside the water container (Figure 6.3a). The flat bottom of the container served as an absolute acoustic reference line. During acquisition, an image processing algorithm automatically extracted the Y-coordinate of this bottom boundary from the US frames, generating a time signal. Concurrently, the tracking system recorded the probe’s physical spatial translation as a secondary time signal. These two outputs, exhibiting distinct sinusoidal waveforms due to the repetitive movement, were computationally compared to determine the precise temporal offset that yielded the maximum mathematical correlation (Figure 6.3b).

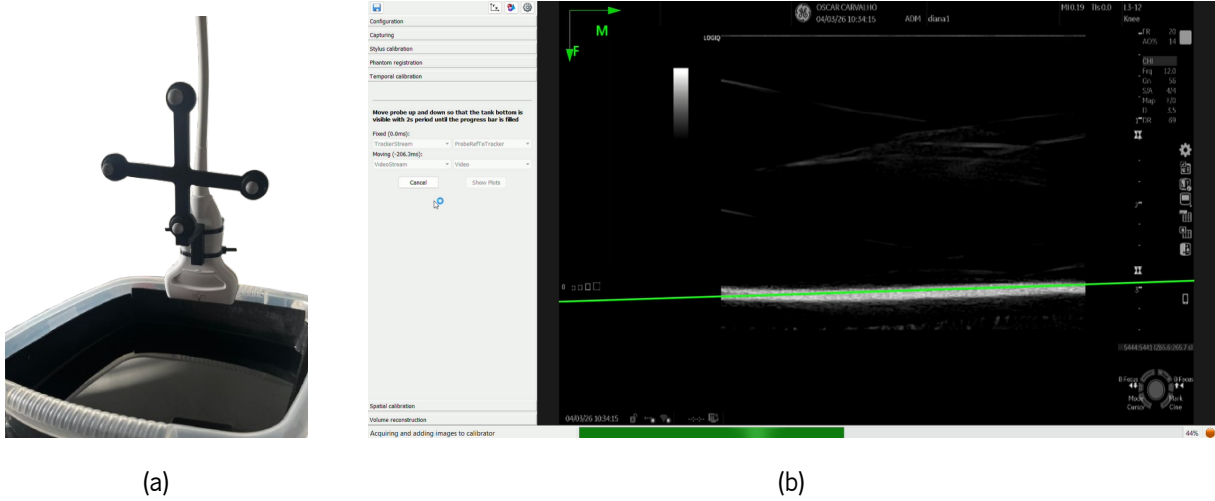


Figure 6.3: Temporal calibration procedure: (a) physical setup for oscillatory motion, and (b) software interface for signal correlation.

6.1.3 Spatial Calibration

The final calibration stage consolidates the system by computing the 'ProbeRefToTracker' transformation matrix, which directly maps the 2D pixel coordinates of the US image into the 3D physical coordinate system of the probe's marker. The comprehensive spatial relationships governing this process are mapped in Figure 6.4.

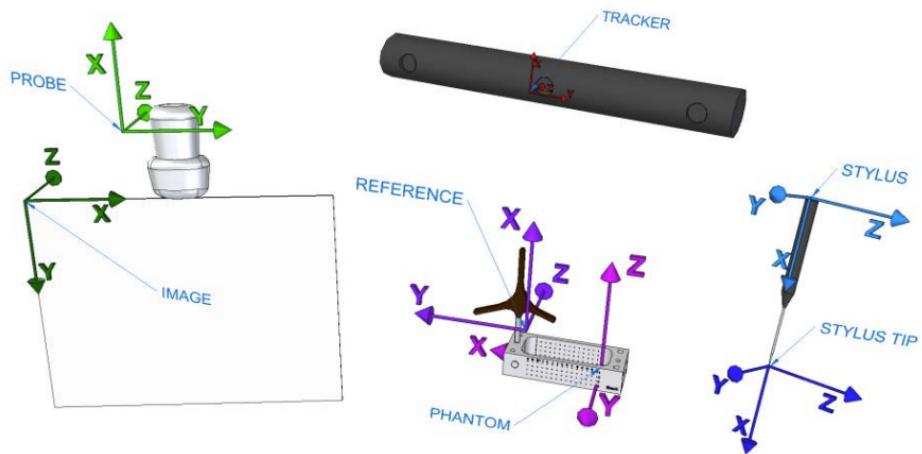


Figure 6.4: Coordinate systems of the components involved in the calibration [116].

This spatial mapping was achieved by executing a freehand US scan along the length of the submerged N-wire fCal phantom (Figure 6.5a). The algorithm continuously correlates the real-time cross-sections of the wires detected in the US images with their theoretical 3D spatial coordinates. To ensure robust

feature extraction and mitigate background noise, specific morphological parameters were iteratively fine-tuned: the wire detection opening radius was constrained to 0.10 mm, and the spatial conversion factor was empirically established at 0.11 mm/pixel. To guarantee statistical relevance for the least-squares calculation, the acquisition was configured to collect 400 consecutive valid frames at a 50 ms recording interval. The real-time algorithmic detection of these wire points is visualised in Figure 6.5b.

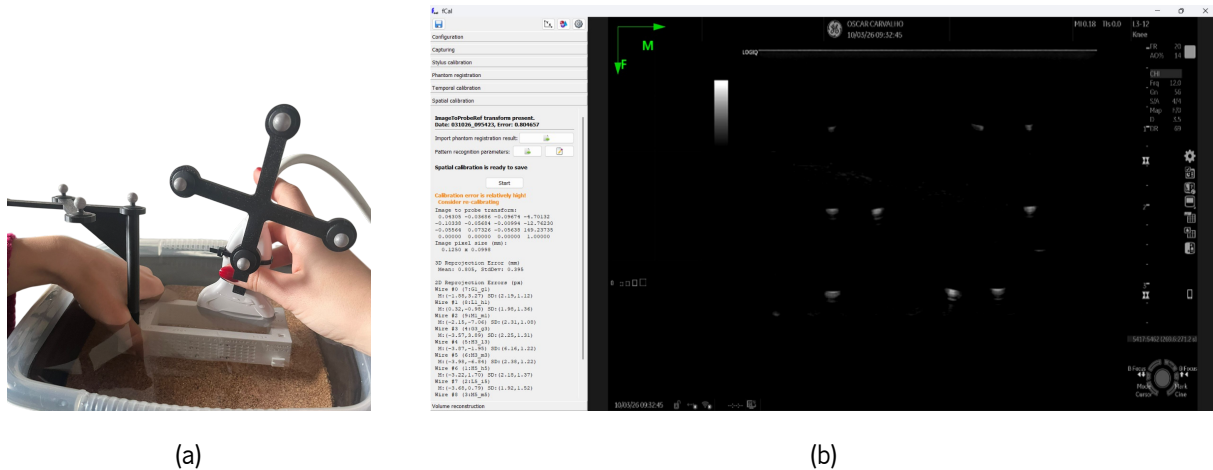


Figure 6.5: Spatial calibration procedure: (a) probe manipulation over the N-wire fCal phantom, and (b) real-time algorithmic detection of the wire cross-sections.

6.2 3D Volume Reconstruction

6.2.1 Experimental Phantoms and Acquisition

To rigorously evaluate the proposed freehand US 3D reconstruction pipeline, data acquisitions were performed under two distinct physical setups. Initial baseline experiments were conducted on a stationary physical femur phantom submerged in a simple water bath (Figure 6.6a). Subsequently, to accurately simulate the acoustic impedance, wave attenuation, and speckle characteristics typical of human soft tissue, acquisitions were performed on a custom tissue-mimicking phantom. This was fabricated using a high-density gelatine and glycerine matrix, with the physical femur deeply embedded within it (Figure 6.6b). The femur phantom was anatomically derived from CT scans of human subjects and did not incorporate integrated tracking references. To prevent any registration bias or spatial drift, the phantoms and their respective containers were rigidly fixed. This experimental setup ensured that both the internal anatomical structures and the global support frame remained completely immobile throughout the entire acquisition process.



(a)



(b)

Figure 6.6: Experimental setup showing the stationary physical femur phantom submerged in: (a) water, and (b) tissue-mimicking.

6.2.2 Data Communication and Module Architecture

Once the calibration protocols were concluded, the system was configured for freehand US scanning. To enable accurate spatial tracking, a dedicated XML configuration file was generated, incorporating the final spatial calibration matrix, the temporal offset, and the PlusOpenIGTLinkServer settings. During an acquisition, the PLUS Server Launcher runs with this configuration, streaming two synchronised data channels, the real-time US images and the 'ProbeRefToTracker' spatial transformations, directly to 3D Slicer via the OpenIGTLink protocol (Figure 6.7).



Figure 6.7: Workflow linking the data acquisition, the PLUS server, and the 3D Slicer client through the OpenIGTLink protocol.

To orchestrate data reception, preprocessing, and automated inference, a custom software module named 'SlicerFreehandUS' was developed within 3D Slicer (Figure 6.8). Upon data stream initialization, this module seamlessly bridges the real-time image feed with the trained DL architecture. Incoming frames undergo the preprocessing pipeline detailed in Chapter 5, and the generated 2D segmentations are automatically mapped to the probe's spatial coordinates, projecting the osseous contours into 3D space.

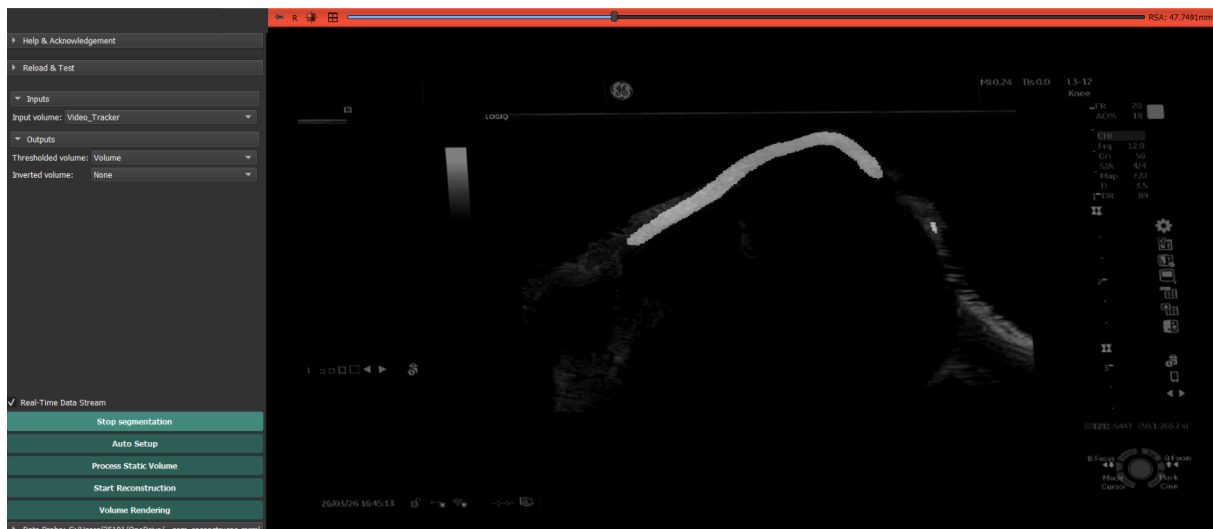


Figure 6.8: 'SlicerFreehandUS' module performing real-time automated segmentation.

6.2.3 Real-Time Volume Reconstruction

To transform the continuous stream of spatially tracked 2D slices into a coherent solid, the SlicerIGT Volume Reconstruction algorithm was programmatically embedded into the 'SlicerFreehandUS' module. Operating in a dynamic 'Live' mode, the system applies linear interpolation between the incoming segmented masks within an automatically bounded Region of Interest (ROI) that remains fixed in the 3D tracking space. To provide the operator with immediate visual feedback during the freehand sweep, an automated volume rendering workflow was integrated. By applying a predefined US-Fetal preset, the rendering node dynamically updates, allowing the operator to monitor the bone reconstruction continuously materialising in the 3D viewer (Figure 6.9).

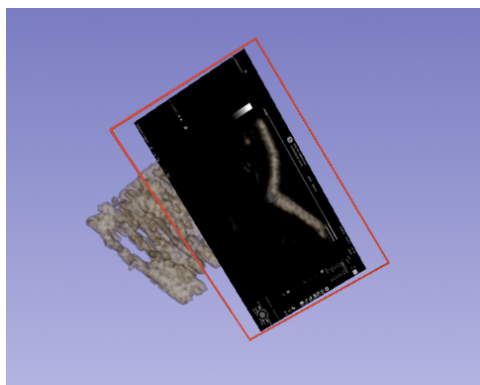


Figure 6.9: 'SlicerFreehandUS' rendering the reconstructed bone surface in real-time.

Finally, to establish an absolute anatomical ground truth for the subsequent registration phase, a highly accurate 3D model was generated from preoperative CT scans of the same human subject used to create the femur phantoms. By applying strict intensity thresholding, 3D Slicer computed the reference volumetric model illustrated in Figure 6.10.



Figure 6.10: Ground truth volumetric reconstruction from preoperative CT data.

6.2.4 Point Cloud Extraction and Refinement

To transition from volumetric visualisation to a mathematically rigorous representation, the raw 3D US reconstruction underwent a multi-stage data conditioning pipeline. As illustrated in the block diagram in Figure 6.11, this workflow is divided into two main phases: visual morphological refinement of the volume, followed by algorithmic spatial filtering of the exported point cloud.

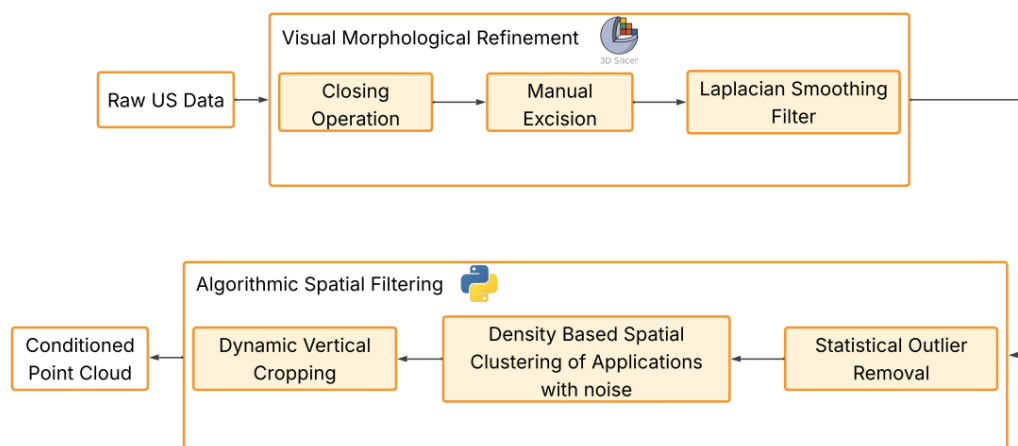


Figure 6.11: Sequential data refinement pipeline: transitioning from raw 3D volumetric reconstruction to a conditioned point cloud.

Initially, a preliminary morphological cleanup was performed using the Segment Editor module in 3D Slicer. This phase involved a closing operation to fill structural gaps, followed by the manual excision of prominent, isolated tracking artefacts. Subsequently, a Laplacian smoothing filter was applied to improve mesh regularity while preserving the overall anatomical boundaries. Despite the visual clarity achieved in the previous phase, the exported point cloud inherently retained acoustic fluctuation noise and outlier predictions at the coordinate level.

Because the subsequent registration algorithms are highly sensitive to spatial anomalies, a dedicated three-stage algorithmic pipeline was implemented to condition the data using a python script. First, to eliminate fluctuation noise, the Open3D statistical outlier removal function [117] was applied, actively discarding points located more than 1.5 standard deviations from the average distance of their 50 nearest neighbours. Secondly, the density-based spatial clustering of applications with noise algorithm was utilised to identify and isolate the primary anatomical mass. By grouping points with a strict spatial connectivity threshold of 3.5 mm, the main continuous cluster representing the femur was successfully extracted. Finally, to remove any remaining reverberation artefacts overlapping the true anatomy, a dynamic vertical crop was programmatically established. By calculating a threshold based on the highest Z-coordinate of the identified bone structure plus a 2 mm safety margin, all spatial points extending beyond this upper limit were mathematically filtered.

The visual impact of this complete refinement pipeline, contrasting the raw exported data against the clean structural reconstruction, is demonstrated in Figure 6.12.

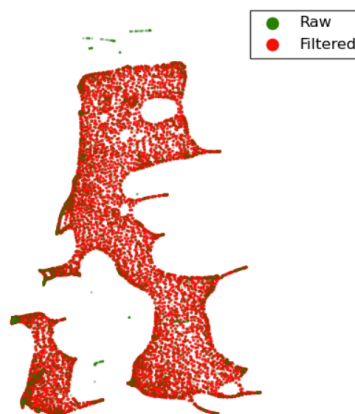


Figure 6.12: Visual comparison demonstrating the raw point cloud (green) and the resulting anatomically filtered structure (red).

6.3 3D Surface Registration

The objective of this pipeline is to establish a non-invasive, surface-based rigid registration between the intraoperative US point cloud and the preoperative CT model. As the bone is physically inaccessible during the procedure, spatial alignment must rely exclusively on the acquired surface geometries.

Based on the literature review (Chapter 3), registration can be performed using classical, probabilistic, combined, or learning-based methods. For this framework, the latter two were excluded: learning-based approaches are unfeasible due to the severe logistical and ethical constraints of acquiring large paired US-CT datasets, while combined methods introduce high intraoperative computational overhead and the risk of propagating alignment errors across their multiple stages. Consequently, to ensure computational efficiency and allow for a rigorous comparative analysis, this study evaluates two distinct foundational approaches: a classical geometric algorithm (for its proven reliability) and a probabilistic algorithm (designed to enhance robustness against acoustic noise and partial overlap).

6.3.1 Cross-Modality Data Harmonisation

A fundamental geometric challenge in US to CT bone registration is the severe partial overlap. Inherently, a US probe captures only a restricted, localised acoustic window of the bone surface, whereas the preoperative CT model represents the complete 3D anatomy. In the context of this study, due to the specific constraints of the physical phantoms and the scanning setups, the acquired US data primarily corresponded to the anterior view of the femur. If left uncorrected, registration algorithms tend to force erroneous correspondences between this localised US patch and distant, unrepresented regions of the CT bone (such as the posterior surface), leading to highly distorted alignments.

To mitigate this discrepancy, the CT-derived point cloud underwent a geometric adaptation phase to simulate the inherent acoustic constraints of the US scan. The algorithm automatically isolated the corresponding anterior view of the CT model and applied a grid-based surface extraction method to isolate a continuous thick outer shell. This effectively replicated the anatomical boundaries accessible to the US waves during the acquisitions (Figure 6.13a). Finally, a voxel downsampling filter with a voxel size of 1.5 mm was applied to both point clouds (Figure 6.13b), equalising the spatial resolution to ensure mathematical balance and improve computational efficiency.

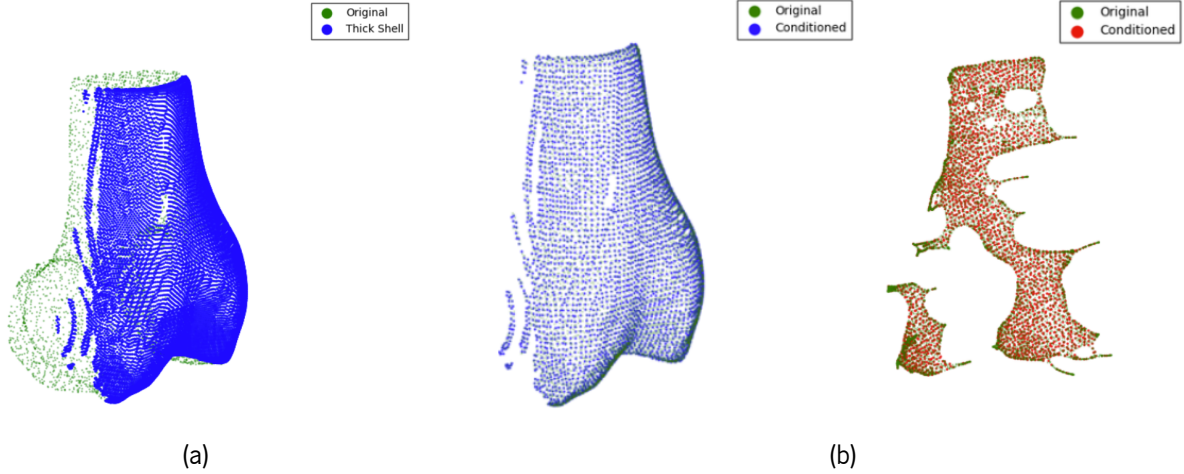


Figure 6.13: Data refinement and normalisation: (a) geometric adaptation of the CT model to extract the anterior acoustic shell, and (b) spatial resolution equalisation for both CT and US point clouds.

6.3.2 Coarse Registration

To provide an optimal initial alignment for the subsequent high-precision algorithms, a robust, two-stage automatic coarse registration workflow was implemented. In this pipeline, the CT-derived point cloud was defined as the fixed target geometry, while the US-derived point cloud acted as the moving source. The objective of this phase is not to achieve a perfect anatomical match, but rather to reduce massive initial spatial discrepancies, transforming a potentially divergent global optimisation problem into a well-conditioned local task.

This hierarchical approach [118] was executed sequentially. Initially, the severe global translation offset was addressed by calculating the spatial centroids (barycentres) of both point clouds. By programmatically aligning these centroids, the system forces the US patch into the immediate spatial neighbourhood of the CT model with negligible computational cost.

While centroid alignment effectively resolves translation, significant rotational misalignments typically persist. To address this, the second stage employed a preliminary, large-scale ICP alignment. Although ICP is standardly associated with fine registration, utilising it in this phase is a recognised multi-resolution strategy. Specifically, a standard point-to-point ICP was executed using an intentionally wide correspondence threshold of 20 mm. This aggressively large capture radius is a crucial design choice: it prevents the algorithm from prematurely locking into local minima, allowing it to pull the US point cloud into the correct general anatomical orientation (Figure 6.14).

By the end of this coarse phase, the major rotational and translational offsets are mitigated, ensuring the data is optimally positioned for the rigorous fine registration algorithms to converge successfully.

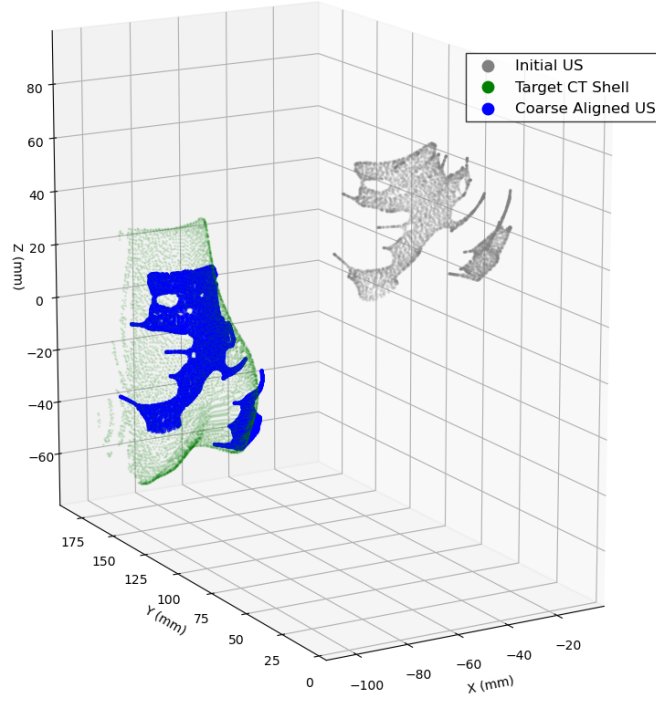


Figure 6.14: Visualisation of the spatial alignment achieved by the two-stage automatic coarse registration algorithm.

6.3.3 Fine Registration

Following the initial coarse alignment, the final high-precision registration was executed. To ensure a robust comparative analysis, this implementation evaluated two distinct mathematical formulations: a classic geometric approach (Point-to-Plane ICP) and a probabilistic framework (CPD).

The **ICP** algorithm [119] is the standard geometric method for rigid registration. While the classic version minimises the Euclidean distance between points, this implementation utilised a more robust Point-to-Plane ICP variant [63; 65]. Instead of forcing direct point matches, this algorithm minimises the orthogonal distance between a source point and the tangent plane of the target surface. This formulation allows the US patch to perform a "sliding fit" along the smooth cortical surface of the CT bone, drastically reducing the chances of converging to local minima. In practice, the optimisation computes the optimal rotation (R) and translation (t) by minimising the following cost function:

$$E_{\text{point to plane}} = \sum_i ((R \cdot x_i + t - y_i) \cdot n_i)^2, \quad (6.1)$$

To compute the necessary unit normal vectors (n_i) at the target point y_i , a hybrid K-Dimensional

(KD) tree search [120] with a 5 mm neighbourhood radius was implemented. This approach ensures stable geometric descriptors even in regions with varying point density, allowing the algorithm to establish accurate correspondences between the localised US patch and the complete CT bone model.

As an alternative to purely geometric alignment, the **CPD** algorithm [121] was implemented. CPD treats registration as a probability density estimation problem, where the source point cloud represents the centroids of a gaussian mixture model (GMM), while the target point cloud represents the observed data. A key advantage of CPD for US data is its motion coherence constraint, which forces points to move together, preserving the overall topological shape of the bone patch by minimising this energy function:

$$E_{CPD}(Y) = - \sum_{n=1}^N \log \sum_{m=1}^M e^{-\frac{1}{2} \left| \frac{x_n - y_m}{\sigma} \right|^2} + \frac{\lambda}{2} \phi(Y) \quad (6.2)$$

In this formulation, the first term corresponds to the negative log-likelihood, $\phi(Y)$ is a regularisation term that enforces motion coherence, and λ controls the trade-off between data fitting and regularisation.

To minimise this energy function, the transformation is computationally solved using the expectation-maximisation algorithm. Crucially for this pipeline, during the expectation step, the algorithm was explicitly configured with a noise weight parameter of $w = 0.4$. This parameter instructs the probabilistic model to expect 40% of the input data to be acoustic noise or outliers, allowing it to maintain an accurate fit even when presented with unrefined US artefacts. The iteration process continued until the maximum likelihood estimation stabilised across the target bone model. The results obtained from these fine registration methods are detailed in Section 6.5.2.

6.4 Evaluation Strategy and Metrics

In the absence of physical fiducial markers, which would enable direct computation of absolute registration errors [122], a rigorous evaluation protocol was established based on the principles of spatial hold-out validation. As highlighted by Roberts et al. [123], when data points exhibit an inherent spatial structure, such as a continuous bone surface, nearby points are naturally correlated. Under these conditions, conventional random splitting of the data leads to overly optimistic and biased error estimates, as the training and test samples are not truly independent. To mitigate this spatial correlation bias, the US point cloud was partitioned into distinct spatial blocks using the K-means clustering algorithm, as illustrated in Figure 6.15.

This spatial partitioning enabled a highly strategic training and testing split based on anatomical relevance. Specifically, the condyles (identified as clusters 2 and 4) are essential for driving the registration

algorithm due to their complex geometry, which provides strong geometric constraints against the rotational ambiguity typical of the femoral shaft. However, to ensure global alignment stability, this distal anchoring must be complemented with spatial information from the proximal cylindrical region (cluster 1) and the central shaft continuity (cluster 0). Consequently, these four regions were utilised to optimise the registration, while cluster 3 was deliberately withheld from the optimisation process to serve exclusively as an independent spatial test set.

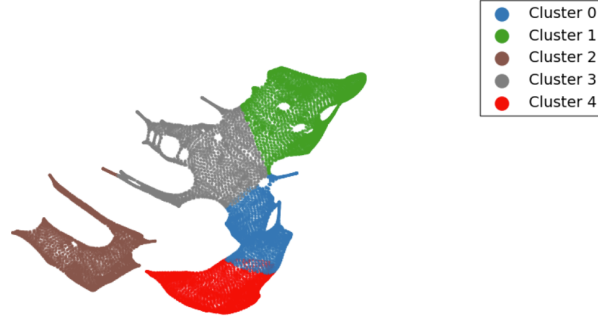


Figure 6.15: Spatial partitioning of the femur point cloud into five distinct anatomical clusters for hold-out validation.

To quantitatively evaluate the accuracy and clinical viability of the proposed US to CT registration pipeline, three primary metrics were selected based on established literature standards: surface-matching accuracy, target positional error, and computational efficiency.

The algorithmic surface-matching accuracy was evaluated using the RMSE. This metric quantifies the average Euclidean distance between the registered US point cloud and the nearest corresponding points on the CT model [53]. The RMSE is mathematically defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|T_{reg}(p_i^{US}) - p_i^{CT}\|^2}, \quad (6.3)$$

where N is the total number of aligned points, $T_{reg}(p_i^{US})$ represents the transformed US points, and p_i^{CT} corresponds to their nearest neighbours on the CT surface.

While RMSE assesses global surface proximity, the TRE was computed to evaluate the true positional error of specific bone surface points [80]. TRE is defined as the Euclidean distance between target points in the CT space and their corresponding positions after applying the calculated registration transformation. For a set of M targets, the mean TRE is calculated as:

$$TRE = \frac{1}{M} \sum_{j=1}^M \|T_{reg}(t_j^{US}) - t_j^{GT}\|, \quad (6.4)$$

where t_j^{US} is the position of the target point in the initial US coordinate system, and t_j^{GT} is its true ground truth position in the CT coordinate system.

Finally, because time efficiency is a critical factor for the viability of intraoperative surgical procedures, the computational execution time of the pipeline was strictly monitored. The processing duration for segmentation, 3D reconstruction, and surface registration was recorded to assess the system's compatibility with real-time navigation constraints.

6.5 Results and Discussion

This chapter details the quantitative outcomes and critical analysis of the implemented US to CT bone registration pipeline. To ensure a cohesive evaluation, the findings are presented and immediately discussed sequentially: first, the foundational accuracy established during system calibration, followed by the structural fidelity and computational efficiency of the 3D reconstruction and surface registration algorithms.

6.5.1 Calibration

The evaluation began with the spatial calibration of the hardware setup. The remarkably low pivot calibration error of 0.26 mm demonstrates exceptional tracking performance, surpassing the 0.5 mm RMS accuracy specified by the optical tracking manufacturer. This confirms that the experimental setup was highly optimised, benefiting from a rigid optical marker mount and strict line-of-sight maintenance, which effectively minimised undesired infrared reflections and tracking jitter. Similarly, subsequent phantom registration error of 0.77 mm indicates an accurate spatial mapping between the measured fiducial positions and their ground-truth reference points. This ensures a highly reliable transformation between the phantom and the tracker coordinate systems. The minor residual error observed in this step is no longer governed by intrinsic tracking noise, but is rather a consequence of the mechanical manufacturing tolerances of the 3D-printed phantom and the inherent human inconsistency during the manual point-picking process.

Following the spatial alignment, the temporal calibration of the system was assessed. Based on the 15 dynamic measurements detailed in Appendix C.1, the temporal analysis (Figure 6.16a) revealed an average latency of 187.31 ± 15.52 ms. This delay represents the cumulative processing overhead of the entire data

acquisition pipeline, encompassing the frame grabber digitisation, optical tracking transmission, and the PLUS Toolkit synchronisation algorithms. Statistically, the proximity between the mean (187.31 ms) and the median (185.90 ms), combined with a low coefficient of variation of 8.29% (despite a single identified outlier of 226.0 ms), characterises a highly stable and predictable temporal distribution.

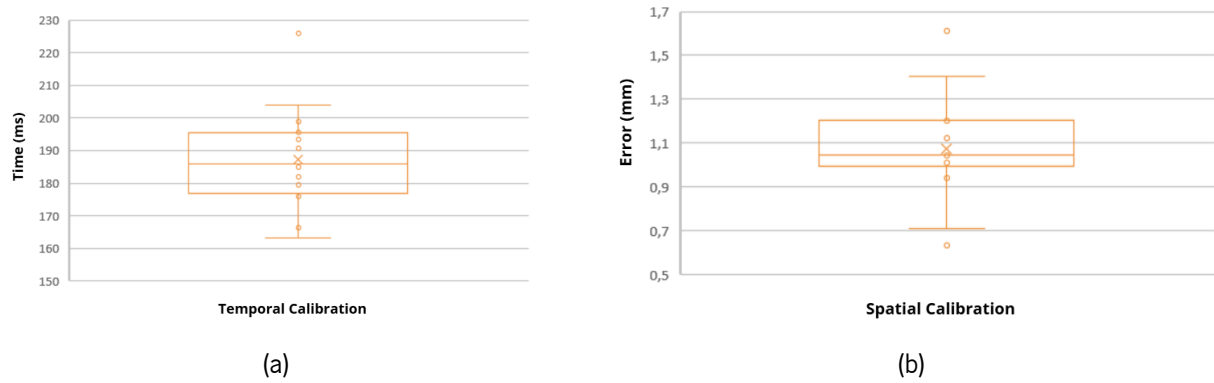


Figure 6.16: Statistical distribution of the system calibration metrics: (a) temporal latency (ms) for US-tracking synchronisation, and (b) spatial registration error (mm) evaluated under dynamic freehand motion. Boxplots illustrate the median, interquartile ranges, and outliers across the 15 independent acquisitions.

While a baseline mean latency of ~ 187 ms would be noticeable for instantaneous visual feedback, its high predictability allows the temporal calibration algorithm to mathematically compensate for the delay during the freehand US sweep. By actively applying this calculated latency to offset the data stream, the temporal misalignment between the tracking and imaging systems is successfully neutralised. This is clearly demonstrated in Figure 6.17, where the calibrated moving signal exhibits robust synchronisation and close morphological agreement with the fixed reference signal across the entire periodic motion.

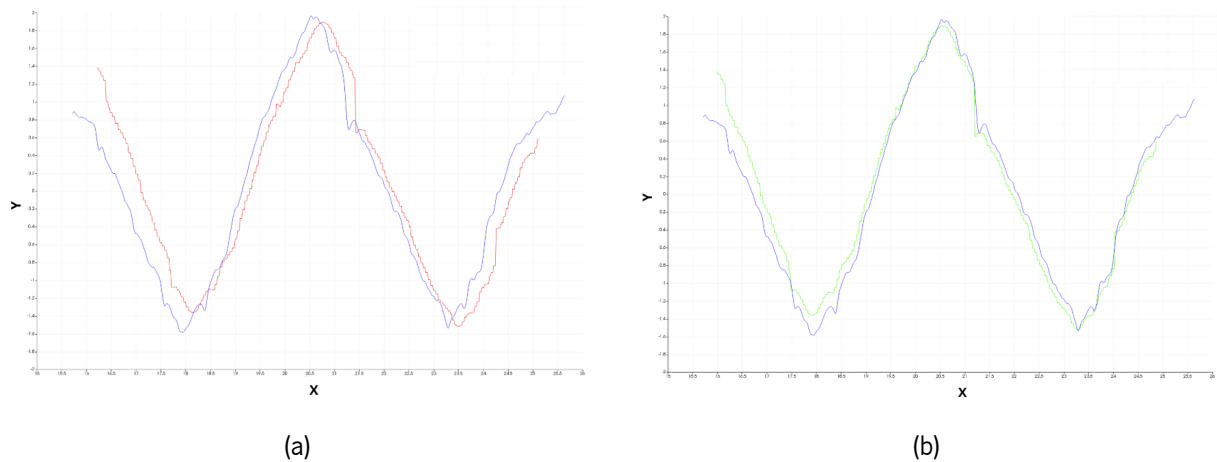


Figure 6.17: Temporal calibration results illustrating the fixed (blue line) and moving (red and green lines) tracking signals (a) before, and (b) after applying the calculated latency value.

Following the temporal alignment, the dynamic spatial calibration of the US system was evaluated. Based on 15 independent acquisitions, the procedure yielded an average spatial error of 1.06 ± 0.25 mm. As illustrated in Figure 6.16b, the resulting error distribution exhibits high symmetry. The tight proximity between the mean (1.06 mm) and the median (1.05 mm) confirms that the calibration pipeline is robust and free from significant systematic bias. In terms of overall accuracy, this result is highly competitive, outperforming the 1.40 mm mean error reported by Carbaja et al. [124] using a traditional N-wire phantom configuration. This reduction in spatial error is primarily attributed to the meticulous mechanical and acoustic design of the experimental setup. Unlike traditional open-water configurations, the integration of cork and EVA foam acoustic insulation significantly suppressed reverberation artefacts within the water tank. Consequently, the signal-to-noise ratio of the N-wire cross-sections was greatly enhanced, enabling the PLUS Toolkit line-detection algorithms to compute the centroids with superior precision. Nevertheless, the calculated coefficient of variation (23.18%) indicates a reasonable degree of dispersion across individual measurements. This variance suggests that, while the system's average accuracy is excellent, instantaneous measurements remain susceptible to transient external factors, such as acoustic acquisition noise or intrinsic hardware limitations (the complete spatial measurements are provided in Appendix C.2).

To properly contextualise this dynamic dispersion, a static spatial calibration test was conducted with the US probe securely immobilised. This baseline test aimed to isolate and quantify the inherent system noise, primarily optical tracking jitter and US image segmentation variance, independent of user motion. Under these static conditions, the system achieved a baseline average error of 0.50 ± 0.20 mm (ranging from 0.25 to 0.75 mm), with a median of 0.51 mm (detailed in Appendix C.3). Crucially, comparing this static baseline with the average dynamic error (1.06 mm) explicitly demonstrates that the physical motion of the freehand sweep, coupled with residual temporal latency, effectively doubles the spatial tracking variance. Consequently, this dynamic calibration variance of 1.06 mm establishes the theoretical lower bound for the system's global accuracy, inevitably propagating through the subsequent 3D reconstruction and ICP alignment phases to influence the final TRE.

6.5.2 3D Reconstruction and Registration

This stage evaluated the transition from 2D US scans to a comprehensively registered 3D spatial representation. Data acquisition was performed on a femur phantom submerged in two distinct acoustic propagation media: water and a tissue-mimicking fluid.

The custom reconstruction algorithm successfully generated a point cloud of the bone surface with an average processing time of 85.40 s, as detailed in Table 6.1. Because the geometric fidelity of the reconstructed surface is ultimately validated through its alignment accuracy, the evaluation of this stage was intrinsically tied to the final spatial registration results.

Table 6.1: Breakdown of average computational execution times for the 3D surface reconstruction pipeline

Pipeline Stage	Time (s)
Data acquisition, preprocessing and 3D reconstruction	35.83
point cloud postprocessing	49.57
Total	85.40

The quantitative analysis of the registration algorithms, presented in Table 6.2, revealed distinct performance profiles. In water, the ICP algorithm outperformed the CPD method across all metrics, achieving a TRE of 1.46 mm and an RMSE of 1.65 mm, compared to CPD's 1.54 mm and 1.88 mm, respectively. Conversely, in the tissue-mimicking medium, the CPD approach appeared to exhibit lower numerical errors, recording a TRE of 0.96 mm. However, as will be critically addressed later in this section, this deceptively low CPD error is the result of a degenerate mathematical convergence rather than true anatomical alignment. In contrast, the ICP method maintained a highly consistent and anatomically faithful performance across both fluids, recording a stable TRE of 1.47 mm and an RMSE of 1.82 mm in the complex tissue-mimicking medium.

Table 6.2: Comparative analysis of spatial registration metrics (ICP vs. CPD) across different acoustic propagation media

Medium	Algorithm	TRE (mm)	RMSE (mm)	Time (s)
Water	ICP	1.46	1.65	2.50
	CPD	1.55	1.88	56.60
Tissue-mimicking	ICP	1.47	1.82	1.60
	CPD	0.96	1.14	20.30

Comparing the reliable ICP results with the current state-of-the-art reviewed in Chapter 3, the mean TRE of 1.46 mm across both media not only demonstrates high clinical competitiveness but actively outperforms several established methodologies. This value is substantially lower than the median TRE of 3.19 mm reported by Pandey et al. [54] for the tibia. Furthermore, the proposed pipeline surpasses the intensity-based results (1.70 mm) achieved by Gueziri et al. [58; 61], while strictly competing with their highly optimised gradient-based approach (1.27 mm). The performance of our ICP pipeline also fits comfortably within the range reported by Liu et al. [80], who observed TREs between 0.58 mm and 1.63 mm for the knee joint, with our mean remaining well below their upper maximum.

While certain studies report significantly smaller errors, such as Min et al. [71] (0.30 mm) for the femur and Schaffert et al. [65; 66] (0.32–0.49 mm) for the spine, these discrepancies are fundamentally attributable to the intraoperative imaging modalities employed. Radiographic and volumetric techniques (e.g., CT and fluoroscopy) inherently provide substantially sharper bone-to-soft-tissue contrast and a more uniform data profile than the highly noisy acoustic boundaries of US. Similarly, the structural reliability of our RMSE (1.65–1.82 mm) proves robust when contrasted with Goerres et al. [49], outperforming their final converged value of 2.00 mm. The gap between our results and those of Barbosa et al. [48] (0.29 mm) and Xie et al. [79] (0.79 mm) is equally expected, as their techniques rely on the direct, physical collection of unambiguous geometric landmarks using a tracked digitising stylus on fully exposed bone, entirely bypassing the complexities of US segmentation.

Beyond structural accuracy, a profound discrepancy between ICP and CPD emerged in computational efficiency. ICP demonstrated outstanding performance, with processing times ranging from just 1.60 to 2.50 s per registration, rendering it nearly 20 times faster than CPD (20.30–56.60 s). This ICP performance rivals the fastest methods in the literature, matching the 1.75 s achieved by Xie et al. [79], and significantly outperforming the times reported by Chan et al. [74] (8.80 s) and Gueziri et al. [58; 61] (5.25–10.79 s). This drastic computational advantage stems directly from our rigorous data conditioning pipeline. By heavily reducing the point cloud density post-reconstruction, the ICP algorithm evaluates far fewer point correspondences per iteration compared to heavy 3D voxel-space methods.

Crucially, the specific numerical superiority of CPD in the tissue-mimicking medium must be recognised as highly misleading. Although probabilistic methods like CPD are often praised for their flexibility in complex scenarios, visual inspection (Figure 6.18) exposes a severe numerical paradox. The CPD algorithm converged to a degenerate solution, mathematically minimising local point-to-point distances at the absolute cost of global anatomical consistency. In contrast, the ICP method maintained anatomical fidelity, proving that its spatial metrics genuinely reflect the true surface alignment.

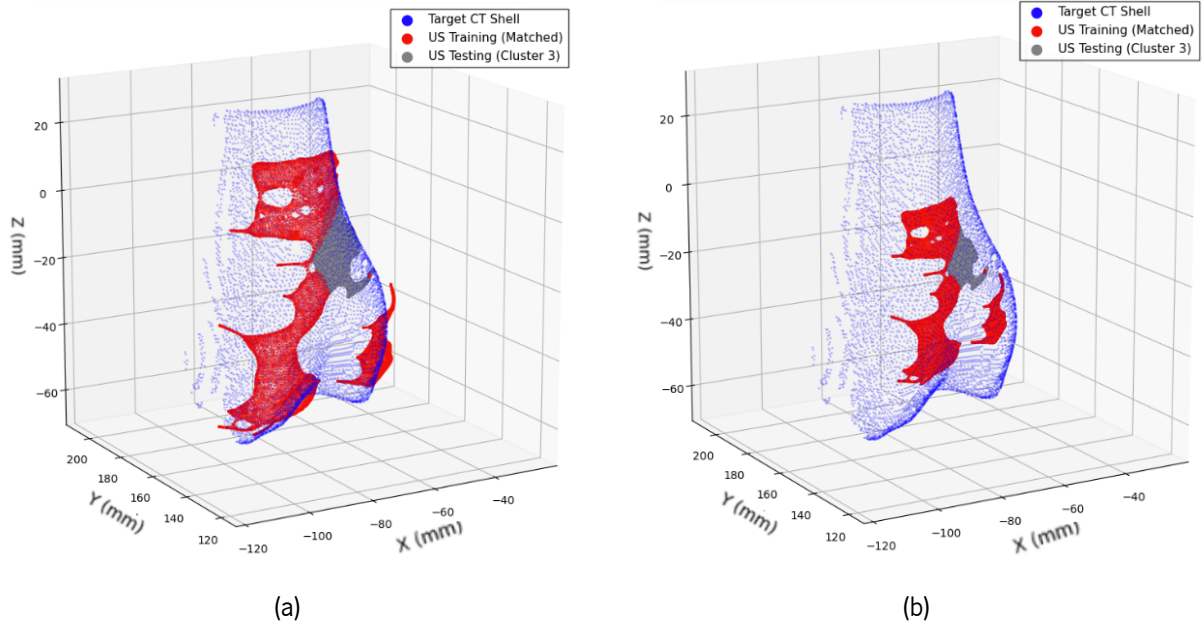


Figure 6.18: Visual comparison of the spatial registration outcomes: (a) anatomically faithful alignment achieved by ICP algorithm, and (b) degenerate mathematical convergence produced by CPD algorithm.

This visual deficiency is an inherent flaw of the algorithm's probabilistic nature, which models point sets as a Gaussian Mixture Model (GMM). Confronted with the acoustic noise, speckle, and irregular point densities typical of tissue-mimicking US, the CPD sought to maximise statistical likelihood rather than preserve geometric integrity [80]. Consequently, the target point cloud collapsed into condensed regions of high probability on the CT model. While this artificially deflates distance-based metrics (TRE/RMSE), it completely sacrifices anatomical coherence [4].

Finally, when factoring in the baseline 85.40 s required for initial point cloud generation (Table 6.2), the total computational latency of the complete ICP pipeline averages 87.45 s. In contrast, employing the CPD algorithm significantly inflates this overall latency, resulting in total execution times ranging from 105.70 to 142.00 s. Returning to the proposed ICP-driven workflow, its end-to-end performance aligns perfectly with the 1–2 minute range reported by Han et al. [62], is notably faster than the times reported by Liu et al. [80] (120.40 s for the femur and up to 254.57 s for the tibia and fibula), and drastically outperforms the 7-minute average reported by Schaffert et al. [65]. To transition towards real-time deployment, future efforts must target the primary bottleneck, the initial 3D reconstruction, by transitioning from end-of-sweep batch processing to incremental point cloud updating.

Ultimately, the proposed non-invasive US to CT bone registration framework, driven by the ICP algorithm, offers a robust, safe, and effective solution for surgical navigation. By successfully eliminating the need for invasive reference markers and intraoperative ionising radiation, while satisfying the strict 1.5

mm clinical limit (achieving a mean TRE of 1.46 mm in under 90 seconds), this system demonstrates high viability for future integration into orthopaedic workflows.

6.6 Conclusion

This chapter provided a comprehensive quantitative evaluation and critical discussion of the proposed non-invasive US to CT bone registration pipeline. Transitioning from tracked 2D US sweeps to a spatially registered 3D femoral surface, the system was rigorously validated under both aqueous and complex tissue-mimicking acoustic conditions.

Building upon the segmented US data obtained in the preceding phase, the system successfully leveraged standard communication protocols (OpenIGTLink) to stream the tracking and imaging data. This culminated in the development and seamless integration of custom software modules within the 3D Slicer environment, providing a unified, real-time visualisation interface during the reconstruction process. From a technical perspective, the rigorous spatial and temporal calibration procedures ensured a highly stable mean spatial error of 1.06 mm and a consistent latency of 187 ms. Consequently, this baseline calibration variance established the theoretical lower bound for the system's global accuracy, propagating into the subsequent registration phase.

The comparative evaluation of the algorithms yielded a critical methodological takeaway: spatial metrics do not always reflect anatomical reality. While probabilistic methods like CPD suffered a degenerate mathematical collapse, producing deceptively low error metrics at the cost of structural coherence, the ICP maintained geometric integrity. This demonstrated that in US-based bone registration, preserving global anatomical fidelity must strictly supersede the mathematical minimisation of point-to-point distances.

Clinically, the ICP-driven pipeline validated its operational viability. The system achieved a mean TRE of 1.46 mm, satisfying the 1.5 mm safety threshold for orthopaedic surgical navigation [61]. Furthermore, with an end-to-end computational latency of only 87 seconds, the temporal footprint falls comfortably within clinically acceptable limits (under 3 min and 20 s [85]), ensuring seamless integration without disrupting the standard surgical workflow.

Ultimately, the findings presented in this chapter establish the proposed US to CT surface registration framework as a highly accurate alternative to current surgical standards. By successfully balancing strict spatial accuracy with computational efficiency, this non-invasive pipeline offers a safe, zero-radiation solution that eliminates the need for invasive fiducial markers, laying a solid foundation for future *in vivo* clinical validation.

Chapter 7

Conclusion

Modern orthopaedic procedures are intrinsically complex, as surgeons must navigate and manipulate bone structures hidden beneath layers of soft tissue. Traditionally, clinical success has relied heavily on the surgeon's spatial intuition and their mental mapping of preoperative imaging to the actual patient anatomy. While surgical expertise is invaluable, this reliance on subjective estimation introduces variability that can compromise surgical outcomes and patient recovery. To provide real-time spatial awareness and standardise these procedures, CAOS systems have become instrumental. However, the foundational step of CAOS remains a major clinical bottleneck. Currently, the most robust registration methods require the invasive implantation of physical fiducial markers into the skeleton, increasing surgical trauma and infection risks. Conversely, non-invasive alternatives heavily depend on continuous intraoperative ionising radiation, such as fluoroscopy. Motivated by this critical gap, this dissertation successfully developed and validated a completely non-invasive, radiation-free 3D bone registration pipeline using US imaging, designed as a highly adaptable framework for diverse orthopaedic interventions.

Guided by a comprehensive review of CAOS systems, this dissertation strategically targeted a significant clinical gap: while CAOS-based navigation has proven successful in spinal surgeries, its application in complex knee joint procedures remains largely underexplored. Aligning with the clinical shift away from intraoperative ionising radiation, the proposed framework adopted US as a safe, real-time alternative to map against preoperative volumetric planning (CT/MRI). To guarantee clinical viability, the system's foundation was built upon current medical gold standards, utilising optical tracking for spatial positioning and classical geometric algorithms, like ICP, as the reliable benchmark for registration, while keeping pace with the progression of learning-based approaches.

Translating this conceptual framework into a functional reality required the development of a meticulously integrated, multi-phase architecture. A critical achievement of this thesis was the seamless orchestration of both hardware and software across the preoperative and intraoperative stages. The workflow successfully coupled high-resolution US imaging with precise optical tracking, intelligently distributing the

demanding computational load between an high-performance computing node and a local workstation. Ultimately, by leveraging OpenIGTLink protocols to connect a PLUS server with 3D Slicer, the project transcended isolated algorithmic testing, delivering a unified, real-time data streaming and visualisation environment that proves the technical viability of the complete US to CT registration pipeline.

A major scientific contribution of this dissertation was the development of a generalised, automated DL architecture capable of segmenting bone surfaces directly from noisy, artifact-heavy 2D US images. Addressing the lack of public data, a novel US dataset was acquired and manually labeled via a custom in-house application under the supervision of an expert musculoskeletal radiologist. Trained on this robust clinical ground truth, a rigorous comparative analysis of state-of-the-art architectures was conducted, encompassing a standard U-Net, a customised U-Net with VGG16, a nnU-Net, and the Vision Transformer SegFormer. The U-Net with VGG16 was identified as the optimal solution, achieving a DSC of 0.8108 with a remarkably low inference time of 28.70 ms per slice. By running up to 7 times faster than the highly complex nnU-Net, this architecture successfully delivered the essential balance between morphological accuracy and the strict real-time constraints dictated by the operating room.

Building on this AI-driven automated bone extraction, the project successfully transitioned into the realm of system engineering and spatial computing. The foundation of the spatial tracking was validated through a rigorous calibration protocol, achieving a highly stable temporal latency of 187.31 ms and a dynamic spatial error of just 1.06 mm. To synthesise these data streams, a custom 'SlicerFreehandUS' module was developed for 3D Slicer. This module acts as the technological heart of the system, seamlessly integrating OpenIGTLink communication, DL inference, and dynamic 3D volume reconstruction into a unified, clinician-friendly graphical interface.

The final and most critical phase of the pipeline, the 3D surface registration between the reconstructed US point cloud and the preoperative CT model, yielded highly significant biomechanical insights. The experimental evaluation exposed a mathematical paradox when using probabilistic methods on acoustic data: while the CPD algorithm produced misleadingly low numerical errors by mathematically collapsing the point cloud, the classical geometric approach, point-to-plane ICP, maintained exceptional anatomical and structural integrity. Ultimately, the ICP-based system achieved a mean TRE of 1.46 mm. This precision successfully remains below the strict 1.5 mm [61] clinical safety threshold required for high-precision orthopaedic navigation. Furthermore, the total computational latency of the entire reconstruction and registration pipeline averaged only 87.45 s, ensuring that the surgical workflow remains highly efficient.

In conclusion, this dissertation proves that US-based surface registration is a highly competitive, clinically viable alternative to the current surgical standards. By successfully satisfying both the rigorous spatial

constraints of orthopaedic navigation and the temporal demands of the operating room, this system delivers a precise and robust alignment. Most importantly, it completely eliminates the need for invasive bone markers and intraoperative radiation, significantly enhancing patient safety and paving the technical way for future in vivo clinical trials across a wide spectrum of orthopaedic procedures.

Based on these achievements, this dissertation provides definitive answers to the primary RQs formulated in Chapter 1:

RQ1: What are the current state-of-the-art techniques for bone registration in surgical navigation, and how do they compare?

Chapter 3 answered this RQ. The systematic review indicated that while current CAOS research predominantly focus on spinal procedures, there is a significant gap regarding procedures related to the knee joint. Preoperatively, volumetric CT remains the gold standard. Intraoperatively, however, a clear clinical shift towards radiation-free modalities was identified, establishing US coupled with optical tracking as the optimal navigation setup. With regard to registration algorithms, classical geometric methods, particularly the ICP algorithm, remain the clinical benchmark for surface registration, demonstrating proven reliability and achieving a minimum TRE of 1.27 mm in the literature.

RQ2: What specific clinical and technical requirements must be met to develop a viable intraoperative US-based registration system?

Chapter 4 answered this RQ. From a clinical perspective, an effective navigation system must achieve a TRE strictly below 1.5 mm and must operate within a highly efficient timeframe (under 3 min and 20 s) to prevent disrupting the surgical workflow. From a technical standpoint, satisfying these clinical demands requires rigorous spatial and temporal calibration to perfectly synchronise the US imaging stream with the optical tracking data. Furthermore, it necessitates highly robust automated bone segmentation and the implementation of standard communication protocols (such as OpenIGTLink) to guarantee seamless multimodal data integration.

RQ3: Which deep learning (DL) algorithm is best suited for extracting bone contours from US images, to reliably reconstruct point clouds for registration?

Chapter 5 answered this RQ. After a comparative analysis of various DL segmentation architectures, U-Net with a VGG16 backbone proved to be the optimal method for surgical navigation. While the highly complex nnU-Net yielded slightly higher overall spatial accuracy, the U-Net with VGG16

provided the essential clinically viable balance. It achieved a highly competitive DSC of 0.8108, alongside an exceptionally fast inference time of 28.70 ms per slice. This performance ensures that the network can extract reliable morphological bone contours from noisy US images in real-time, satisfying the strict temporal constraints of the operating room.

RQ4: How do the spatial accuracy and computational performance of the proposed non-invasive registration pipeline align with the clinical demands of surgical navigation?

Chapter 6 answered this RQ. The proposed ICP-based registration framework successfully met and exceeded all established criteria when tested under both aqueous and tissue-mimicking conditions. From a clinical perspective, it completely eliminates the need for invasive fiducial markers and intraoperative ionising radiation, while achieving a competitive mean TRE of 1.46 mm, below the strict 1.5 mm safety threshold. Furthermore, the overall computational latency averaged just 87.45 s, perfectly aligning with time-sensitive surgical workflows. Technically, the system showcased robust engineering integration, backed by precise spatial (1.06 mm error) and temporal (187.31 ms latency) calibration. Ultimately, leveraging OpenIGTLink protocols, the pipeline culminated in the custom-developed 'SlicerFreehandUS' module, providing surgeons with a unified, real-time 3D reconstruction and navigation environment.

7.1 Future Work

While this dissertation successfully establishes a robust foundation for non-invasive, US-based surgical navigation, the transition from a controlled laboratory environment to a fully dynamic clinical setting presents several promising avenues for future research. Specifically, future work comprises the following research directions: (i) to extend the registration methodology to encompass the tibia, performing initial validation on animal bones to evaluate the algorithm's performance on true cortical bone reflectivity, prior to pre-clinical cadaveric studies; (ii) to progress towards in vivo trials by targeting an opportunistic study with patients who have already undergone clinical CT scans, allowing for the evaluation of the registration pipeline without additional radiation exposure; (iii) to investigate the feasibility of MRI to US registration, aiming for a completely radiation-free workflow from preoperative planning to intraoperative navigation; (iv) to significantly expand the custom US dataset by incorporating multi-centre and multi-vendor acquisitions, along with pathological cohorts, ensuring robust algorithmic generalisation; (v) to explore more advanced DL approaches, such as spatiotemporal networks that leverage the continuous video nature of freehand US, aiming to further optimise the segmentation accuracy and temporal coherence established by the

U-Net with VGG16; (vi) to develop non-invasive, real-time bone tracking approaches to provide continuous dynamic guidance during joint manipulation, thereby eliminating the need for bone-anchored reference trackers and ensuring the procedure remains strictly non-invasive; (vii) to integrate patient-specific acoustic models, such as shear-wave elastography, to mathematically compensate for depth estimation errors caused by the standard 1540 m/s speed-of-sound assumption across varying tissue densities; (viii) to integrate biomechanical tissue deformation models or utilise force-sensing probes to compensate for the soft-tissue compression inevitably caused by physical pressure during freehand acquisition, mitigating localised alignment deviations; (ix) to disseminate the achieved breakthroughs in high-impact peer-reviewed journals and establish direct partnerships with orthopaedic surgical centres, bridging the gap between theoretical engineering and practical clinical application.

Bibliography

- [1] Ipek Hacıhaliloglu. 3d ultrasound for orthopedic interventions. In *Advances in Experimental Medicine and Biology*, volume 1093, pages 113–129. Springer, 2018.
- [2] Amelia Alarcón Rivero, Deborah Cabrera Rodríguez, Marlon Carbonell González, and Rosali Santiago Roibal. Clinical outcomes and complications of orthopedic surgical interventions in adult patients. *Salud, Ciencia y Tecnología - Serie de Conferencias*, 4:1771, 12 2025.
- [3] Fatemehzahra Darzi and Thomas Bocklitz. A review of medical image registration for different modalities. *Bioengineering*, 11:786, 8 2024.
- [4] Peter Bröbner, Benjamin Hohlmann, Kristian Welle, and Klaus Radermacher. Ultrasound-based registration for the computer-assisted navigated percutaneous scaphoid fixation. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 70:1064–1072, 9 2023.
- [5] Monan Wang, Donghui Li, Xiping Shang, and Jian Wang. A review of computer-assisted orthopaedic surgery systems. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 16:1–28, 10 2020.
- [6] Chul-Ho Kim, Yong-Beom Park, and Suk Ho Baek. Clinical and radiological outcomes of computer-assisted navigation in primary total knee arthroplasty for patients with extra-articular deformity: Systematic review and meta-analysis. *Clinics in Orthopedic Surgery*, 16:430, 2024.
- [7] Benjamin Hohlmann, Peter Broessner, and Klaus Radermacher. Ultrasound-based 3d bone modelling in computer assisted orthopedic surgery—a review and future challenges. *Computer Assisted Surgery*, 29(1), 2024.
- [8] M. Ciganovic, F. Ozdemir, F. Pean, P. Fuernstahl, C. Tanner, and O. Goksel. Registration of 3d freehand ultrasound to a bone model for orthopedic procedures of the forearm. *International Journal of Computer Assisted Radiology and Surgery*, 13(6):827–836, Jun 2018.

- [9] C. Che, T. S. Mathai, and J. Galeotti. Ultrasound registration: A review. *Methods*, 115:128–143, Feb 2017.
- [10] Xingqi Fan, Qiyang Zhu, Puxun Tu, Leo Joskowicz, and Xiaojun Chen. A review of advances in image-guided orthopedic surgery. *Physics in Medicine & Biology*, 68:02TR01, 1 2023.
- [11] N. Masoumi, C. J. Belasso, M. O. Ahmad, H. Benali, Y. Xiao, and H. Rivaz. Multimodal 3d ultrasound and ct in image-guided spinal surgery: public database and new registration algorithms. *International Journal of Computer Assisted Radiology and Surgery*, 16(4):555–565, Apr 2021.
- [12] Eli Kamara, Zachary P. Berliner, Matthew S. Hepinstall, and H. John Cooper. Pin site complications associated with computer-assisted navigation in hip and knee arthroplasty. *The Journal of Arthroplasty*, 32:2842–2846, 9 2017.
- [13] Ziyang Song. Comparative analysis of mri, ct, and ultrasound imaging modalities: Technological advancements and clinical applications. *Highlights in Science, Engineering and Technology*, 149:82–88, 7 2025.
- [14] Mehrdad Salehi, Raphael Prevost, José-Luis Moctezuma, Nassir Navab, and Wolfgang Wein. *Precise Ultrasound Bone Registration with Learning-Based Segmentation and Speed of Sound Calibration*, pages 682–690. Springer, Cham, 2017.
- [15] Donald A.. Neumann, Elisabeth Roen. Kelly, Craig L.. Kiefer, Kimberly. Martens, and Claudia M.. Grosz. *Kinesiology of the musculoskeletal system : foundations for rehabilitation*. Elsevier, 2017.
- [16] Keith L.. Moore, A. M. R.. Agur, and Arthur F.. Dalley. *Clinically oriented anatomy*. Wolters Kluwer, 2018.
- [17] Breann K. Tisano, Jay P. Shah, and Avneesh Chhabra. Patellar and quadriceps mechanism: Clinical, imaging, and surgical considerations. *Medical Radiology*, Part F812:381 – 406, 2023.
- [18] Jerrold T.. Bushberg. *The essential physics of medical imaging*. Wolters Kluwer Health/Lippincott Williams & Wilkins, 2012.
- [19] J. Boone, J. Brink, S. Edyvean, W. Huda, W. Leitz, C. McCollough, and M. McNitt-Gray. Basics of computed-tomography technology. *Journal of the ICRU*, 12(1):13–23, 2012.
- [20] Preface, executive summary and glossary. *Annals of the ICRP*, 37:9–34, 4 2007.

- [21] A. Berger. Magnetic resonance imaging. *BMJ*, 324(7328):35, 2002.
- [22] R. J. M. Van Geuns, P. A. Wielopolski, H. G. de Bruin, B. J. Rensing, P. M. Van Ooijen, M. Hulshoff, and P. J. de Feyter. Basic principles of magnetic resonance imaging. *Progress in Cardiovascular Diseases*, 42(2):149–156, 1999.
- [23] Michael C.D. Tayler. Nuclear spin relaxation. *Nuclear Magnetic Resonance*, 50:53 – 81, 2024.
- [24] D. L. Longo and T. R. Harrison. *Harrison's Principles of Internal Medicine*. McGraw-Hill, New York, 18th edition, 2012.
- [25] Simon R.. Cherry, James A.. Sorenson, and Michael E.. Phelps. *Physics in nuclear medicine*. Elsevier/Saunders, 2012.
- [26] Miles N.. Wernick and John N.. Aarsvold. *Emission tomography : the fundamentals of PET and SPECT*. Elsevier Academic Press, 2004.
- [27] Paul Suetens. *Fundamentals of Medical Imaging*. Cambridge University Press, Cambridge, UK, 3rd edition, 2017.
- [28] Klaus Strobel and Katrin Stumpe. Pet/ct in musculoskeletal infection. *Seminars in Musculoskeletal Radiology*, 11:353–364, 12 2007.
- [29] Michael J Arnold, Christopher E Jonas, and Rachel E Carter. Point-of-care ultrasonography. *American family physician*, 101:275–285, 3 2020.
- [30] G. Kossoff. Basic physics and imaging characteristics of ultrasound. *World Journal of Surgery*, 24(2):134–142, 2000.
- [31] Pascal Sautot, Philippe Cinquin, Stephane Lavallee, and Jocelyne Troccaz. Computer assisted spine surgery: A first step toward clinical, application in orthopaedics. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1071–1072. IEEE, 10 1992.
- [32] A. Sahovaler, M. J. Daly, H. H. L. Chan, P. Nayak, S. Tzelnick, M. Arkhangorodsky, J. Qiu, R. Weersink, J. C. Irish, P. Ferguson, and J. S. Wunder. Automatic registration and error color maps to improve accuracy for navigated bone tumor surgery using intraoperative cone-beam ct. *JB JS Open Access*, 7(2):e21.00140, May 2022.

- [33] J. D. Chang, I. S. Kim, A. M. Bhardwaj, and R. N. Badami. The evolution of computer-assisted total hip arthroplasty and relevant applications. *Hip Pelvis*, 29(1):1–14, Mar 2017.
- [34] R. Phillips. Phillips, r.: The accuracy of surgical navigation for orthopaedic surgery. *current orthopaedics* 21(3), 180-192. *Current Orthopaedics - CURR ORTHOPAED*, 21:180–192, 06 2007.
- [35] S. W. On, S. W. Cho, S. Y. Park, S. M. Yi, I. Y. Park, S. H. Byun, J. C. Kim, and B. E. Yang. Advancements in computer-assisted orthognathic surgery: A comprehensive review and clinical application in south korea. *Journal of Dentistry*, 146:105061, 2024. Epub 2024 May 9.
- [36] Leo Joskowicz and Eric J. Hazan. *Computer-Aided Orthopedic Surgery: Incremental Shift or Paradigm Change?*, pages 21–30. Springer Singapore, Singapore, 2018.
- [37] Adam J. Karkenny, Justin R. Mendelis, David S. Geller, and Jorge A. Gomez. The role of intraoperative navigation in orthopaedic surgery. *Journal of the American Academy of Orthopaedic Surgeons*, 27(19):e849–e858, 2019.
- [38] Z. Lin, C. Lei, and L. Yang. Modern image-guided surgery: A narrative review of medical image processing and visualization. *Sensors (Basel)*, 23(24):9872, 2023.
- [39] Z. M. C. Baum, Y. Hu, and D. C. Barratt. Meta-learning initializations for interactive medical image registration. *IEEE Transactions on Medical Imaging*, 42(3):823–833, March 2023.
- [40] J. Liu, G. Singh, S. Al'Aref, B. Lee, O. Oleru, J. K. Min, S. Dunham, M. R. Sabuncu, and B. Mosadegh. Image registration in medical robotics and intelligent systems: Fundamentals and applications. *Advanced Intelligent Systems*, 1(8):1900048, 2019.
- [41] X. Jiang, J. Ma, G. Xiao, Z. Shao, and X. Guo. A review of multimodal image matching: Methods and applications. *Information Fusion*, 73:22–71, 2021.
- [42] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang. Deep learning in medical image registration: a review. *Physics in Medicine and Biology*, 65(20):20TR01, October 22 2020.
- [43] J. Kubicek, F. Tomanec, M. Cerny, D. Vilimek, M. Kalova, and D. Oczka. Recent trends, technical concepts and components of computer-assisted orthopedic surgery systems: A comprehensive review. *Sensors (Basel)*, 19(23):5199, Nov 2019.

- [44] Felix Zimmermann, Jochen Franke, Sven Y. Vetter, and Paul A. Grützner. Computer-assisted procedures in orthopedics and trauma surgery—where do we stand?; [computergestützte verfahren in der orthopädie und unfallchirurgie – wo stehen wir?]. *Chirurgie (Germany)*, 94(4):292 – 298, 2023.
- [45] David Hernandez, Roja Garimella, Adam E. M. Eltorai, and Alan H. Daniels. Computer-assisted orthopaedic surgery. *Orthopaedic Surgery*, 9(2):152 – 158, 2017.
- [46] Kevin K. Mathew, Kevin B. Marchand, John M. Tarazi, Hytham S. Salem, William DeGouveia, Joseph O. Ehiorobo, Nipun Sodhi, and Michael A. Mont. Computer-assisted navigation in total knee arthroplasty. *Surgical technology international*, 36:323 – 330, 2020.
- [47] Qais Naziri, Patrick J. Mixa, Daniel P. Murray, Roby Abraham, Bashir A. Zikria, Akhilesh Sastry, and Preetesh D. Patel. Robotic-assisted and computer-navigated unicompartmental knee arthroplasties: A systematic review. *Surgical technology international*, 32:271 – 278, 2018.
- [48] Roberto M. Barbosa, Luís Serrador, Bruno Santos, M. V. Silva, Elena De Momi, and Cristina Santos. Intraoperative bone registration: An implementation in orthopaedic surgery using polaris vica system. In *2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG)*, pages 1–4, 2017.
- [49] Joseph Goerres, Ali Uneri, Michael Jacobson, Ben Ramsay, Thilanka De Silva, Michael Ketcha, Rehan Han, Amir Manbachi, Sebastian Vogt, Gerhard Kleinszig, JP Wolinsky, GM Osgood, and Jeffrey H Siewerdsen. Planning, guidance, and quality assurance of pelvic screw placement using deformable image registration. *Physics in Medicine & Biology*, 62(23):9018–9038, 2017.
- [50] Joseph Goerres, Ali Uneri, Thilanka De Silva, Michael Ketcha, Sansanee Reaungamornrat, Michael Jacobson, Sebastian Vogt, Gerhard Kleinszig, G Maxwell Osgood, JP Wolinsky, and Jeffrey H Siewerdsen. Spinal pedicle screw planning using deformable atlas registration. *Physics in Medicine & Biology*, 62(7):2871–2891, 2017.
- [51] Lambros Vlachopoulos, Felipe Carrillo, Christian Gerber, Gábor Székely, and Philipp Färnstahl. A novel registration-based approach for 3d assessment of posttraumatic distal humeral deformities. *The Journal of Bone and Joint Surgery. American Volume*, 99(23):e127, 2017.
- [52] Z Fanti, F Torres, E Hazan-Lasri, A Gastelum-Strozzi, L Ruiz-Huerta, A Caballero-Ruiz, and FA Cosío. Improved surface-based registration of ct and intraoperative 3d ultrasound of bones. *Journal of Healthcare Engineering*, 2018:2365178, 2018.

- [53] Kai Niu, J Homminga, Viola I Sluiter, A Sprengers, and Nico Verdonschot. Feasibility of a-mode ultrasound based intraoperative registration in computer-aided orthopedic surgery: A simulation and experimental study. *PLoS One*, 13(6):e0199136, 2018.
- [54] Prashant Pandey, Pierre Guy, Alan J Hodgson, and Rafeef Abugharbieh. Fast and automatic bone segmentation and registration of 3d ultrasound to ct for the full pelvic anatomy: a comparative study. *International Journal of Computer Assisted Radiology and Surgery*, 13(10):1515–1524, 2018.
- [55] Carolina Raposo and João P. Barreto. 3d registration of curves and surfaces using local differential information. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9300–9308, Salt Lake City, USA, June 2018. IEEE.
- [56] Thilanka De Silva, Ali Uneri, Xiaojie Zhang, Michael Ketcha, Rehan Han, Neil Sheth, Annethea Martin, Sebastian Vogt, Gerhard Kleinszig, Alex Belzberg, Daniel M Sciubba, and Jeffrey H Siewerdsen. Real-time, image-based slice-to-volume registration for ultrasound-guided spinal intervention. *Physics in Medicine & Biology*, 63(21):215016, 2018.
- [57] Lifeng Wang, Tianmiao Wang, Hongpeng Liu, Lei Hu, Zhonghao Han, Wenyong Liu, Na Guo, Yansong Qi, and Yongsheng Xu. An automated calibration method of ultrasonic probe based on coherent point drift algorithm. *IEEE Access*, 6:8657–8665, 2018.
- [58] H. E. Gueziri, S. Drouin, C. X. B. Yan, and D. L. Collins. Toward real-time rigid registration of intra-operative ultrasound with preoperative ct images for lumbar spinal fusion surgery. *International Journal of Computer Assisted Radiology and Surgery*, 14(11):1933–1943, Nov 2019.
- [59] Houssem-Eddine Gueziri and D. Louis Collins. Fast registration of ct with intra-operative ultrasound images for spine surgery. In *Lecture Notes in Computer Science (Including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11397 of *5th International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging, CSI 2018, held in conjunction with Medical Image Computing and Computer-Assisted Intervention, MICCAI 2018*, pages 29–40. Springer, Granada, September 2019.
- [60] D. Guha, R. Jakubovic, M. K. Leung, H. J. Ginsberg, M. G. Fehlings, T. G. Mainprize, A. Yee, and V. X. D. Yang. Quantification of computational geometric congruence in surface-based registration for spinal intra-operative three-dimensional navigation. *PLoS One*, 14(8):e0207137, Aug 2019.

- [61] H. E. Gueziri, C. X. B. Yan, and D. L. Collins. Open-source software for ultrasound-based guidance in spinal fusion surgery. *Ultrasound in Medicine and Biology*, 46(12):3353–3368, Dec 2020.
- [62] R. Han, A. Uneri, M. Ketcha, R. Vijayan, N. Sheth, P. Wu, P. Vagdargi, S. Vogt, G. Kleinszig, G. M. Osgood, and J. H. Siewerdsen. Multi-body 3d-2d registration for image-guided reduction of pelvic dislocation in orthopaedic trauma surgery. *Physics in Medicine Biology*, 65(13):135009, Jul 2020.
- [63] He Liu and Ferdinando Rodriguez Y Baena. Automatic markerless registration and tracking of the bone for computer-assisted orthopaedic surgery. *IEEE Access*, 8:42010–42020, 2020.
- [64] Zhe Min, Jiaole Wang, and Max Q.-H. Meng. Robust generalized point cloud registration with orientational data based on expectation maximization. *IEEE Transactions on Automation Science and Engineering*, 17(1):207–221, 2020.
- [65] R. Schaffert, J. Wang, P. Fischer, A. Maier, and A. Borsdorf. Robust multi-view 2-d/3-d registration using point-to-plane correspondence model. *IEEE Transactions on Medical Imaging*, 39(1):161–174, Jan 2020. Epub 2019 Jun 13.
- [66] Roman Schaffert, Jian Wang, Peter Fischer, Anja Borsdorf, and Andreas Maier. Learning an attention model for robust 2-d/3-d registration using point-to-plane correspondences. *IEEE Transactions on Medical Imaging*, 39(10):3159–3174, 2020.
- [67] Junhao Zhu, Anwen Zhou, and Qijie Zhao. Registration method of virtual model and real bone based on feature points for surgical navigation. In *2020 12th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 2, pages 83–86, 2020.
- [68] A. Chan, B. Coutts, E. Parent, and E. Lou. Development and evaluation of ct-to-3d ultrasound image registration algorithm in vertebral phantoms for spine surgery. *Annals of Biomedical Engineering*, 49(1):310–321, Jan 2021. Epub 2020 Jun 12.
- [69] R. Han, A. Uneri, R.C. Vijayan, P. Wu, P. Vagdargi, N. Sheth, S. Vogt, G. Kleinszig, G.M. Osgood, and J.H. Siewerdsen. Fracture reduction planning and guidance in orthopaedic trauma surgery via multi-body image registration. *Medical Image Analysis*, 68:101917, Feb 2021. Epub 2020 Nov 30.
- [70] Xue Hu, He Liu, and Ferdinando Rodriguez Y. Baena. Markerless navigation system for orthopaedic knee surgery: A proof of concept study. *IEEE Access*, 9:64708–64718, 2021.

- [71] Zhe Min, Delong Zhu, Jianbang Liu, Hongliang Ren, and Max Q.-H. Meng. Aligning 3d curve with surface using tangent and normal vectors for computer-assisted orthopedic surgery. *IEEE Transactions on Medical Robotics and Bionics*, 3(2):372–383, 2021.
- [72] Minhua Wei, Chaoqun Wang, Haifeng Zhou, Xiao Cheng, Shaolong Kuang, and Lining Sun. A two-stage point cloud registration method for knee joint replacement navigation. In *2021 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*, pages 42–46, 2021.
- [73] K. Abe, M. Oba, N. Kobayashi, S. Higashihira, H. Choe, T. Tezuka, H. Ike, and Y. Inaba. Accuracy of computer navigation-assisted arthroscopic osteochondroplasty for cam-type femoroacetabular impingement using the model-to-image registration method. *American Journal of Sports Medicine*, 50(5):1272–1280, Apr 2022.
- [74] A. Chan, E. Parent, J. Mahood, and E. Lou. 3d ultrasound navigation system for screw insertion in posterior spine surgery: a phantom study. *International Journal of Computer Assisted Radiology and Surgery*, 17(2):271–281, Feb 2022.
- [75] Zhe Min, Jiaole Wang, and Max Q.-H. Meng. Joint rigid registration of multiple generalized point sets with anisotropic positional uncertainties in image-guided surgery. *IEEE Transactions on Automation Science and Engineering*, 19(4):3612–3627, 2022.
- [76] R. R. Naik, A. Hoblidar, S. N. Bhat, N. Ampar, and R. Kundangar. A hybrid 3d–2d image registration framework for pedicle screw trajectory registration between intraoperative x-ray image and preoperative ct image. *Journal of Imaging*, 8(7):185, 2022.
- [77] C. Liu, Y. Song, X. Ma, and T. Sun. Accurate and robust registration method for computer-assisted high tibial osteotomy surgery. *International Journal of Computer Assisted Radiology and Surgery*, 18(2):329–337, Feb 2023.
- [78] Pragyan Shrestha, Chun Xie, Hidehiko Shishido, Yuichi Yoshii, and Itaru Kitahara. X-ray to ct rigid registration using scene coordinate regression. In Hayit Greenspan, Anant Madabhushi, Parvin Mousavi, Septimiu Salcudean, James Duncan, Tanveer Syeda-Mahmood, and Russell Taylor, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, pages 781–790, Cham, 2023. Springer Nature Switzerland.
- [79] X. Xie, M. Zhu, B. He, and J. Xu. Image-guided navigation system for minimally invasive total hip

- arthroplasty (mitha) using an improved position-sensing marker. *International Journal of Computer Assisted Radiology and Surgery*, 18(12):2155–2166, Dec 2023.
- [80] C. Liu, W. Wang, T. Sun, and Y. Song. Soft-tissue sound-speed-aware ultrasound-ct registration method for computer-assisted orthopedic surgery. *Medical Biology Engineering and Computing*, 62(11):3385–3396, Nov 2024.
- [81] Zhuoyuan Li, Xuquan Ji, Chuantao Wang, Wenyong Liu, Feiyu Zhu, and Jiliang Zhai. 2d/3d-mgr: A 2d/3d medical image registration framework based on drr. *IEEE Access*, 12:124365–124374, 2024.
- [82] P. U. Pandey, P. Guy, E. K. Bliven, K. A. Lefavre, and A. J. Hodgson. An integrated system using ultrasound-based registration for surgical navigation of iliosacral screw insertions. *IEEE Transactions on Biomedical Engineering*, pages 120–135, Jul 2024.
- [83] Zhengyan Zhang, Ang Zhang, Jiewen Lai, Hongliang Ren, Rui Song, Yibin Li, Max Q.-H. Meng, and Zhe Min. Ghmm: Learning generative hybrid mixture models for generalized point set registration in computer-assisted orthopedic surgery. *IEEE Transactions on Medical Robotics and Bionics*, 6(3):1285–1295, 2024.
- [84] Seong Kim and Peter Kazanzides. Fiducial-based registration with a touchable region model. *International Journal of Computer Assisted Radiology and Surgery*, 12(2):277–289, 2017.
- [85] Nadia C. Sciberras, Mohammed Almustafa, Benjamin R.K. Smith, David J. Allen, Frederic Picard, and Angela H. Deakin. A randomized controlled trial to compare component placement in navigated total knee arthroplasty using original and streamlined registration processes. *Arthroplasty Today*, 3:111–117, 6 2017.
- [86] Ron Kikinis, Steve D. Pieper, and Kirby G. Vosburgh. *3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support*, pages 277–289. Springer New York, 2014.
- [87] M. Jorge Cardoso, Wenqi Li, Richard Brown, et al. Monai: An open-source framework for deep learning in healthcare, 2022.
- [88] Andras Lasso, Tamas Heffter, Adam Rankin, Csaba Pinter, Tamas Ungi, and Gabor Fichtinger. Plus: Open-source toolkit for ultrasound-guided intervention systems. *IEEE Transactions on Biomedical Engineering*, 61:2527–2537, 10 2014.

- [89] SonoSkills and Hitachi Medical Systems Europe. Ultrasound cases. <https://www.ultrasoundcases.info/cases/musculo-skeletal-joints-and-tendons/knee/>. Accessed on November, 2024.
- [90] Amanda Nelson. Development of an AI/ML-ready knee ultrasound dataset in a population-based cohort, 2023. Accessed on November, 2024.
- [91] Chair for Computer Aided Medical Procedures (CAMP) Technical University of Munich. Open-Access Dataset for the Low-limb-Leg multilabel dataset. <https://www.cs.cit.tum.de/camp/publications/leg-3d-us-dataset/>. Accessed on November, 2024.
- [92] Hridayi Patel and Ilker Hacihaliloglu. Improved automatic bone segmentation using large-scale simulated ultrasound data to segment real ultrasound bone surface data. In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, pages 288–294, 2020.
- [93] Pezhman Foroughi, Emad Boctor, M.J. Swartz, R.H. Taylor, and Gabor Fichtinger. P6d-2 ultrasound bone segmentation using dynamic programming. pages 2523 – 2526, 12 2007.
- [94] Dezhi Sun, Stefano Stramigioli, and Kenan Niu. Hcce-cunet-based multi-class musculoskeletal segmentation for robotic ultrasound system. *IEEE Transactions on Medical Robotics and Bionics*, 7:1728–1738, 11 2025.
- [95] T. Happila, A. Rajendran, Guhan S, Siranjeevi G, Sooraj Krishna S, and Shibu George D. Finger bone segmentation based on deep supervision model using d2fc-dn. In *2025 International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1332–1339. IEEE, 3 2025.
- [96] Ruixuan Li, Ayoob Davoodi, Yuyu Cai, Kenan Niu, Gianni Borghesan, Nicola Cavalcanti, Aidana Masalimova, Fabio Carrillo, Christoph J. Laux, Mazda Farshad, Philipp Fürnstahl, and Emmanuel Vander Poorten. Robot-assisted ultrasound reconstruction for spine surgery: from bench-top to pre-clinical study. *International Journal of Computer Assisted Radiology and Surgery*, 18:1613–1623, 5 2023.
- [97] Kian Hashtrudi-Zaad, Tamas Ungi, Chris Yeung, Zachary Baum, Pavel-Dumitru Cernelev, Anthony N. Hage, Christopher Schlenger, and Gabor Fichtinger. Expert-guided optimization of ultrasound segmentation models for 3d spine imaging. In *2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 680–685. IEEE, 8 2024.

- [98] M. Yuan, B. Jie, R. Han, J. Wang, Y. Zhang, Z. Li, J. Zhu, R. Zhang, and Y. He. Automatic segmentation of the midfacial bone surface from ultrasound images using deep learning methods. *International Journal of Oral and Maxillofacial Surgery*, 54:663–669, 7 2025.
- [99] Luohong Wu, Nicola A. Cavalcanti, Matthias Seibold, Giuseppe Loggia, Lisa Reissner, Jonas Hein, Silvan Beeler, Arnd Viehöfer, Stephan Wirth, Lilian Calvet, and Philipp Färnstahl. Ultrabones100k: A reliable automated labeling method and large-scale dataset for ultrasound-based bone surface extraction. *Computers in Biology and Medicine*, 194:110435, 8 2025.
- [100] Nora Baka, Sieger Leenstra, and Theo van Walsum. Ultrasound aided vertebral level localization for lumbar surgery. *IEEE Transactions on Medical Imaging*, 36:2138–2147, 10 2017.
- [101] Jasper M. van der Zee, Aimon M. Rahman, Kevin Klein Gunnewiek, Marijn A. J. Hiep, Matthijs Fitski, Ilker Hacihaliloglu, Ahmed Z. Alsinan, Vishal M. Patel, Annemieke S. Littooi, and Alida F. W. van der Steeg. Evaluation of unext for automatic bone surface segmentation on ultrasound imaging in image-guided pediatric surgery. *Bioengineering*, 12:1008, 9 2025.
- [102] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, pages 234–241. In: Navab N, Hornegger J, Wells WM, Frangi AF (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, Cham, 2015.
- [103] Ahmed Z. Alsinan, Vishal M. Patel, and Ilker Hacihaliloglu. Automatic segmentation of bone surfaces from ultrasound using a filter-layer-guided cnn. *International Journal of Computer Assisted Radiology and Surgery*, 14:775–783, 5 2019.
- [104] Shuying Liu and Weihong Deng. Very deep convolutional neural network based image classification using small training sample size. In *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 730–734. IEEE, 11 2015.
- [105] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. IEEE, 6 2009.
- [106] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, 2021.

- [107] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. Segformer: simple and efficient design for semantic segmentation with transformers. In *Proceedings of the 35th International Conference on Neural Information Processing Systems, NIPS '21*, Red Hook, NY, USA, 2021. Curran Associates Inc.
- [108] Godau P et al (2024) Maier-Hein L, Reinke A. Metrics reloaded: recommendations for image analysis validation. *Nature Methods*, 21:195–212, 2 2024.
- [109] Shervin Minaee, Yuri Y. Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2021.
- [110] Pranav Rajpurkar, Emma Chen, Oishi Banerjee, and Eric J. Topol. Ai in health and medicine. *Nature Medicine*, 28:31–38, 1 2022.
- [111] Puyang Wang, Michael Vives, Vishal M. Patel, and Ilker Hacihaliloglu. Robust real-time bone surfaces segmentation from ultrasound using a local phase tensor-guided cnn. *International Journal of Computer Assisted Radiology and Surgery*, 15:1127–1135, 7 2020.
- [112] Bolun Zeng, Li Chen, Yuanyi Zheng, and Xiaojun Chen. Adaptive multi-dimensional weighted network with category-aware contrastive learning for fine-grained hand bone segmentation. *IEEE Journal of Biomedical and Health Informatics*, 28:3985–3996, 7 2024.
- [113] Maha Ezzelarab, Yuyue Zhou, and Abhilash Rakkunedeth Hareendranathan. Automated wrist ultrasound image bone enhancement and segmentation using deep learning. In *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–4. IEEE, 7 2025.
- [114] Tamas Ungi, Hastings Greer, Kyle R. Sunderland, Victoria Wu, Zachary M. C. Baum, Christopher Schlenger, Matthew Oetgen, Kevin Cleary, Stephen R. Aylward, and Gabor Fichtinger. Automatic spine ultrasound segmentation for scoliosis visualization and measurement. *IEEE Transactions on Biomedical Engineering*, 67:3234–3241, 11 2020.
- [115] Benjamin Hers, Maria Bonta, Siyi Du, Kishore Mulpuri, Emily K. Schaeffer, Antony J. Hodgson, and Rafeef Garbi. Segformer3d: Improving the robustness of deep learning model-based image segmentation in ultrasound volumes of the pediatric hip. *Ultrasound in Medicine & Biology*, 51:751–758, 5 2025.

- [116] PLUS Toolkit Team. Plus applications user manual: Temporal calibration algorithm. <http://perk-software.cs.queensu.ca/plus/doc/nightly/user/AlgorithmTemporalCalibration.html>, 2025. Accessed on January, 2025.
- [117] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Open3D: A modern library for 3D data processing. *arXiv:1801.09847*, 2018.
- [118] Shuifa Sun, Kun Xia, Ning Wei, Yongheng Tang, Yaobin Zou, and Yirong Wu. 3d point cloud coarse registration algorithm based on center of gravity and centroid transformation. In *2022 Euro-Asia Conference on Frontiers of Computer Science and Information Technology (FCSIT)*, pages 46–50. IEEE, 12 2022.
- [119] Paul J. Besl and Neil D. McKay. Method for registration of 3-d shapes. pages 586–606, 4 1992.
- [120] Jon Louis Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18:509–517, 9 1975.
- [121] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:2262–2275, 12 2010.
- [122] J.M. Fitzpatrick, J.B. West, and C.R. Maurer. Predicting error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, 17:694–702, 1998.
- [123] David R. Roberts, Volker Bahn, Simone Ciuti, Mark S. Boyce, Jane Elith, Gurutzeta Guillera-Aroita, Severin Hauenstein, José J. Lahoz-Monfort, Boris Schröder, Wilfried Thuiller, David I. Warton, Brendan A. Wintle, Florian Hartig, and Carsten F. Dormann. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40:913–929, 8 2017.
- [124] Guillermo Carbajal, Andras Lasso, Álvaro Gómez, and Gabor Fichtinger. Improving n-wire phantom-based freehand ultrasound calibration. *International Journal of Computer Assisted Radiology and Surgery*, 8:1063–1072, 11 2013.

Appendix A

Design of the Probe Reference

This appendix details the computer-aided design (CAD) of the custom reference attached to the ultrasound (US) probe. Figure A.1 illustrates the complete technical drawings, developed in SolidWorks, providing the exact geometric dimensions and cross-sectional profiles required for the manufacturing and precise spatial configuration of the retroreflective markers used by the OptiTrack optical tracking system.

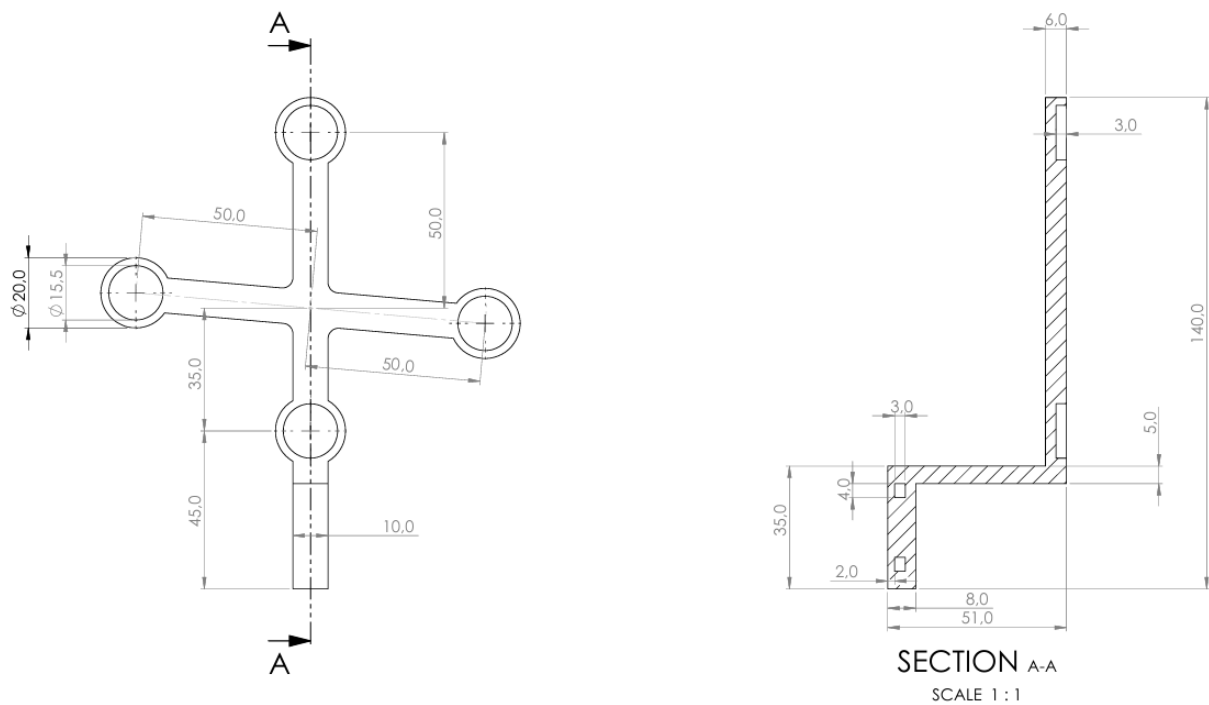


Figure A.1: Detailed technical drawing of the custom-designed US probe reference. The schematic includes the primary top-down view and the corresponding structural cross-section engineered for the integration of OptiTrack markers. All dimensional units are provided in millimetres.

Appendix B

Design of the Calibration Phantom Reference

This appendix details the CAD of the custom reference developed for the spatial calibration phantom. Figure B.1 illustrates the precise structural geometry, exact metric dimensions, and cross-sectional profiles required to manufacture this component. This specific geometry ensures stable integration of the OptiTrack retroreflective markers.

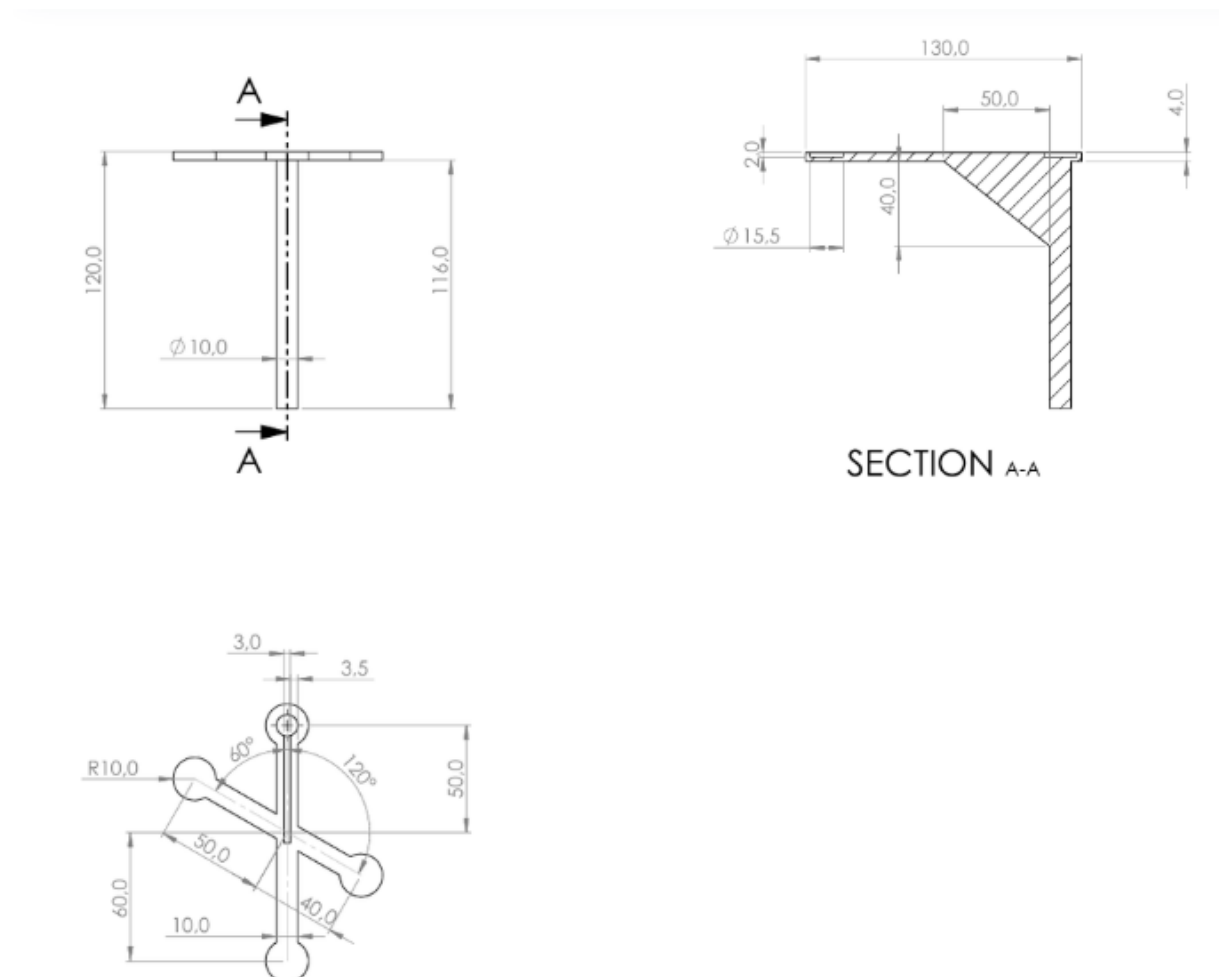


Figure B.1: Detailed technical drawing of the custom-designed reference frame for the calibration phantom. The schematic includes multiple structural views, such as the cross-section, explicitly engineered to accommodate the OptiTrack markers. All dimensional units are provided in millimetres.

Appendix C

Detailed Calibration Measurements

This appendix provides the raw data acquired during the temporal and spatial calibration procedures. The tables below detail the individual measurements that support the statistical analysis in Chapter 6.

C.1 Temporal Calibration

Table C.1 presents the individual latency measurements recorded during the temporal calibration. These values quantify the time delay (in milliseconds) between the ultrasound image stream and the optical tracking data acquisition, a critical parameter for ensuring accurate real-time synchronisation.

Table C.1: Individual latency measurements (ms) recorded during the temporal calibration procedure

Measurement	Latency (ms)
1	190.6
2	177.0
3	203.9
4	179.5
5	198.9
6	175.9
7	163.2
8	186.2
9	166.5
10	185.9
11	185.1
12	193.4
13	226.0
14	195.6
15	181.9

C.2 Spatial Calibration

This section presents the raw spatial error metrics derived from the spatial calibration procedures. Table C.2 details the individual error measurements recorded during the standard dynamic calibration, reflecting the system's accuracy under continuous probe motion. For comparative analysis, Table C.3 provides the baseline error measurements acquired during a static test (with the probe securely immobilised). This static baseline is crucial for isolating the inherent optical tracking noise from motion-induced calibration artefacts.

Table C.2: Individual spatial error measurements (mm) obtained during the dynamic calibration process

Measurement	Error (mm)
1	1.02
2	1.05
3	1.03
4	1.20
5	0.64
6	0.71
7	1.22
8	1.01
9	1.12
10	1.05
11	1.41
12	1.05
13	1.61
14	0.94

Table C.3: Baseline spatial error measurements (mm) acquired during the static calibration test

Measurement	Error (mm)
1	0.248279
2	0.329151
3	0.631211
4	0.381159
5	0.749002
6	0.643654