CHALMERS
UNIVERSITY OF TECHNOLOGY

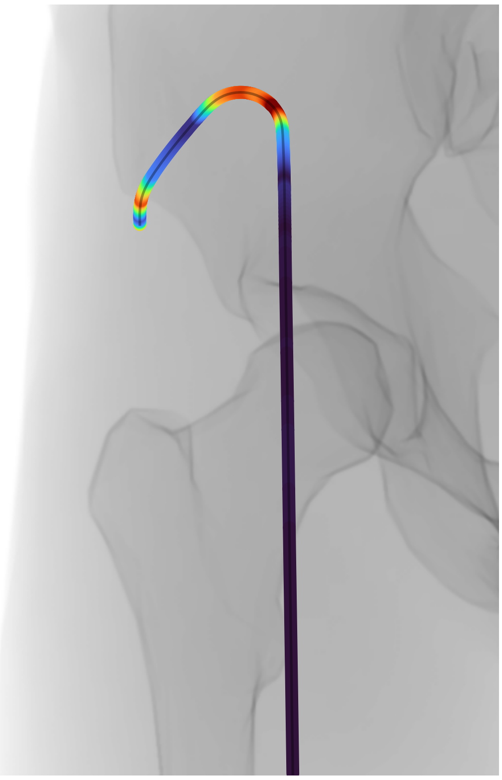


Image Processing Framework for Accuracy Measurements in Real-Time Endovascular Simulation

Master's thesis in Biomedical Engineering

ADAM FRANZÉN
LINNÉA SWANMARK

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Master's thesis 2026

Image Processing Framework for Accuracy Measurements in Real-Time Endovascular Simulation

Comparing Virtual Simulation with Physical Surgical Procedures

ADAM FRANZÉN
LINNÉA SWANMARK

Department of Electrical Engineering
Division of Signal processing and Biomedical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2026

Image Processing Framework for Accuracy Measurements in Real-Time Endovascular Simulation
Comparing Virtual Simulation with Physical Surgical Procedures
ADAM FRANZÉN
LINNÉA SWANMARK

Supervisor: Anders Ekström, Mentice
Supervisor: David Ohlin Saletti, Mentice
Supervisor: Fredrik Sjunnesson, Mentice
Supervisor: Ida Häggström, Department of Electrical Engineering
Examiner: Ida Häggström, Department of Electrical Engineering

Master's Thesis 2026
Department of Electrical Engineering
Division of Signal processing and Biomedical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Curvature heatmap overlay of a catheter tip configuration against a simulated fluoroscopic background..

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2026

Image Processing Framework for Accuracy Measurements in Real-Time Endovascular Simulation
ADAM FRANZÉN
LINNÉA SWANMARK
Department of Electrical Engineering
Chalmers University of Technology

Abstract

Training surgeons to navigate catheters and guidewires through blood vessels is challenging, and simulators are increasingly used to practice these procedures safely. However, no standard method currently exists to objectively measure how realistically a simulator reproduces the behavior of real surgical instruments. This thesis presents an image-based framework for the quantitative comparison of physical and simulated catheter configurations, combining camera-based image acquisition, centerline extraction, geometric alignment, and shape comparison metrics.

Physical experiments were performed by imaging catheters and guidewires both on flat surfaces and inside a vascular phantom (Physical SIM). The same instrument configurations were then reproduced in the Mentice simulator (VIST), allowing a direct comparison between the physical and virtual setups. The framework was further extended with an automated optimization module that searches for simulator stiffness parameters that best reproduce the observed physical catheter behavior using Bayesian optimization.

Validation experiments demonstrated high accuracy in catheter length estimation, achieving a mean absolute error of 0.117 cm and a coefficient of determination exceeding $R^2 > 0.99$. It also remained consistent regardless of the catheter shape. The alignment procedure showed that the extracted centerlines converge to a close geometric correspondence after registration. Optimization experiments further showed that simulator stiffness parameters could be tuned to reproduce physical catheter configurations with high geometric similarity, although discrepancies remained for certain catheter-guidewire combinations. A direct comparison between the physical and virtual environments was additionally limited by fundamental differences in anatomy representation.

The results demonstrate that the proposed framework enables quantitative and repeatable evaluation of endovascular simulator realism. Among the evaluated metrics, RMS error and curvature analysis were found to be the most informative: RMS error for localizing spatial deviations along the catheter shaft, and curvature analysis for capturing mechanical differences in instrument bending behavior. The framework provides a foundation for future validation and calibration of catheter mechanics in anatomically realistic simulation environments.

Keywords: endovascular, simulation, catheter, guidewire, validation, centerline, alignment, optimization, curvature, phantom.

Acknowledgements

We express our sincere gratitude to our supervisors at Mentice, Anders Ekström, David Ohlin Saletti, and Fredrik Sjunnesson, for their invaluable guidance, expertise, and continuous support throughout this project. Thank you for sharing your knowledge, providing constructive feedback, and encouraging us during the development of this thesis. Your insights and commitment have been essential to its completion.

We would also like to thank Ellinor Claesson at Mentice for her enthusiasm, encouragement, and valuable discussions. Her willingness to brainstorm ideas and share her insights contributed greatly to the development of this thesis.

We are grateful to Mentice for giving us the opportunity to carry out this thesis project. We greatly appreciated the welcoming and collaborative work environment as well as the opportunity to gain practical experience and insight into the development of medical simulation. Working alongside skilled professionals has been inspiring and educational and has contributed significantly to our professional growth.

Finally, we would like to thank our supervisor and examiner at Chalmers University of Technology, Ida Häggström, for her valuable feedback, guidance, and support throughout the project.

Adam Franzén & Linnéa Swanmark, Gothenburg, June 2026

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

AD	Average distance
CDF	Cumulative distribution function
DSC	Dice similarity coefficient
HD	Hausdorff distance
HFV	Hessian-based Frangi vesselness
MAE	Mean Absolute Error
MHD	Modified Hausdorff distance
RMS	Root Mean Square
SIFT	Scale-Invariant Feature Transform
SVD	Singular Value Decomposition

Nomenclature

Below is the nomenclature of symbols and variables used throughout this thesis.

Points and Coordinates

$\mathbf{p}_w$	3D point in the world coordinate frame, $(x_w, y_w, z_w)^\top$
$\mathbf{p}_c$	3D point in the camera coordinate frame, $(x_c, y_c, z_c)^\top$
$\tilde{\mathbf{p}}_w$	Homogeneous world point, $(x_w, y_w, z_w, 1)^\top$
$\tilde{\mathbf{p}}_c$	Homogeneous camera point, $(x_c, y_c, z_c, 1)^\top$
$\tilde{\mathbf{p}}_i$	Homogeneous in-plane calibration target point, $(x_{w,i}, y_{w,i}, 1)^\top$
$\mathbf{x}_s$	2D pixel coordinate, $(x_s, y_s)^\top$
$\tilde{\mathbf{x}}_s$	Homogeneous pixel coordinate, $(x_s, y_s, 1)^\top$
(x_c, y_c)	Normalized camera coordinates (pre-distortion)
$(\hat{x}_c, \hat{y}_c)$	Radially corrected normalized camera coordinates
r_c^2	Squared radial distance from the principal point, $x_c^2 + y_c^2$
$(\Delta x_c, \Delta y_c)$	Tangential distortion displacement

Camera Parameters

$\mathbf{K}$	3×3 intrinsic (calibration) matrix
f	Focal length (isotropic case)
$f_x,\ f_y$	Focal lengths in horizontal and vertical pixel units
(c_x, c_y)	Principal point
s	Skew coefficient
$\mathbf{R}$	3×3 rotation matrix, $\mathbf{R} \in SO(3)$
$\mathbf{t}$	Translation vector, $\mathbf{t} \in \mathbb{R}^3$
$\mathbf{P}$	3×4 camera matrix, $\mathbf{K}[\mathbf{R} \mid \mathbf{t}]$

Distortion Parameters

κ_1, κ_2	Radial distortion coefficients
p_1, p_2	Tangential distortion coefficients
$\mathbf{d}$	Distortion parameter vector, $(\kappa_1, \kappa_2, p_1, p_2)^\top$

Calibration and Error Metrics

$\mathbf{H}_j$	3×3 homography matrix for image j
$\mathbf{r}_1$, $\mathbf{r}_2$	First and second columns of rotation matrix $\mathbf{R}$
$\tilde{\mathbf{x}}_{s,ij}$	Detected pixel coordinate of corner i in image j
$\hat{\mathbf{x}}_{s,ij}$	Projected (predicted) pixel coordinate of corner i in image j
ϵ_{RMS}	RMS reprojection error
M	Number of calibration images
N_j	Number of detected corners in image j
N	Total number of corner observations, $\sum_{j=1}^{M} N_j$

Catheter Kinematics

$\mathbf{r}(s,t)$	Catheter centerline space curve, parameterized by arc length and time
s	Arc length parameter
$\mathbf{R}(s,t)$	Orthonormal frame defining cross-section orientation, $\mathbf{R} \in SO(3)$
$\mathbf{d}_3$	Local tangent direction, $[0,0,1]^\top$
$\boldsymbol{\kappa}$	Curvature-torsion vector
$\hat{(\cdot)}$	Skew-symmetric matrix operator
$\mathbf{n}$	Internal force vector
$\mathbf{m}$	Internal moment vector
$\mathbf{f}$	External force density
$\mathbf{l}$	External moment density
ρ	Material density
A	Cross-sectional area
I	Second moment of area
$\mathbf{K}$	Stiffness matrix

$\boldsymbol{\kappa}_0$	Intrinsic (rest) curvature
E	Young's modulus

Geometric Analysis

$\mathbf{C}(u)$	B-spline centerline curve
u	Normalized arc-length parameter, $u \in [0, 1]$
p	B-spline degree
$\mathbf{P}_i$	B-spline control points
$N_{i,p}(u)$	B-spline basis functions
$\mathbf{U}$	Knot vector, $\{u_0, u_1, \ldots, u_m\}$

Procrustes Alignment

c	Procrustes scaling factor
$\mathbf{x}_i$, $\mathbf{y}_i$	Corresponding source and target point sets
e^2	Mean squared alignment error

Accuracy Metrics

x_i^{sim}	i-th centerline point of the simulated instrument
x_i^{real}	i-th centerline point of the physical reference instrument
N	Number of centerline points
$d(p, G)$	Minimum Euclidean distance from point p to point set G
G_{sim}	Simulated instrument centerline or segmentation
G_{real}	Physical reference instrument centerline or segmentation

Contents

List of Figures

List of Tables

1

Introduction

This chapter introduces the clinical and technical background of endovascular simulation, establishes the need for objective accuracy evaluation, and defines the research problem, aim, research questions, and scope of the thesis.

1.1 Introduction to endovascular interventions

Image-guided endovascular interventions are a form of minimally invasive surgery in which medical instruments, such as catheters and guidewires, are navigated through the vascular system using real-time imaging, most often fluoroscopy or ultrasound [1]. Compared to open surgery, these procedures generally result in reduced patient trauma, shorter recovery times, and improved clinical outcomes, which has contributed to their increased use in modern healthcare [2].

Despite these advantages, endovascular procedures are technically demanding. Physicians must accurately control long and flexible instruments within complex vascular geometries while interpreting two-dimensional images of three-dimensional anatomy [1, 3]. Small errors in instrument positioning or timing can lead to complications, putting high demands on technical skill and experience [1].

1.2 Endovascular simulation

Due to the complexity of endovascular procedures, simulation-based training has become an important element of medical training [4, 5]. High-fidelity endovascular simulators provide a safe environment for physicians and students to practice procedures and explore complex scenarios without risk to patients.

Compared to traditional training methods, simulation enables standardized evaluation and objective performance assessment. Previous studies have shown that simulator-based training can improve technical performance and clinician confidence, contributing to improved patient safety [5, 6, 7].

A key concept in medical simulation is fidelity, which describes how closely a simulator represents real clinical conditions and how well skills learned in simulation transfer to real procedures [8]. In endovascular simulation, achieving high fidelity is challenging, as small inaccuracies in geometry, or instrument behavior can affect perceived realism and training effectiveness [9].

1.3 Motivation and Problem Statement

There is currently no standardized methodology for objectively measuring the accuracy and realism of endovascular simulators [10]. Assessment is often predominantly based on subjective expert feedback [9], in which clinicians judge whether the simulation appears realistic based on their experience. Such assessments are difficult to standardize and reproduce, and may vary between observers.

Objective accuracy measures, such as spatial deviation, temporal delay, and geometric similarity, are well established in other technical fields, particularly in computer vision and robotics [11], but are not widely applied in endovascular simulation. The absence of standardized accuracy metrics makes it difficult to compare simulation systems and highlights the need for objective, quantitative evaluation criteria.

1.4 Industrial context

This thesis is conducted in collaboration with Mentice AB, a Gothenburg-based company developing high-fidelity endovascular simulators. As simulation systems become increasingly advanced, there is a growing need for objective and data-driven validation methods. Although expert evaluations remain important, they are often time-consuming and difficult to integrate.

From an industrial perspective, an image-based accuracy measurement framework would enable repeatable comparison between virtual simulation environments and physical reference systems. Such a framework could support continuous validation and assist Mentice in fine-tuning their simulation models.

1.5 Aim and Research Questions

The aim of this thesis is to develop an image-based framework for measuring the technical accuracy of endovascular simulation systems. The framework enables quantitative comparison between Mentice simulation and physical reference data.

A central aim of the project is to propose and evaluate specific accuracy metrics to determine their suitability for this domain. The project aims to move beyond subjective evaluation by validating data-driven metrics that are interpretable and suitable for integration into simulator development workflows.

The long-term vision of this work is to contribute to the development of standardized and objective accuracy evaluation methods for endovascular simulators. In an industrial context, the proposed framework may also serve as a foundation for continuous validation and testing as simulation systems evolve over time.

To achieve this, the thesis seeks to answer the following research questions:

RQ1 How accurately do virtual endovascular simulations reproduce the spatial behavior of instruments compared to physical reference systems in two-dimensional image space?
RQ2 Which image-based accuracy metrics are most suitable for capturing relevant deviations in catheter or guidewire navigation?
RQ3 How consistent are the proposed accuracy measurements across multiple experiments?
RQ4 What practical limitations of uncertainty arise when comparing virtual and physical endovascular simulations using image-based methods?

1.6 Scope and Limitations

The scope is intentionally limited to ensure feasibility within the time frame of a master's thesis and to maintain a clear focus on quantitative accuracy assessment.

A specific part of the scope is the evaluation of the accuracy metrics themselves. The thesis does not assume a single standard of truth, but rather investigates the suitability of different mathematical definitions for endovascular validation.

The framework is evaluated through offline analysis of images from virtual simulation environments and controlled physical experiments using phantom models. Real-time implementation in live clinical imaging systems is considered outside the scope of this work; the framework is intended specifically for post-processing validation and benchmarking rather than for direct clinical application.

Accuracy analysis is performed in two-dimensional image space, reflecting the projection-based nature of imaging used in clinical settings. Comprehensive three-dimensional reconstruction and full 3D validation are not required outcomes of this thesis.

1.7 Contributions

This thesis contributes to the development of a more objective and quantitative evaluation framework for endovascular simulation systems. By introducing and systematically applying measurable image-based accuracy metrics, the work reduces reliance on subjective expert assessment and improves the reproducibility, transparency, and comparability of simulation performance.

In addition, the thesis provides a structured methodology for extracting and parameterizing instrument trajectories from both virtual and physical data, and evaluates the suitability of different metrics for capturing clinically relevant deviations. The proposed framework is designed to support continuous validation and benchmarking within an industrial simulation environment.

2

Theory

This chapter presents the theoretical background required for the methods developed in this work. It covers medical imaging principles, catheter kinematics, and fundamental concepts in computer vision and geometric analysis.

The chapter introduces X-ray fluoroscopy as the primary imaging modality, followed by models describing catheter behavior. Then it outlines the principles of image formation and camera calibration, as well as techniques for segmentation and centerline extraction. Finally, methods for geometric representation, alignment, and accuracy evaluation are presented.

2.1 Medical Imaging and Catheters

Medical imaging and catheter kinematics constitute two fundamental components of endovascular simulation, as they underpin both the visualization of the anatomy and the modeling of instrument behavior.

2.1.1 X-ray Fluoroscopy

X-ray fluoroscopy is an imaging modality commonly used during endovascular procedures [1]. Despite concerns regarding both acute and long-term radiation exposure, it provides real-time imaging that is essential for procedural guidance. The system consists of an X-ray source and detector that acquire two-dimensional projection images at high frame rates (typically around 30 Hz).

Although soft tissue contrast is limited, the use of contrast agents and the inherent visibility of surgical instruments allow clinicians to identify relevant structures. However, since fluoroscopy produces 2D projection images of inherently 3D anatomy, accurate interpretation of spatial relationships and instrument positions can be challenging. Additionally, registering these 2D images with pre-acquired 3D datasets is a complex task.

To address these limitations, some procedures employ biplane fluoroscopy systems, where two X-ray sources and detectors are used simultaneously to provide improved spatial understanding of anatomy and instrument positioning.

2.1.2 Catheter Kinematics

Catheter kinematics can be described in terms of the translation, rotation, and deformation of the instrument as it navigates through the vascular system. These motions are controlled from the proximal end and are influenced by interactions with the vessel walls, including contact and friction effects [3].

To model these behaviors more accurately, the catheter can be represented as a continuous, deformable structure. In such approaches, the instrument is treated as a flexible rod subject to external forces and boundary conditions [12, 13]. The catheter is often discretized into multiple segments, each describing the local position and orientation, while the overall deformation is governed by the balance of internal elastic forces and externally applied loads.

A common and physically grounded approach is to model the catheter using beam theory or Cosserat rod formulations, where bending, torsion, and stretching are explicitly considered. These models capture the nonlinear behavior of the catheter as it interacts with the surrounding anatomy and provide a continuous description of its shape and motion.

In the Cosserat rod framework, the catheter centerline is described by a space curve $\mathbf{r}(s,t)$, parameterized by arc length s and time t, together with an orthonormal frame $\mathbf{R}(s,t) \in SO(3)$ that defines the orientation of each cross-section [12, 13]. The kinematics of the rod are given by:

$$\frac{\partial \mathbf{r}}{\partial s} = \mathbf{R}\mathbf{d}_3, \quad \frac{\partial \mathbf{R}}{\partial s} = \mathbf{R}\hat{\boldsymbol{\kappa}}, \tag{2.1}$$

where $\mathbf{d}_3 = [0,0,1]^T$ represents the local tangent direction and $\boldsymbol{\kappa}$ is the curvature-torsion vector. The operator $\hat{(\cdot)}$ denotes the skew-symmetric matrix associated with a vector.

The dynamics of the catheter are governed by the balance of linear and angular momentum:

$$\rho A \frac{\partial^2 \mathbf{r}}{\partial t^2} = \frac{\partial \mathbf{n}}{\partial s} + \mathbf{f}, \tag{2.2}$$

$$\rho I \frac{\partial \boldsymbol{\omega}}{\partial t} = \frac{\partial \mathbf{m}}{\partial s} + \mathbf{r}_s \times \mathbf{n} + \mathbf{l}, \tag{2.3}$$

where $\mathbf{n}$ and $\mathbf{m}$ are the internal force and moment vectors, $\mathbf{f}$ and $\mathbf{l}$ are external force and moment densities, ρ is the density, A is the cross-sectional area, and I is the second moment of area.

The material behavior of the catheter is described through constitutive relations:

$$\mathbf{m} = \mathbf{K}(\boldsymbol{\kappa} - \boldsymbol{\kappa}_0), \quad \mathbf{n} = EA(\mathbf{r}_s - \mathbf{d}_3), \tag{2.4}$$

where $\mathbf{K}$ is the stiffness matrix, $\boldsymbol{\kappa}_0$ is the intrinsic curvature, and E is the Young's modulus.

Interactions with vessel walls introduce additional complexity, as contact forces, friction, and geometric constraints significantly influence catheter motion [12, 13]. These effects can lead to phenomena such as buckling, looping, and sudden changes in direction, which are important to capture in realistic simulations. Overall, the Cosserat rod formulation provides a powerful framework for modeling catheter deformation, enabling accurate representation of its mechanical behavior in endovascular procedures.

2.2 Image Formation

The process in which a three-dimensional scene is mapped onto a two-dimensional image plane is governed by the geometry of perspective projection. The most widely used idealization is the pinhole camera model, see Figure 2.1 for an illustration, in which all rays of light from the scene pass through a single point before being recorded on the image plane [14]. A 3D scene point $\mathbf{p}_c = (x_c, y_c, z_c)^\top$, expressed in the camera coordinate frame, is projected onto the sensor plane through the central projection

$$\bar{\mathbf{x}} = \mathcal{P}_z(\mathbf{p}_c) = \begin{pmatrix} x_c/z_c \\ y_c/z_c \\ 1 \end{pmatrix}, \tag{2.5}$$

which divides the lateral coordinates by the depth z_c [14]. For a camera with a single focal length f and the principal point at the image origin, the corresponding sensor coordinates (x_s, y_s) are

$$x_s = f\,\frac{x_c}{z_c}, \qquad y_s = f\,\frac{y_c}{z_c}. \tag{2.6}$$

The nonlinearity introduced by the perspective division is removed by expressing the 3D point in homogeneous form $\tilde{\mathbf{p}}_c = (x_c, y_c, z_c, 1)^\top$, making projection a linear matrix operation [14].

2.3 Camera Calibration

Camera calibration is the process of estimating the parameters that relate 3D world coordinates to 2D pixel coordinates in the image. These parameters are divided into intrinsic parameters, which are internal to the camera, and extrinsic parameters, which describe the pose of the camera relative to a world reference frame [14]. An accurately calibrated camera is a prerequisite for any quantitative image-based measurement, as it enables metric quantities such as distances and angles to be inferred from pixel observations.

2.3.1 Camera Model

The projection from a 3D world point $\mathbf{p}_w$ to a homogeneous sensor coordinate $\tilde{\mathbf{x}}_s = (x_s, y_s, 1)^\top$ is described by the camera matrix [14]

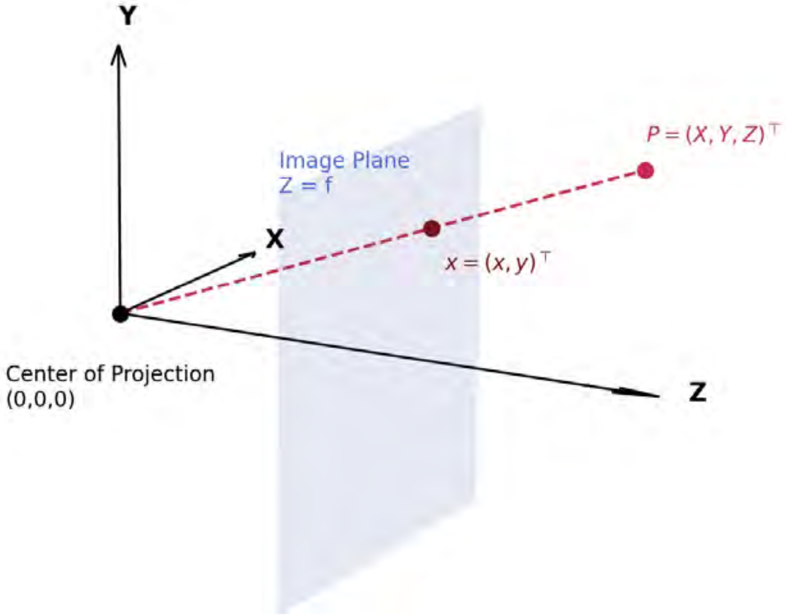


Figure 2.1: Geometry of the pinhole camera model. A 3D camera-centered point $\mathbf{P}$ is projected through the optical center onto the image plane located at distance f along the Z-axis.

$$\tilde{\mathbf{x}}_s = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{p}_w = \mathbf{P}\,\mathbf{p}_w, \tag{2.7}$$

where $\mathbf{R} \in SO(3)$ and $\mathbf{t} \in \mathbb{R}^3$ are the extrinsic parameters encoding the camera orientation and position with respect to the world, and $\mathbf{P} = \mathbf{K}[\mathbf{R} \mid \mathbf{t}]$ is the 3×4 camera matrix [14]. The matrix $\mathbf{K}$ is the calibration matrix, which encodes the camera intrinsics as [14]

$$\mathbf{K} = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.8}$$

where f_x and f_y are the focal lengths expressed in horizontal and vertical pixel units, (c_x, c_y) is the principal point, and s is a skew coefficient that accounts for non-perpendicularity of the pixel axes [14]. For most modern digital cameras the skew is negligibly small and is set to zero in practice. For a camera with $f_x = f_y = f$ and zero skew, $\mathbf{K}$ reduces to three degrees of freedom (f, c_x, c_y), while the extrinsic parameters $(\mathbf{R}, \mathbf{t})$ contribute six additional degrees of freedom.

2.3.2 Lens Distortion

Real lenses deviate from the ideal pinhole model due to imperfections in their manufacture and design. These deviations, collectively termed lens distortion, introduce systematic errors in image-based measurements and must therefore be corrected prior to analysis [14].

The dominant form is radial distortion, which displaces image points toward or away from the principal point by an amount that increases with radial distance. Lenses that pull points toward the centre exhibit barrel distortion, while those that push them outward exhibit pincushion distortion, as illustrated in Figure 2.2 [15].

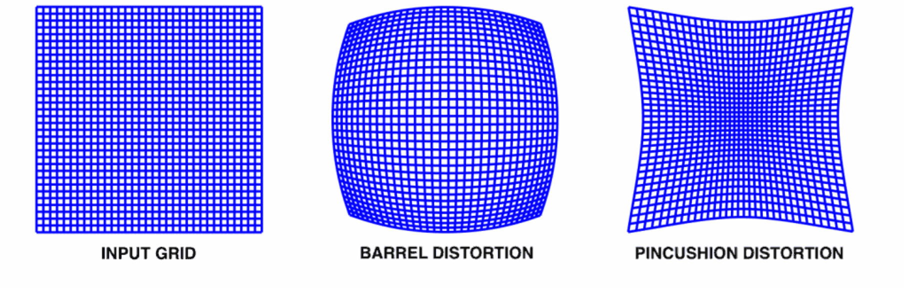


Figure 2.2: The effect of barrel distortion and pincushion distortion on a square grid.

If we let (x_c, y_c) denote the coordinates obtained after perspective division but before scaling by f and shifting by (c_x, c_y) [14]:

$$x_c = \frac{\mathbf{r}_x \cdot \mathbf{p} + t_x}{\mathbf{r}_z \cdot \mathbf{p} + t_z}, \qquad y_c = \frac{\mathbf{r}_y \cdot \mathbf{p} + t_y}{\mathbf{r}_z \cdot \mathbf{p} + t_z}, \tag{2.9}$$

where $\mathbf{r}_x$, $\mathbf{r}_y$, $\mathbf{r}_z$ are the three rows of $\mathbf{R}$ and $\mathbf{p}$ is the 3D world point [14]. The radial distortion model is then [14]

$$\hat{x}_c = x_c\Big(1 + \kappa_1 r_c^2 + \kappa_2 r_c^4\Big), \qquad \hat{y}_c = y_c\Big(1 + \kappa_1 r_c^2 + \kappa_2 r_c^4\Big), \tag{2.10}$$

where $r_c^2 = x_c^2 + y_c^2$ and κ_1, κ_2 are the radial distortion coefficients [14]. After the radial distortion step, the final pixel coordinates are recovered as [14]

$$x_s = f\hat{x}_c + c_x, \qquad y_s = f\hat{y}_c + c_y. \tag{2.11}$$

In addition to radial distortion, a small tangential distortion (see Figure 2.3 [15]) arises from slight misalignment between the lens and sensor planes [14]. The explicit correction terms, modeled by two additional coefficients p_1 and p_2, are given by [16, 17, 18]

$$\Delta x_c = 2p_1 x_c y_c + p_2\Big(r_c^2 + 2x_c^2\Big), \qquad \Delta y_c = p_1\Big(r_c^2 + 2y_c^2\Big) + 2p_2 x_c y_c, \tag{2.12}$$

and equation (2.11) generalizes to

$$x_s = f\left(\hat{x}_c + \Delta x_c\right) + c_x, \qquad y_s = f\left(\hat{y}_c + \Delta y_c\right) + c_y. \tag{2.13}$$

The complete distortion parameter vector $\mathbf{d} = (\kappa_1, \kappa_2, p_1, p_2)^\top$ is estimated jointly with the intrinsic parameters during calibration, ensuring consistency with the OpenCV implementation used in practice [17, 18].

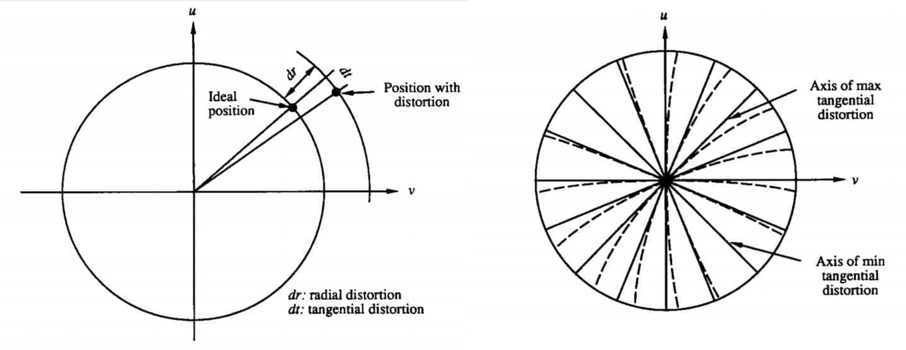


Figure 2.3: Effect of tangential distortion. Solid lines indicate the undistorted case; dashed lines show distortion to the left and right. Both tangential and radial distortion displace a pixel from its ideal position.

2.3.3 Checkerboard Calibration

The standard method for estimating $\mathbf{K}$ and $\mathbf{d}$ is the planar checkerboard technique introduced by Zhang [17]. The approach requires a planar calibration target with a regular pattern of known geometry and a set of images acquired from different viewpoints.

The calibration proceeds as follows. For each image, the inner corner positions are detected using a sub-pixel-accurate corner detector, yielding a set of 2D pixel observations $\{\tilde{\mathbf{x}}_{s,i}\}$. The corresponding 3D world coordinates $\{\mathbf{p}_{w,i}\}$ are known from the target geometry. Because the target is planar, the world coordinate system can be defined so that all corners satisfy $z_{w,i} = 0$, which reduces the projection to a homography.

For each image j, a 3×3 homography $\mathbf{H}_j$ is estimated that maps the homogeneous planar target coordinates $\tilde{\mathbf{p}}_i = (x_{w,i},\, y_{w,i},\, 1)^\top$ to the observed image coordinates:

$$\tilde{\mathbf{x}}_{s,i} \;\sim\; \mathbf{H}_j\, \tilde{\mathbf{p}}_i. \tag{2.14}$$

Since $\mathbf{H}_j = \mathbf{K}[\mathbf{r}_1 \;\; \mathbf{r}_2 \;\; \mathbf{t}_j]$, each homography provides two linear constraints on $\mathbf{K}$ by exploiting the orthonormality of the rotation columns $\mathbf{r}_1$ and $\mathbf{r}_2$. With images from at least three distinct viewpoints, a closed-form solution for $\mathbf{K}$ is obtained via decomposition of the image of the absolute conic [17, 11].

This closed-form solution initializes a subsequent nonlinear refinement in which all intrinsic parameters, distortion coefficients, and per-image extrinsic parameters are jointly optimized by minimizing the total reprojection error, using the Levenberg-Marquardt algorithm [19, 20].

2.3.3.1 Quality Metrics

The quality of a camera calibration is assessed by the reprojection error: the Euclidean distance between each detected corner $\tilde{\mathbf{x}}_{s,ij}$ and its projection $\hat{\mathbf{x}}_{s,ij}$ through the calibrated model. The overall quality is summarized by the root mean square (RMS) reprojection error,

$$\epsilon_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\|\tilde{\mathbf{x}}_{s,ij} - \hat{\mathbf{x}}_{s,ij}\right\|_2^2}, \tag{2.15}$$

where M is the number of calibration images, N_j is the number of detected corners in image j, and $N = \sum_{j=1}^{M} N_j$ is the total number of observations. A well-converged calibration typically yields $\epsilon_{\mathrm{RMS}} < 0.5$ pixels for standard imaging setups; values below 1 pixel are generally considered acceptable for quantitative applications [11].

2.4 Image Segmentation

To identify the catheter structure against the background, a Hessian-based Frangi Vesselness (HFV) filtering approach is used [21]. This method analyzes the local curvature of image intensity to enhance tubular structures while suppressing non-tubular features. The filter operates at the voxel level and is based on the decomposition of the eigenvalue of the local Hessian matrix of the image. The local structure at each voxel and scale is characterized by the sign and magnitude of the corresponding Hessian eigenvalues. By appropriately selecting and combining these eigenvalues within a vesselness function, tubular structures such as catheters and guidewires are enhanced, whereas non-vascular structures and background noise are effectively suppressed. The 2D Frangi vesselness function for detecting vessels with bright contrast is defined as:

$$v(\overline{x}, s) = \begin{cases} 0\ if\ \lambda_2(\overline{x}, s) > 0 \\ exp(-\frac{R_b^2(\overline{x},s)}{2\beta_1^2})(1 - exp(-\frac{S^2(\overline{x},s)}{2\beta_2^2})),\ otherwise \end{cases} \tag{2.16}$$

where $\lambda_1(\overline{x}, s)$ and $\lambda_2(\overline{x}, s)$ are the eigenvalues of the Hessian matrix at position $\overline{x}$ on the scale s, ordered so that $(|\lambda_1(\overline{x}, s)| < |\lambda_2(\overline{x}, s)|)$. The term $R_b(\overline{x}, s) = \frac{|\lambda_1(\overline{x},s)|}{|\lambda_2(\overline{x},s)|}$ represents measure of blobness, distinguishing tubular structures from blob-like structures, while $S(\overline{x}, s) = \sqrt{\lambda_1^2(\overline{x}, s) + \lambda_2^2(\overline{x}, s)}$ quantifies the structureness, reflecting the overall strength of the local second-order structure.

2.5 Centerline Extraction

Centerline extraction is a key step in catheter geometry analysis, as it enables the characterization of instrument shape, length, and curvature. By reducing the segmented catheter to a one-dimensional representation, the centerline provides a compact and meaningful description of its spatial trajectory.

In this work, the centerline is extracted from segmented images through a multi-step process consisting of skeletonization, graph-based path extraction, and subsequent refinement.

2.5.1 Skeletonization

Skeletonization reduces a binary object to a one-pixel-wide representation while preserving its overall topology and connectivity. In this thesis, the segmented catheter mask consists of foreground pixels representing the catheter and background pixels representing the surrounding image. For tubular objects, the resulting skeleton provides an approximation of the object's medial axis.

Skeletonization is performed using the Zhang-Suen thinning algorithm [22]. The algorithm iteratively removes boundary pixels from the foreground object while preserving connectivity. For each foreground pixel P_1 with 8-neighbors $P_2, \ldots, P_9$, let $N(P_1)$ denote the number of non-zero neighbors and $S(P_1)$ denote the number of $0 \to 1$ transitions in the ordered neighborhood.

The algorithm consists of two sub-iterations that are repeated until no further pixels can be removed. In the first sub-iteration, P_1 is removed if

$$2 \leq N(P_1) \leq 6, \quad S(P_1) = 1, \quad P_2 P_4 P_6 = 0, \quad P_4 P_6 P_8 = 0.$$

In the second sub-iteration, P_1 is removed if

$$2 \leq N(P_1) \leq 6, \quad S(P_1) = 1, \quad P_2 P_4 P_8 = 0, \quad P_2 P_6 P_8 = 0.$$

These conditions ensure that only removable boundary pixels are deleted, while endpoints and pixels required to preserve object connectivity are retained.

2.5.2 Graph-based Path Finding

A skeletonized binary object consists of discrete pixels that describe the object's medial structure, but these pixels are not inherently ordered. To obtain a continuous centerline, the skeleton is represented as a graph in which each pixel is a node and edges connect neighboring pixels with weights encoding Euclidean distance. The longest geodesic path between endpoint pairs then yields an ordered coordinate sequence corresponding to the object's primary axis, while naturally suppressing spurious branches and noise-induced artifacts.

2.6 Geometric Analysis

The transformation of discrete, pixel-based skeleton representations into continuous geometric curves is a prerequisite for further analysis. B-splines are used to parameterize the catheter centerline. A B-spline curve $\mathbf{C}(u)$ of degree p is defined as a piecewise polynomial function constructed from a linear combination of control points $\mathbf{P}_i$ and normalized B-spline basis functions $N_{i,p}(u)$ [23]:

$$\mathbf{C}(u) = \sum_{i=0}^{n} N_{i,p}(u) \mathbf{P}_i, \quad 0 \leq u \leq 1 \tag{2.17}$$

The basis functions are defined recursively over a nonperiodic, nonuniform knot vector $\mathbf{U} = \{u_0, u_1, \ldots, u_m\}$ using the Cox-de Boor recursion formula.

2.7 Procrustes Alignment

The comparison of geometric shapes requires a framework to remove variations caused by arbitrary positioning, orientation, and scale. This process is formally known as Procrustes Analysis.

Procrustes alignment finds the optimal similarity transformation (scaling factor c, rotation matrix $\mathbf{R}$, and translation vector $\mathbf{t}$) that maps a source point set onto a target point set by minimizing the mean squared error between corresponding landmarks [24]:

$$e^2(\mathbf{R}, \mathbf{t}, c) = \frac{1}{n} \sum_{i=1}^{n} \|\mathbf{y}_i - (c\mathbf{R}\mathbf{x}_i + \mathbf{t})\|^2 \tag{2.18}$$

The optimal parameters are computed in closed form via singular value decomposition of the cross-covariance matrix of the two point sets, following the method of Umeyama [24]. An illustration of a similarity transformation can be seen in Figure 2.4.

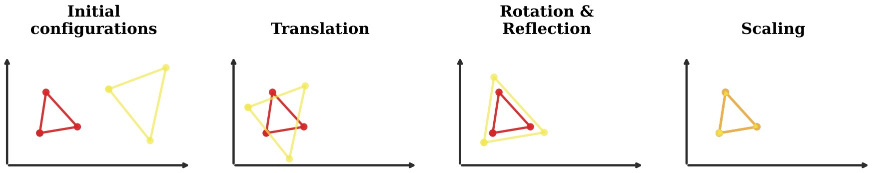


Figure 2.4: Illustration of the sequential similarity transformations applied during Procrustes superposition: translation aligns centroids, rotation and reflection align orientation, and scaling resolves differences in coordinate space magnitudes.

2.8 Accuracy Metrics

The following accuracy measurements have previously been used to evaluate guidewire simulation in endovascular interventions [2]. These metrics will be reviewed and, where appropriate, incorporated into the proposed framework to support quantitative evaluation of simulation accuracy. Furthermore, discussions with Mentice have identified curvature analysis as a specific requirement. Although distance-based metrics quantify spatial deviations, curvature analysis enables evaluation of the instrument's mechanical behavior by assessing whether the simulated catheter bends and deforms realistically at specific locations compared to the physical reference.

2.8.1 Root Mean Square Distance

The root mean square (RMS) error is a widely used metric to quantify the average magnitude of deviation between the simulated and physical trajectories [2, 12]. It provides a single scalar value representing the overall discrepancy between two sets of corresponding points. For a guidewire represented by N ordered centerline points, the RMS error is defined in equation 2.19:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(||x_i^{sim} - x_i^{real}||^2)}, \tag{2.19}$$

where x_i^{sim} denotes the i-th centerline point of the simulated guidewire, and x_i^{real} denotes the corresponding centerline point of the reference (physical) guidewire. The norm $||\cdot||$ represents the Euclidean distance between corresponding points.

2.8.2 Hausdorff Distance and Modified Hausdorff Distance

The Hausdorff Distance (HD) and the Modified Hausdorff Distance (MHD) are commonly used to quantify the discrepancy between two contours. While HD measures the maximum deviation between two point sets (i.e., the worst-case mismatch), MHD captures the average deviation, providing a more robust estimate of the overall positional error in the simulation [2].

MHD reflects the average local distribution of positional errors along the guidewire path. Given two ordered point sets corresponding to the simulated guidewire (G_{sim}) and the real guidewire (G_{real}), the MHD is defined in equation 2.20.

$$MHD(G_{real}, G_{sim}) = \max\left[\underset{p\in S_{real}}{\text{mean}}\, d(p, G_{sim}), \underset{p\in S_{sim}}{\text{mean}}\, d(p, G_{real})\right] \tag{2.20}$$

Similarly, the HD represents the maximum local positional error along the guidewire path. HD is defined in equation 2.21.

$$HD(G_{real}, G_{sim}) = \max\left[\max_{p\in S_{real}} d(p, G_{sim}), \max_{p\in S_{sim}} d(p, G_{real})\right], \tag{2.21}$$

In both formulations, $d(p, G)$ denotes the minimum Euclidean distance between a point P and the point set G, defined as:

$$d(p, G) = \min_{x\in G}||p - x||, \tag{2.22}$$

where $||\cdot||$ represents the Euclidean norm. Thus, MHD characterizes the average discrepancy between the real and simulated guidewire centerlines, whereas HD identifies the largest deviation between them.

2.8.3 Average Distance and Maximum Distance

The Average Distance (AD) and Maximum Distance quantify the mean and peak spatial separation between the simulated and real guidewire centerlines, respectively [2].

The AD represents the mean Euclidean distance between corresponding ordered centerline points of the simulated and real guidewires. For a guidewire consisting of N centerline points, AD is defined as:

$$AD = \frac{1}{N} \sum_{i=0}^{N-1} ||x_i^{sim} - x_i^{real}||, \tag{2.23}$$

where x_i^{sim} and x_i^{real} denote the i-th centerline points of the simulated and real guidewires, respectively, and $|| \cdot ||$ represents the Euclidean norm.

The Maximum Distance corresponds to the largest Euclidean distance between corresponding centerline points of the two guidewires in the 2D medical image [2, 25]. It reflects the peak local deviation between the simulated and physical trajectories.

2.8.4 Dice Similarity Coefficient

The Dice Similarity Coefficient (DSC) is used to quantify the degree of overlap between the segmented regions of the virtual and physical instruments [2, 25]. It is a widely adopted metric for evaluating spatial similarity between two binary images. The DSC is defined as twice the area of overlap between the simulated guidewire segmentation (G_{sim}) and the real guidewire segmentation (G_{real}) divided by the sum of their individual areas, as expressed in Equation 2.24.

$$DSC = \frac{2|G_{real} \cap G_{sim}|}{|G_{real}| + |G_{sim}|} \tag{2.24}$$

where $| \cdot |$ denotes the area (or number of pixels) of the segmented region, and $\cap$ represents the intersection between the two segmentations.

3

System Overview

Quantitative comparison between physical catheter configurations and their simulated counterparts requires a workflow that is both reproducible and physically grounded. The same preprocessing assumptions, the same geometric representations, and the same metric definitions must apply consistently across every experiment. To satisfy this requirement, a software framework was developed to integrate physical image acquisition, image-based geometric analysis, and simulation control into a single configurable pipeline.

The framework was implemented in Python, using PySide6 for the graphical interface, OpenCV and scikit-image for image processing, SciPy for geometric operations, and Optuna for Bayesian optimization. Communication with the Mentice VIST software is handled via a gRPC interface, which allows the framework to control simulation state, instrument parameters, and image capture during parameter search.

3.1 Processing Pipeline

The processing pipeline converts each raw image input into a physically scaled, continuous geometric representation of the catheter centerline. It is designed to be deterministic for a fixed configuration, with all geometric outputs expressed in millimeters rather than pixels, and modular so that outputs can be inspected and cached independently.

The pipeline consists of five phases, illustrated in Figure 3.1: input and physical prior derivation from the catheter French size, preprocessing (undistortion and region-of-interest restriction), segmentation and centerline extraction, physical scale conversion and centerline refinement, and finally spline fitting with geometric feature extraction. Each input produces a structured output containing the spline centerline coordinates, curvature profile, radius of curvature, total length, and the millimeter-per-pixel scale factor. Implementation details are given in Section 4.3.

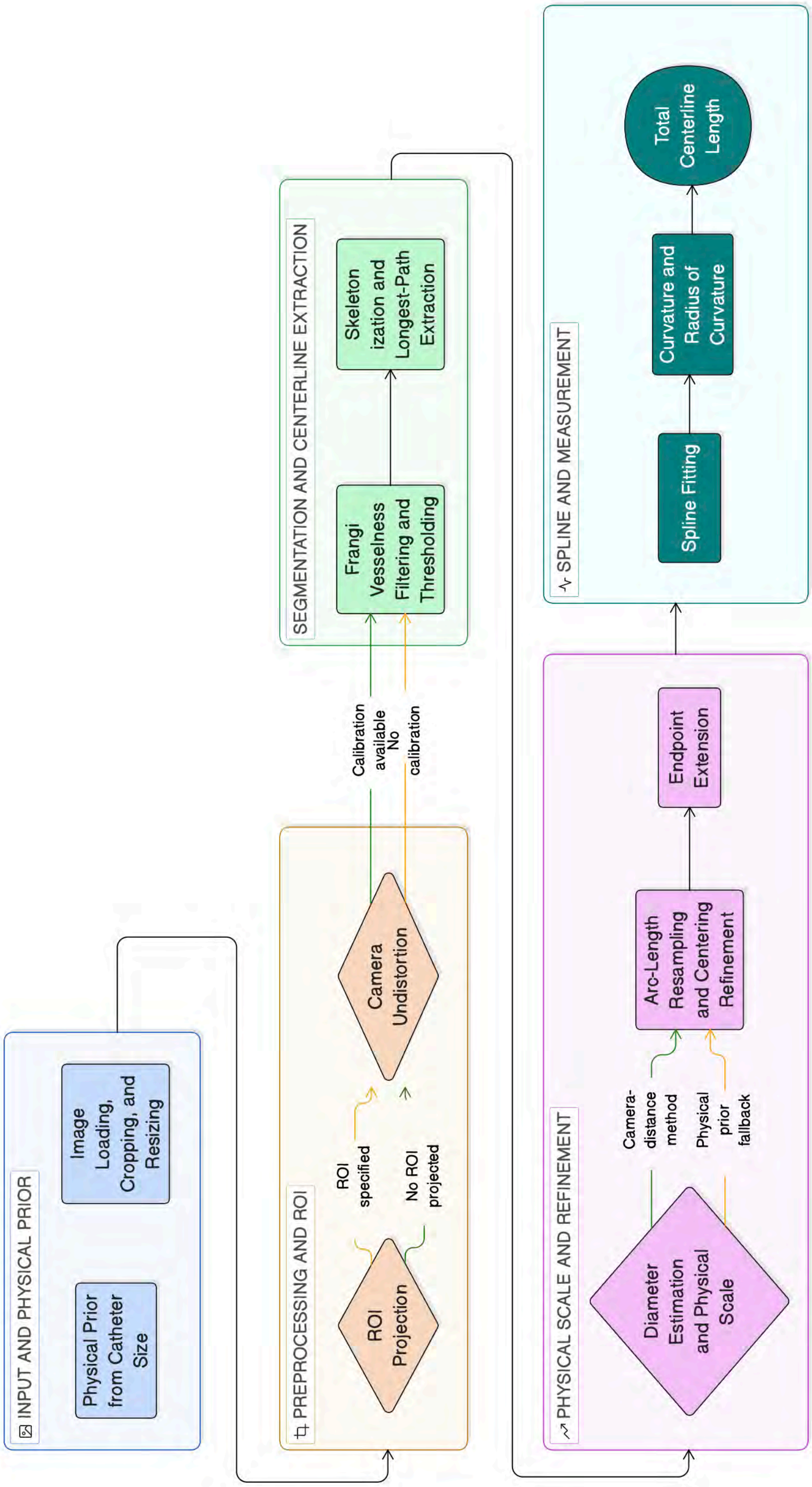


Figure 3.1: Overview of the processing pipeline, organized into five phases: input and physical prior, preprocessing, segmentation and centerline extraction, physical scale and refinement, and geometric modeling and feature extraction. Diamond nodes indicate branching decisions; rectangular nodes indicate processing steps.

3.2 Comparison Module

The comparison module takes two pipeline outputs, one from a physical catheter image and one from a simulated image, and produces a quantitative assessment of their geometric correspondence. Because both inputs share an identical data structure and have passed through the same pipeline, any measured discrepancy reflects a genuine geometric difference rather than a preprocessing artifact. The module computes metrics across three levels of sensitivity: positional error along the centerline, area-level segmentation overlap, and curvature profile similarity. The details of the implementation are given in Section 4.4.

4

Method

This methodology consists of four main stages. First, physical and simulated fluoroscopic images of catheters and guidewires are acquired using controlled experimental setups. Second, the images are processed to extract catheter centerlines and geometric representations. Third, alignment and quantitative comparison metrics are applied to evaluate similarity between physical and simulated configurations. Finally, an automated stiffness sweep framework combined with Bayesian optimization is used to identify simulation parameters that best reproduce the observed physical catheter behavior. Each stage is designed to address one or more of the four research questions defined in Section 1.5.

4.1 Hardware Configuration

The image acquisition system consists of a Raspberry Pi Camera Module 3 mounted on a rigid overhead rig above the experimental surface, connected to a Raspberry Pi 5 Model B via a CSI ribbon cable. The framework runs on a Dell Precision 7530 computer connected to the Pi over a direct Ethernet link and triggers image capture remotely over SSH. This setup was used consistently across all experiments in this thesis: validation (Section 4.5), physical SIM vs. VIST (Section 4.6) comparison, and the stiffness sweep optimization protocol (Section 4.7).

The camera is mounted perpendicular to the image plane at a fixed height of 392 mm, which was measured and kept constant across all experiments. The background surface was chosen per experiment to either be white or black, which maximized the contrast of the catheter in the image. Lighting was ambient room lighting and was not actively controlled. The camera setup consists of the Raspberry Pi camera module 3, Raspberry Pi 5 B and the camera mount, as can be seen in figure 4.1.

4.1.1 Raspberry Pi Camera Module 3

The Raspberry Pi Camera Module 3 is used for all image acquisition in this work. It is based on the Sony IMX708 sensor and features an 11.9 MP resolution of 4608×2592 pixels and an integrated autofocus system [26]. The autofocus capability is relevant for this application because catheter thickness varies across instruments, and consistent focus across the full catheter length is necessary for reliable segmentation. An integrated infrared filter is also included, which suppresses IR contamination under mixed lighting conditions. The camera supports up to 50 fps

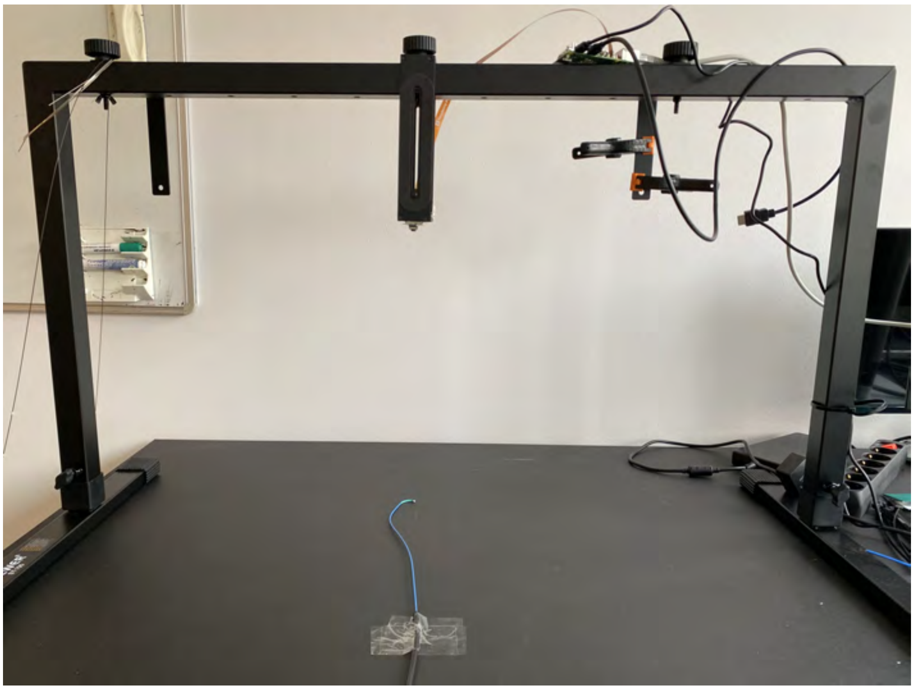


Figure 4.1: Camera setup used for capturing images of the catheters and guidewires, showing the overhead rig with the camera mounted at a fixed height of 392 mm above the image plane.

at Full HD resolution, though only still-image capture is used in this work.

4.1.2 Raspberry Pi 5 B

The Raspberry Pi 5 Model B serves as the dedicated compute unit for camera control and image transfer [27]. It features an ARM Cortex-A76 CPU running at 2.4 GHz and 1 GB of RAM. Rather than running the full analysis framework on the Pi, the Pi is used solely for camera control and image acquisition. This keeps processing off the Pi and avoids the memory constraints of running a full image analysis pipeline on embedded hardware.

4.2 Camera Calibration

The camera was calibrated using a planar checkerboard target with a (7×9) inner-corner layout and a square size of 20.0 mm, following Zhang's method (Section 2.3.3) as implemented in OpenCV. A total of 22 images were acquired at the native camera resolution of (4608×2592) pixels, with varying checkerboard pose to constrain both intrinsic and distortion parameters. The resulting parameters were subsequently applied to undistort all images prior to geometric analysis and length estimation. The summary of the calibration results can be found in Section 5.1.

4.3 Processing Pipeline Implementation

The processing pipeline described in Section 3.1 was configured and applied consistently across all experiments in this thesis. The following describes the specific implementation decisions made for each pipeline phase and the rationale behind parameter choices given the imaging conditions. Default configuration parameters are summarized in Table 4.1.

Table 4.1: Pipeline configuration parameters and their default values.

Parameter	**Default**	**Unit**	**Rationale**
Processing width	2160	px	Balances sub-mm feature resolution with computation time
Frangi σ range	[3, 5)	px	Covers expected catheter cross-sections at processing resolution (see Step 5)
Frangi threshold	0.070	-	Empirically tuned to separate catheter ridge response from background noise
Spline samples	10 000	-	Provides dense enough sampling for smooth curvature estimation
Resampling step	0.025	mm	Sub-pixel physical spacing; avoids aliasing in curvature computation
Curvature window	0.25	mm	Suppresses pixel-grid discretization noise while preserving clinically relevant bending features
Spline tolerance	0.15	px	Allows minor smoothing of skeleton noise without distorting true bends

For each input, the pipeline accepts a raw media path, catheter size in French units, calibration mode and optional camera calibration file, optional object distance for pinhole-based scaling, and an optional rectangular or polygonal ROI in raw-image coordinates. It returns a structured analysis tuple

$$(\text{img},\ \text{mask\_visual},\ \text{path\_centered},\ x_s,\ y_s,\ u,\ \kappa_{\text{mm}},\ r_{\text{mm}},\ L_{\text{mm}},\ m),$$

where x_s, y_s are spline centerline coordinates, u is the normalized arc-length parameter, κ_{mm} is curvature in mm^{-1}, r_{mm} is radius of curvature, L_{mm} is total centerline length, and m is the estimated mm-per-pixel scale factor. This output structure is shared across all analysis modes, ensuring that the comparison module receives identically formatted data regardless of whether the input originated from a physical or simulated image, which directly supports the measurement consistency required to answer RQ3 and RQ4.

Step 1: Physical Prior from Catheter Size

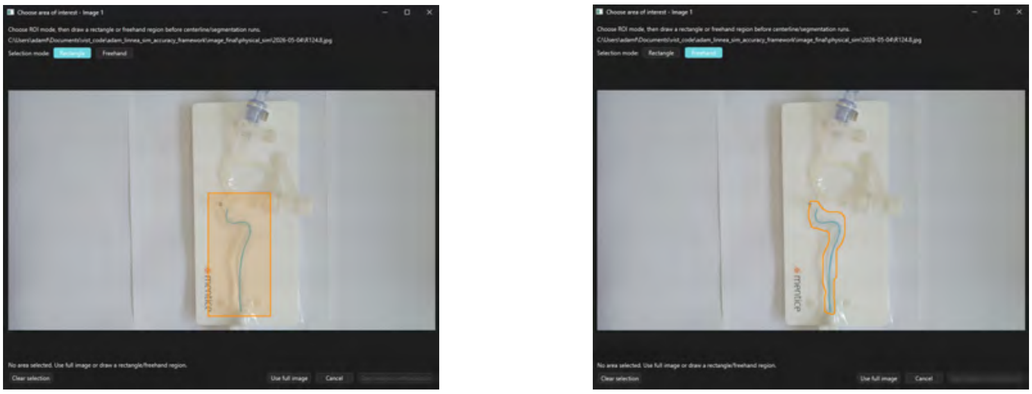
Before any image processing begins, a physical diameter prior is derived from the catheter's French size using the standard conversion $d_{\text{mm}} = Fr/3$. This prior serves two roles: it anchors the Frangi sigma range to the expected catheter width in pixels (Step 5), and it provides a fallback scale estimate when camera-based calibration is unavailable (Step 7). Using a physical prior rather than a fixed pixel assumption ensures that the pipeline generalizes across the range of catheter sizes used in this study without manual re-tuning between experiments.

Step 2: Image Loading, Cropping, and Resizing

The input image is resized to the configured processing width while preserving aspect ratio, with optional pre- and post-crop passes to remove border artifacts. The post-crop is particularly important for simulated frames from VIST, which typically carry black margins that would otherwise bias the Frangi ridge-polarity detection in Step 5. Resizing to a fixed width normalizes the spatial frequency content of the image, ensuring that the Frangi sigma range corresponds to consistent physical scales regardless of whether the input comes from the Raspberry Pi camera or the VIST simulator.

Step 3: ROI Projection

If a user-defined ROI (Figure 4.2) has been specified in raw-image coordinates, it is geometrically projected into the processed-image coordinate space by composing the pre-crop offset, the resize scale factor, and the post-crop offset. Restricting processing to the ROI reduces the influence of background structures and is particularly important when the field of view includes non-catheter features such as phantom walls or guide-tube connectors, as encountered in the physical SIM experiments described in Section 4.6.



Choose ROI mode, then draw a rectangle or freehand region before centerline/segmentation runs.

No area selected. Use full image or draw a rectangle/freehand region.

Figure 4.2: Region of interest selection interface, allowing the user to restrict processing to the relevant anatomical and instrument area. Supports both rectangle and free hand control.

Step 4: Camera Undistortion

When a camera calibration file is available, radial and tangential lens distortion is corrected prior to segmentation using the intrinsic and distortion parameters estimated during checkerboard calibration (Section 2.3). Undistortion is applied at this stage, before any geometric measurements, because all subsequent length and curvature estimates assume a rectilinear projection model (Section 2.2). If the calibration file cannot be loaded, the original image is retained and processing continues, with a corresponding reduction in metric accuracy for length and curvature outputs.

Step 5: Frangi Vesselness Filtering and Thresholding

Catheter enhancement is performed using Hessian-based Frangi vesselness filtering (Section 2.4). Ridge polarity (dark-on-light or light-on-dark) is determined automatically from the global image brightness, allowing the same pipeline to process both standard camera images and simulated fluoroscopy frames without manual configuration.

The Frangi sigma range, specified in the configuration as pixel values, is grounded in the physical catheter diameter. The expected catheter radius in pixels at the processing resolution is

$$\sigma_{\text{px}} = \frac{d_{\text{mm}}/2}{m},$$

and the configured sigma range is chosen to bracket this value across the supported French-size range. A lower bound is enforced for numerical stability of the Hessian eigenvalue decomposition.

The resulting vesselness map is binarized at the configured threshold, the ROI mask from Step 3 is applied, and connected-component analysis retains only the dominant structure, discarding smaller fragments and noise regions.

Step 6: Skeletonization and Longest-Path Extraction

The binary catheter mask is reduced to a one-pixel-wide medial axis using the Zhang-Suen thinning algorithm (Section 2.5.1), and the skeleton is converted to a graph in which each pixel is a node and edge weights encode Euclidean distance between neighbors (Section 2.5.2). The primary centerline is then extracted as the longest geodesic path between detected endpoint pairs. When no topological endpoints exist, for example when skeletonization produces a closed loop due to segmentation noise, a farthest-node strategy is applied to the largest connected component instead. The result is an ordered pixel-coordinate polyline approximating the catheter centerline.

Step 7: Diameter Estimation and Physical Scale

Accurate conversion from pixel to physical coordinates requires an estimate of the mm-per-pixel scale factor m. The Euclidean distance transform of the binary mask is sampled at each centerline point to estimate the local vessel radius. The median value provides a robust diameter estimate that is insensitive to local segmentation failures:

$$d_{\mathrm{px}} = 2\,\mathrm{median}(r_{\mathrm{DT}}).$$

The scale factor is then computed by one of two methods depending on available calibration data. When a camera calibration file has been loaded and an object distance Z is provided, the pinhole projection model (Section 2.2) gives

$$m = \frac{Z}{f_{\mathrm{px,scaled}}}, \qquad f_{\mathrm{px,scaled}} = f_x \cdot \frac{W_{\mathrm{proc}}}{W_{\mathrm{calib}}},$$

where f_x is the calibrated horizontal focal length and $W_{\mathrm{proc}}/W_{\mathrm{calib}}$ accounts for the resize applied in Step 2. When no object distance is available, the physical prior from Step 1 is used as a fallback:

$$m = \frac{d_{\mathrm{mm}}}{d_{\mathrm{px}}}.$$

The camera-distance method is preferred when available because it does not assume that the segmented catheter diameter is free of thresholding bias. The diameter fallback is nonetheless robust in practice because the median distance transform is largely insensitive to moderate threshold variations.

Step 8: Arc-Length Resampling and Centering Refinement

The raw skeleton polyline is non-uniformly spaced due to the discrete pixel grid. To obtain a geometrically consistent representation, the path is resampled uniformly in arc length at a physical step size of $\Delta s_{\mathrm{px}} = \Delta s_{\mathrm{mm,target}}/m$. Each resampled point is then refined by shifting it toward the local maximum of the squared distance-transform field, effectively snapping it to the medial axis of the catheter cross-section. This centering step is particularly important for thick catheter segments where the skeleton may lie off-center due to asymmetric segmentation boundaries, and it reduces systematic bias in curvature estimates near high-curvature regions.

Step 9: Endpoint Extension

Morphological thinning algorithms systematically underestimate the length of tubular objects because the skeleton terminates approximately one local radius before the true mask boundary. To compensate, the centerline is extrapolated at both ends by projecting along the local tangent direction until the catheter mask boundary is reached. This correction is small in absolute terms, typically on the order of one catheter radius, but it is consequential for tip-region curvature estimates and for total-length measurements, where a consistent end-to-end definition is necessary for valid cross-experiment comparison.

Step 10: Spline Fitting

A smoothing B-spline (Section 2.6) is fitted to the refined centerline to obtain a continuous, differentiable curve free of pixel-grid discretization artifacts. Near-duplicate points introduced by the resampling and snapping steps are removed prior to fitting.

Given the remaining noisy centerline observations $\{\mathbf{y}_j\}_{j=1}^{m}$, the spline fitting is formulated as a regularized optimization problem. Extending classic smoothing spline methodology [28] to a parametric framework [29], the objective is to find a curve $\mathbf{C}(u)$ that minimizes a cost function balancing data fidelity and global smoothness:

$$\mathcal{L} = \sum_{j=1}^{m} w_j \|\mathbf{C}(u_j) - \mathbf{y}_j\|^2 + \lambda \int_{u_0}^{u_m} [\mathbf{C}''(u)]^2 du \tag{4.1}$$

where w_j represents a weighting factor for individual points and λ is the global smoothing parameter.

To ensure the spline honors the physical extents and boundary constraints of the device, the endpoint weights w_j at the proximal and distal boundaries are significantly elevated relative to the interior points. The choice of λ is critical; it must be large enough to suppress discretization artifacts from the pixel grid but small enough to preserve the high-frequency bending moments of the catheter. To achieve this across varying inputs, the smoothing parameter is set adaptively as:

$$\lambda = N_{\text{pts}} \cdot (\tau_{\text{px}})^2, \tag{4.2}$$

where N_{pts} is the number of input points and τ_{px} is the configured spline tolerance. This form scales the allowed total residual with the number of input points, ensuring that the effective per-point tolerance remains constant regardless of catheter length.

Step 11: Curvature and Radius of Curvature

Curvature is computed analytically from the spline's first and second parametric derivatives:

$$\kappa_{\text{px}}(u) = \frac{|x'(u)y''(u) - y'(u)x''(u)|}{(x'(u)^2 + y'(u)^2)^{3/2}},$$

and converted to physical units via $\kappa_{\text{mm}}(u) = \kappa_{\text{px}}(u)/m$. A moving-average smoother is applied over a physically fixed window of 0.25 mm to suppress residual numerical noise in the second derivative without attenuating clinically relevant bending features. The radius of curvature $r_{\text{mm}}(u) = 1/\kappa_{\text{mm}}(u)$ is retained as a more interpretable quantity for regions of low curvature.

Step 12: Total Centerline Length

Total catheter length is computed by summing the Euclidean distances between consecutive spline sample points and converting to physical units:

$$L_{\text{mm}} = \sum_{i} \|\mathbf{p}_{i+1} - \mathbf{p}_i\|_{\text{px}} \cdot m.$$

Using the dense spline representation rather than the raw skeleton polyline reduces length bias from pixel-grid quantization, particularly for curved segments where a coarse polyline would systematically underestimate arc length. The accuracy of this length estimate is validated experimentally in Section 4.5.1, directly addressing RQ3 regarding measurement consistency across different catheter configurations.

For simulated images, instrument length is not extracted through image processing. Instead, the simulator has direct access to the instrument's three-dimensional state and computes the visible length analytically from its physics model.

4.4 Comparison and Metrics Implementation

The comparison module receives two pipeline outputs: a reference object derived from a physical catheter image and a simulated object derived from a VIST simulator frame. Both carry identical data structures because they have passed through the same upstream pipeline described in Section 3.1. This guarantees that any measured discrepancy between the two reflects a genuine geometric difference between the physical and simulated catheter configurations, rather than an artifact of inconsistent preprocessing. Before metric computation, both centerlines are resampled to a common point count under a shared arc-length parameterization, ensuring that pointwise distances are computed between geometrically corresponding locations rather than arbitrarily indexed points.

4.4.1 Alignment

Before metric computation, the simulated centerline is registered to the reference using Procrustes alignment (Section 2.7). Two alignment modes are provided to accommodate different experimental contexts.

The default Procrustes-only mode minimizes the global RMS correspondence error and is appropriate when neither centerline has a geometrically privileged anchor point. This mode was used for the validation experiments (Section 4.5), and for the physical SIM versus VIST comparison (Section 4.6), where the goal is to assess overall shape similarity regardless of absolute position.

The endpoint-lock mode addresses the fact that tracking errors in complex medical devices accumulate non-uniformly along the device length. It fixes the proximal endpoint of the simulated centerline to that of the reference before alignment, using it as an anchor. Applying the resulting transformation to the full configuration places both centerlines in a shared coordinate frame anchored at the insertion point, so that metric values reflect the relative spatial divergence of the distal segment rather than global shape difference. This mode is used in stiffness-sweep contexts, where the catheter insertion point is a fixed physical constraint and the distal shape is the quantity under evaluation.

The choice of alignment mode is a methodological decision that must remain consistent within any single experiment, as different modes produce different metric values for the same input pair. This represents one of the practical sources of uncertainty relevant to RQ4.

4.4.2 Metric Computation

The full metric suite is computed from the aligned, resampled centerline pair and the corresponding segmentation masks, with each metric designed to capture a complementary aspect of geometric deviation.

Positional metrics - RMS error, mean correspondence distance, maximum correspondence distance, Hausdorff distance, and modified Hausdorff distance (Section 2.8) - characterize the spatial deviation between centerlines at different levels of sensitivity to outliers. RMS and mean distance summarize the average discrepancy across the full centerline, while Hausdorff distance captures the worst-case deviation. Modified Hausdorff distance provides a more robust alternative by averaging over local maxima rather than taking the global maximum.

Mask-level overlap is evaluated using the Dice similarity coefficient (Section 2.8.4). Because the Dice score operates on the full segmentation area rather than the one-dimensional centerline, it captures thickness and boundary errors that centerline-distance metrics cannot detect.

Curvature profiles from both inputs are compared pointwise, yielding four scalar metrics: MAE, RMS, maximum absolute difference, and the arc-length location of peak discrepancy. To support visual inspection alongside these quantitative measures, curvature and bending radius heatmaps are rendered by mapping the computed profile values onto the extracted centerline and overlaying them on the corresponding input image, allowing regions of mechanical discrepancy to be localized directly in image space.

4.5 Validation of the Algorithms

This validation step is necessary to address RQ3 and RQ4 before cross-experiment comparisons are performed. For the measurements to be consistent across multiple catheters and configurations, the image-based length estimation algorithm must first be confirmed to be accurate and stable. Similarly, understanding the measurement error characteristics - in particular, how error varies with object length and shape - is essential for interpreting the uncertainty associated with all subsequent metric values. A method that exhibits systematic bias or shape-dependent error would introduce a confound into the comparison between physical and simulated catheter geometries.

4.5.1 Length Estimation

The accuracy of the image-based length estimation algorithm was evaluated by comparing measured catheter lengths against known physical reference lengths. A catheter was placed straight alongside a ruler on a flat surface and photographed using the Raspberry Pi camera module, as shown in Figure 4.3. Seventeen measurements were taken across a range of 1.0 to 25.0 cm, with the ground truth defined by the ruler markings. For each image, the pipeline extracted the catheter centerline and computed the total arc length in millimeters, which was then compared against the corresponding reference value. Agreement between the two methods was assessed using mean absolute error, mean relative error, and Bland-Altman analysis.

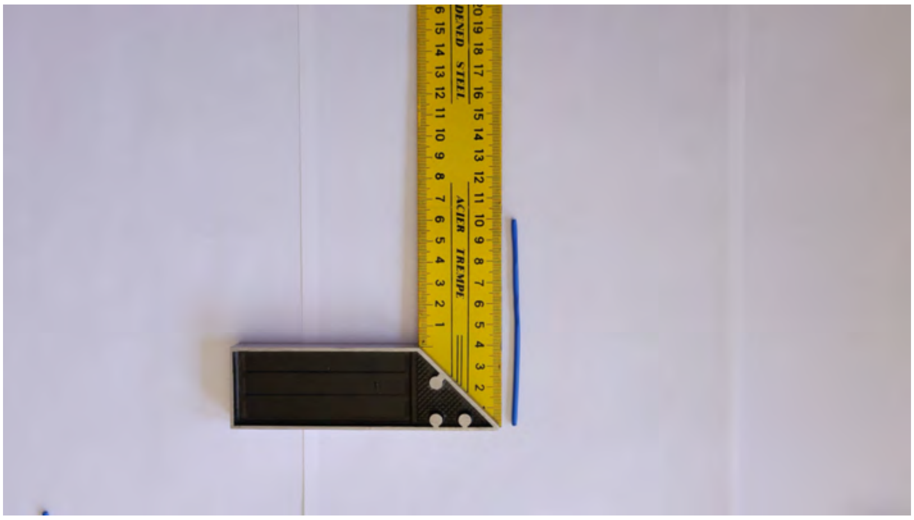
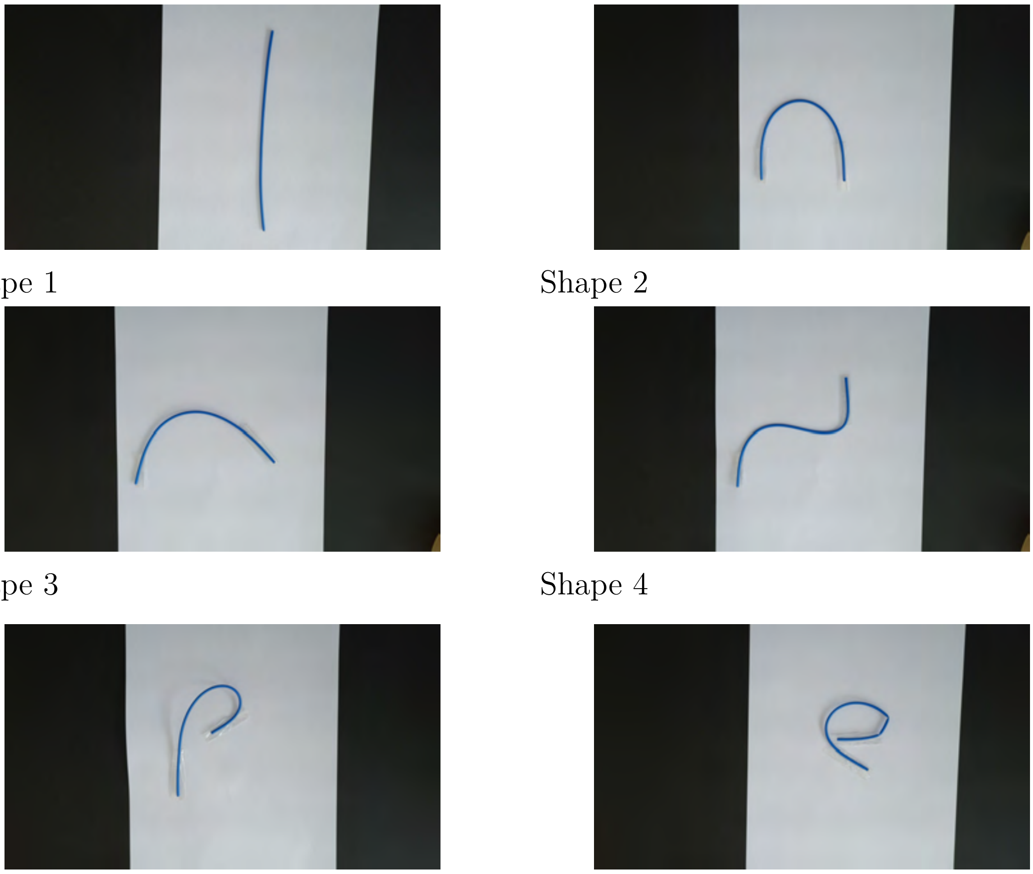


Figure 4.3: Experimental setup for length validation, showing a catheter placed alongside a ruler on a flat surface.

4.5.2 Measurement Consistency Across Shape Variations

To evaluate whether the algorithm produces consistent length estimates regardless of catheter shape, a single catheter of known physical length 20 cm was bent into six distinct configurations and imaged separately for each, as illustrated in Figure 4.4. Configurations were chosen to span a range from nearly straight to tightly curved. Since the same physical catheter was used throughout, the true arc length remained constant across all configurations. The pipeline was applied to each image independently, and the resulting length estimates were compared against the known reference length. This test was designed to confirm that the measurement is invariant to bending deformation, which is a necessary property for the framework to generalize across different catheter configurations in the main experiments.



Shape 1

Shape 2

Shape 3

Shape 4

Shape 5

Shape 6

Figure 4.4: The six bending configurations used to evaluate measurement consistency, each showing the same 20 cm catheter bent into a distinct shape.

4.5.3 Alignment

The alignment procedure was validated by imaging a catheter in one configuration, then physically rotating it and capturing a second image, as illustrated in Figure 4.5. This produced two images of the same catheter with a known rotational difference between them. Both images were processed through the pipeline to extract their respective centerlines, after which the alignment algorithm was applied to register the second centerline onto the first. The quality of alignment was assessed by comparing the RMS correspondence error before and after the transformation, and by visually inspecting the overlay of the aligned centerlines.

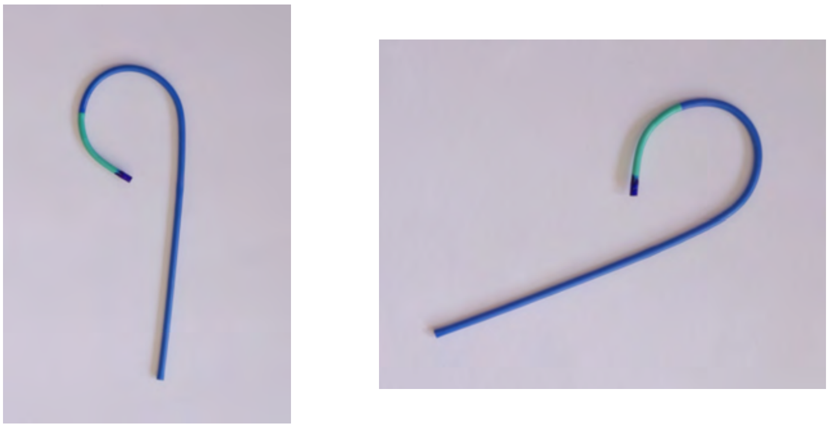


Figure 4.5: Images used for alignment validation, showing the catheter in its original orientation (left) and after physical rotation (right).

4.6 Physical SIM vs. VIST

The physical simulator (SIM) was compared with the VIST simulator by performing equivalent instrument manipulation procedures in both systems. The experiments were conducted using a 4-Fr micro catheter. Both the physical phantom and the VIST case were based on the same neurovascular anatomy, representing the internal carotid artery (*a. carotis interna*) including the carotid siphon and Circle of Willis. Both models are shown in Figure 4.6.

The catheter was first advanced through the physical SIM phantom until the tip reached the carotid siphon. It was then progressively withdrawn, and a still image was captured at each of 13 evenly spaced insertion depths.

The same 13 insertion depths were then replicated in VIST by manually positioning the catheter to the corresponding depth in the simulation, using the insertion length derived from the physical images as a reference. The fluoroscopy projection angle in VIST was matched to the camera viewing angle used during physical acquisition, ensuring that the two image modalities captured comparable two-dimensional projections of the catheter geometry. At each position, the simulator was allowed to reach a settled state before a fluoroscopic frame was captured, ensuring that the recorded image reflected a stable equilibrium configuration of the instrument.

The resulting 13 image pairs were processed through the pipeline described in Section 3.1 to extract catheter centerlines and geometric descriptors. The two-image comparison mode of the framework was used to apply the comparison module described in Section 3.2 to each pair, computing the full metric suite. Prior to metric computation, both centerlines were aligned using Procrustes alignment without endpoint locking, as no fixed anatomical anchor point could be reliably identified across both environments. Length matching was applied where the visible instrument extents differed between the two systems.

(a) Physical SIM phantom. Injected with coffee to highlight the vessel structure.

(b) Corresponding VIST anatomy model.

Figure 4.6: The neurovascular phantom used in the physical SIM experiments (left) and the corresponding anatomy model in VIST (right), both representing the internal carotid artery including the carotid siphon and Circle of Willis.

4.7 Stiffness Sweep and Bayesian Optimization

A central challenge in endovascular simulator validation is that the mechanical behavior of a catheter is governed by its stiffness distribution along the shaft, and this distribution is not directly observable from image alone. Determining which stiffness parameters cause the simulator to reproduces given physical catheter configuration therefore requires searching a continuous, multi-dimensional parameter space. Manual search is impractical because each candidate configuration requires a full simulation run, a fluoroscopic frame capture, and a complete comparison pipeline evaluation, making the cost per trial high relative to the number of trials that would be needed for exhaustive coverage. The stiffness sweep module addresses this by automating the search process, integrating directly with both the VIST and the comparison pipeline.

4.7.1 Stiffness Parameterization

The VIST catheter stiffness model is parameterized by three control points (y_1, y_2, y_3) that define the bending stiffness at fixed positions along the catheter shaft, from proximal to distal. The positions of the two intermediate control points, x_2 and x_3, can either be fixed at their default locations or included as additional search variables which reflects the physical requirement that anchor positions are monotonically ordered along the shaft. Search bounds for all variables are configured in the launcher before the sweep begins.

The stiffness values assigned to these control points are not expressed in absolute physical units. Rather, the simulator interprets them as relative scaling factors applied to an internal baseline stiffness, so that a value of 1.0 corresponds to baseline behavior, while values above or below 1.0 represent proportionally stiffer or more flexible configurations respectively. The shape of the resulting stiffness profile along the shaft - and therefore the simulated bending response - is governed by the ratios between the control point values, not by their absolute magnitudes. This relative parameterization means that the search space explored during optimization is best understood as a space of stiffness profiles rather than a space of physically measurable material constants. Thus, the parameter values recovered by the optimizer should be interpreted as the profile that best reproduces the observed catheter geometry under the given simulator conditions, rather than as direct measurements of the physical catheter's mechanical properties. The control point layout along the shaft is illustrated in Figure 4.7a, and the resulting piecewise-linear stiffness profile is shown in Figure 4.7b.

4.7.2 Simulator Control Interface

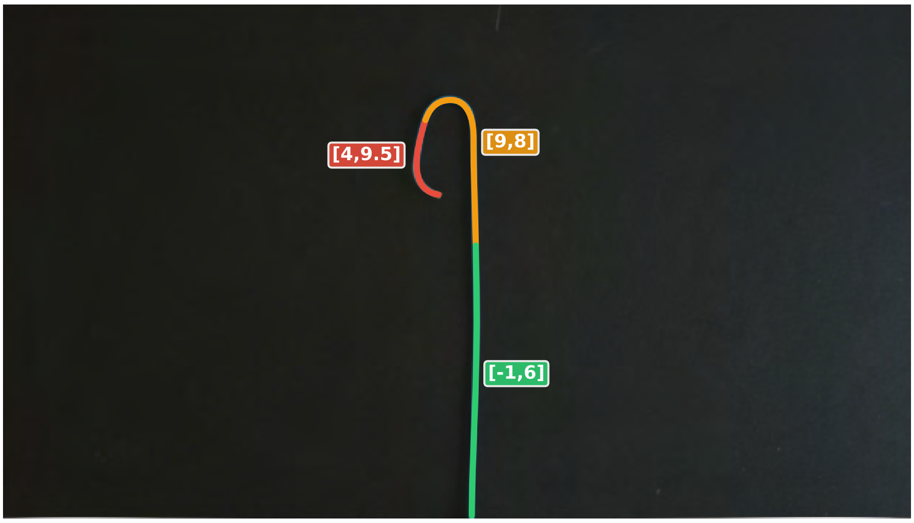
The stiffness sweep communicates with the VIST simulator through a gRPC interface, which serves as the control and data channel between the Python sweep runner and the simulation application. All instrument positioning, state management, and image capture are performed through this interface. Figure 4.8 summaries the full control sequence; the subsections below describe each phase in detail.

Startup

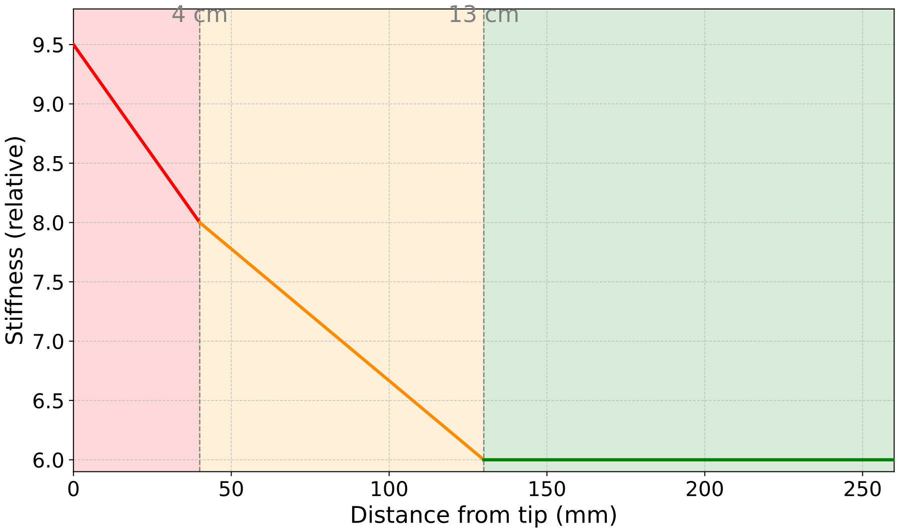
Before any sweep trials begin, the runner verifies that the simulator is running. Once the simulator is active, a single gRPC channel is opened to the simulator process, through which all subsequent communication occurs. The runner then waits until the simulator reports that it has reached a running state before proceeding.

Device Preparation

The instrument is prepared in a fixed sequence. The instrument is moved to its target insertion position automatically using the length derived from the reference image (Section 4.3). Settling is confirmed by polling the instrument state until the position error below configured thresholds (0.30 cm). Once settled, the instrument



(a) Key profile control coordinates.



(b) Relative stiffness profile along the shaft.

Figure 4.7: VIST catheter stiffness parameterization curves and corresponding control segment boundaries.

state is saved as a baseline snapshot that will be restored at the start of each trial.

Fluoroscopy Stream

A continuous fluoroscopic frame stream is opened at the start of the sweep and kept active across all trials. The X-ray source is activated, and an initial set of frames is discarded to allow the image to stabilize before any captures are used for analysis.

Per-Trial Execution

For each candidate stiffness configuration, the following sequence is executed. The instrument state is first restored to the saved baseline, ensuring that each trial be-

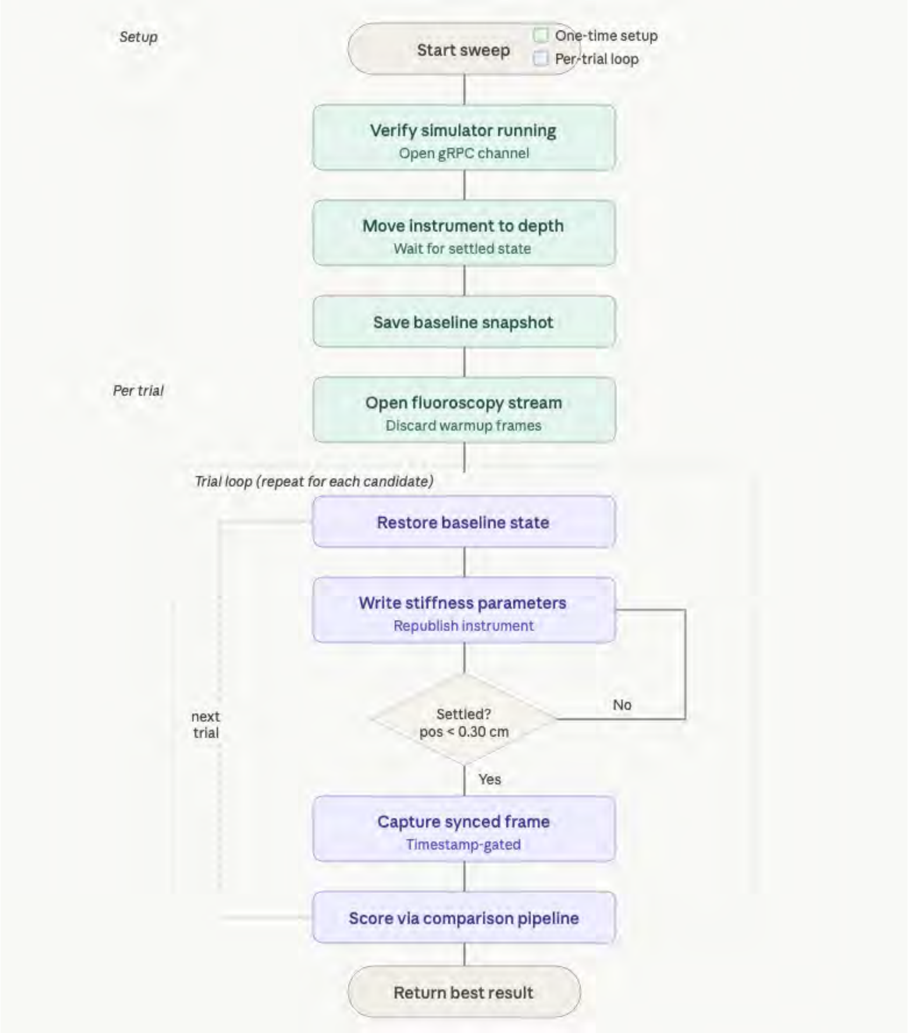


Figure 4.8: Flowchart of the simulator control interface. Teal nodes represent one-time setup operations performed before the first trial; purple nodes represent the per-trial execution loop. The settling check repeats until position error fall below their configured thresholds before frame capture proceeds.

gins from an identical starting condition. The stiffness parameters are then written to a local configuration file, which the simulator reads when the instrument is next published. The instrument is republished with the new stiffness values and moved to the target position, after which the runner again waits for the settled-state conditions to be met.

Once the instrument is settled, a synchronized frame is captured from the fluoroscopy stream. Synchronization is achieved by reading the current simulation time and advancing through the frame stream until a frame whose timestamp meets or exceeds the target time is found. This timestamp-gated capture ensures that the recorded frame reflects the instrument's settled configuration rather than a transient state. The captured frame is saved as an image file and passed to the comparison pipeline for scoring.

4.7.3 Bayesian Optimization

Bayesian optimization is employed to search the stiffness parameter space efficiently: because each trial requires a complete simulation run and image capture, minimizing the number of trials needed to locate a good parameter set has direct practical value. Trial proposals are generated by Optuna using a Tree-structured Parzen Estimator (TPE) sampler, which builds a probabilistic model of the objective function from completed trials and proposes new candidates in regions of the parameter space that are likely to improve on the current best result. The trial schedule begins with a set of random startup trials to initialize the model, followed by model-guided optimization trials.

4.7.4 Composite Objective Function

Although each trial yields several geometric error measures, the optimizer requires a single scalar objective. The composite score f is defined as a weighted combination of four terms:

$$f = 0.30 \cdot \varepsilon_{\mathrm{RMS}} + 0.25 \cdot \varepsilon_{\mathrm{P95}} + 0.25 \cdot \varepsilon_{\kappa} + 0.20 \cdot \varepsilon_{\mathrm{tip}} \tag{4.3}$$

where $\varepsilon_{\mathrm{RMS}}$ is the root-mean-square correspondence error, $\varepsilon_{\mathrm{P95}}$ is the 95th-percentile correspondence error, ε_{κ} is the curvature mismatch, and $\varepsilon_{\mathrm{tip}}$ is the tip position error. All four terms are computed over the matched segment to ensure a fair spatial comparison, as described in Section 3.2.

The distribution of weights across these multiple metrics was chosen to establish a balanced optimization objective. This strategy prevents the optimization loop from over-indexing on or exploiting edge cases in a single dimension of accuracy such as finding parameter sets that minimize global distance ($\varepsilon_{\mathrm{RMS}}$) at the expense of severe localized shape distortions or poor tip placement, and thereby ensuring that the optimized simulation parameters reflect true physical behavior across all measured geometric attributes.

The largest weight (0.30) is assigned to the RMS term because it summarizes global correspondence uniformly along the matched length and is the metric used as the ranking criterion in manual comparison workflows; giving it the highest influence ensures that the optimizer's selection pressure is consistent with that evaluation mode. The P95 term receives equal weight to the curvature term (0.25 each) to penalize large localized deviations, errors that the RMS alone would underweight, while preserving sensitivity to shape-level features such as bending turns. The tip error is assigned the smallest weight (0.20) because the tip position is already captured partially by both the RMS and P95 terms; a separate term is included to retain an explicit signal on distal placement without over-penalizing tip discrepancy.

4.7.5 Interpretation and Scope

The stiffness sweep and Bayesian optimization module is a parameter-search tool, not an evaluation method in its own right. Its output is a ranked set of simulated candidate configurations, of which the best-matching is carried forward for metric analysis using the comparison framework of Section 3.2. The interpretation of whether the best-found parameters constitute an acceptable match to the physical reference, and what the residual discrepancy implies about the simulator's mechanical model, is addressed in the Results and Discussion chapters.

4.7.6 Physics Model Validation Protocol

The optimization experiments described above are not only used to find stiffness parameters that reproduce a given catheter configuration; they are also structured as a protocol for evaluating whether the VIST physics model generalizes correctly when instrument components are varied independently. The central question is whether stiffness parameters recovered by optimizing against one instrument combination remain predictive when a different, physically distinct instrument is introduced without re-optimization. If the simulator's physics model is faithful, the mechanical response of one component to a change in another should follow naturally from the underlying equations of motion, rather than requiring the parameters to be re-calibrated for every new combination.

Because all stiffness values in VIST are relative scaling factors rather than absolute physical quantities (Section 4.7.1), one guidewire stiffness value must be fixed arbitrarily to anchor the scale. In this protocol, $\mathrm{GW}_{\mathrm{ref}}$ is assigned a fixed stiffness of $[-1, 6]$ in Stages 1 and 3, where it serves as the anchor against which the catheter stiffness is optimized, and is never itself optimized. The choice of $[-1, 6]$ is arbitrary as any fixed value would serve the same purpose, since what the optimizer recovers are ratios between component stiffnesses, not their absolute magnitudes.

In all stages, reference images are captured using the Raspberry Pi camera setup described in Section 4.1.1, with the physical instrument placed in our setup. Stage 4 therefore constitutes a direct test of whether the simulator reproduces real-world catheter behavior under a novel instrument combination, rather than a self-consistency check within the simulation environment.

The protocol is carried out across two catheters with different shapes (Catheter A and Catheter B) and two guidewires ($\mathrm{GW}_{\mathrm{ref}}$ and $\mathrm{GW}_{\mathrm{test}}$, where $\mathrm{GW}_{\mathrm{test}}$, in our case, is a physically stiffer variant of $\mathrm{GW}_{\mathrm{ref}}$, confirmed by tactile assessment):

1. **Catheter A optimization with $\mathbf{GW}_{\mathbf{ref}}$.** $\mathrm{GW}_{\mathrm{ref}}$ is inserted to the tip of Catheter A and a reference image is captured with the Raspberry Pi camera. The Bayesian optimizer searches the catheter stiffness parameter space to find the configuration that best reproduces this image in VIST, with $\mathrm{GW}_{\mathrm{ref}}$ held at its fixed anchor value of $[-1, 6]$. The resulting stiffness parameters for Catheter A are fixed for all subsequent stages.
2. **$\mathbf{GW}_{\mathbf{test}}$ optimization with Catheter A.** $\mathrm{GW}_{\mathrm{ref}}$ is replaced with $\mathrm{GW}_{\mathrm{test}}$ while the optimized Catheter A parameters from Stage 1 are retained. A new reference image is captured and the optimizer searches only the guidewire

stiffness parameter space. The recovered stiffness value for $\mathrm{GW}_{\mathrm{test}}$ is noted and fixed for Stage 4.

3. **Catheter B optimization with $\mathbf{GW}_{\mathbf{ref}}$.** A physically different Catheter B is inserted with $\mathrm{GW}_{\mathrm{ref}}$ and a reference image is captured. The optimizer searches the catheter stiffness space independently of Catheter A, yielding a separate optimized parameter set for Catheter B. $\mathrm{GW}_{\mathrm{ref}}$ is again held at its fixed anchor value of $[-1, 6]$.
4. **Cross-validation: Catheter B with $\mathbf{GW}_{\mathbf{test}}$.** The optimized stiffness parameters from Stage 3 (Catheter B) and Stage 2 ($\mathrm{GW}_{\mathrm{test}}$) are combined without any further optimization, and the simulator is run under this combined configuration. A reference image of Catheter B with $\mathrm{GW}_{\mathrm{test}}$ inserted is captured with the Raspberry Pi camera, and the resulting simulated image is compared against it using the full comparison pipeline. RMS correspondence error and visual centerline alignment are the primary evaluation criteria.

The logic of this protocol is that each component's stiffness parameters are estimated from an experiment in which only that component varies. Stage 4 then tests whether the simulator correctly predicts the emergent behavior of a novel combination assembled from those independently estimated parameters. A low correspondence error in Stage 4 would indicate that the physics model composes instrument interactions correctly, while a high error would suggest that the model's response to a given catheter depends on the specific guidewire present during parameter estimation, implying that the stiffness parameters are not truly transferable across combinations. This is therefore a test of the compositional fidelity of the simulator's mechanical model rather than of its ability to fit a single configuration.

5

Results

This chapter presents the results obtained in this thesis. The findings are organized according to the main components of the proposed framework and the defined evaluation metrics.

5.1 Estimated Camera Parameters

The estimated intrinsic and distortion parameters are summarized in Table 5.1, together with the mean reprojection error.

Table 5.1: Summary of calibration results.

Quantity	**Value**
f_x	4117.49 px
f_y	4111.04 px
c_x	2296.80 px
c_y	1310.74 px
$(k_1, k_2, p_1, p_2, k_3)$	$(0.01796,\ 0.29686,\ 0.00358,\ 0.00057,\ -0.78619)$
Mean reprojection error	0.240 px

5.2 Validation of the Algorithms

This section presents the validation results for the algorithms developed in this work. The evaluation focuses on the accuracy and robustness of the catheter length estimation and alignment methods used throughout the comparison framework.

5.2.1 Length Estimation

The length estimation revealed a strong agreement with reference values, with a mean absolute error of 0.117 ± 0.107 cm and a mean relative error of 1.89% (range: -3.67% to 7.00%). The measured lengths clustered closely around the ideal 1:1 relationship line with ($R^2 > 0.99$).

Absolute errors ranged from -0.23 cm to $+0.18$ cm, indicating both positive and negative deviations from reference values. Performance varied across measurement ranges:

- **Small measurements** (1-5 cm): Mean relative error of 1.7%, with higher variability.
- **Large measurements** (>15 cm): Mean relative error of 0.34%, demonstrating improved accuracy for extended catheter segments.

Bland-Altman analysis revealed a mean bias of +0.035 cm, indicating that measured values slightly overestimate real length. The 95% limits of agreement were −0.197 cm to +0.266 cm, and the bias distribution remained approximately symmetric across the measurement range.

Figure 5.1 provides a visual summary of the validation results, including the real-versus-measured comparison, error distributions, and Bland-Altman agreement analysis.

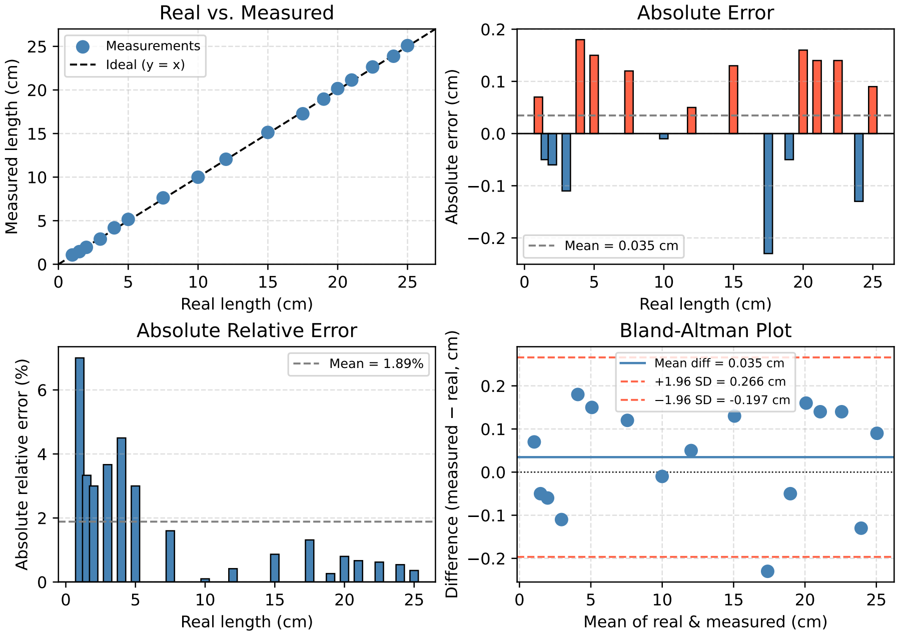


Figure 5.1: Measurement validation analysis. **(a)** Real versus measured catheter length comparison showing measurements relative to the ideal 1:1 line. **(b)** Absolute error distribution across the measurement range with mean error indicated. **(c)** Relative error percentage showing proportional deviation from true measurements. **(d)** Bland-Altman plot displaying agreement between methods with mean bias and 95% limits of agreement indicated by horizontal lines.

Figure 5.2 shows an example used for validation of the catheter length estimation based on the extracted centerline. In this example, the catheter has a known physical length of 10 cm, allowing the estimated length obtained from the algorithm to be compared against the ground truth.

5.2.2 Measurement Consistency Across Shape Variations

The robustness of the algorithm was further evaluated using a single catheter with a constant. The resulting output lengths and corresponding absolute errors for each

Figure 5.2: Example image illustrating validation of the catheter length estimation based on the extracted centerline. The catheter shown in the image has a physical length of 10 cm.

shape are summarized in Table 5.2.

Table 5.2: Algorithm output for a 20 cm catheter under different bending shapes

Configuration	Actual Length (cm)	Measured Length (cm)	Error (cm)
Shape 1	20	20.24	+0.24
Shape 2	20	20.00	0.00
Shape 3	20	19.93	-0.07
Shape 4	20	20.05	+0.05
Shape 5	20	20.12	+0.12
Shape 6	20	20.13	+0.13
Mean (± SD)	**20**	**20.08 ± 0.11**	**0.10 (MAE)**

The algorithm produced a mean measured length of 20.08 cm across all trials, with a standard deviation of 0.11 cm. The Mean Absolute Error (MAE) was calculated at 0.10 cm. The recorded measurements ranged from a minimum of 19.93 cm to a maximum of 20.24 cm, representing a maximum relative error of 1.2% across the tested deformations.

5.2.3 Alignment

The alignment procedure produced the transformation parameters summarized in Table 5.3.

As shown in Figure 5.4, the centerlines from Figure 5.3 are initially misaligned due to rotational differences. Despite this initial discrepancy, the alignment method successfully compensates for the transformation, resulting in a close correspondence

Table 5.3: Transformation parameters obtained from the alignment procedure.

Parameter	**Value**
Offset_x	-47.4169
Offset_y	90.3882
Rotation	$-62.9570°$
Scale	0.991412
$\text{RMS}_{\text{before}}$	229.7905
$\text{RMS}_{\text{after}}$	1.0422

between the two centerlines.

The aligned centerline demonstrates a strong agreement with the reference, indicating that the proposed alignment approach is effective in correcting geometric differences.

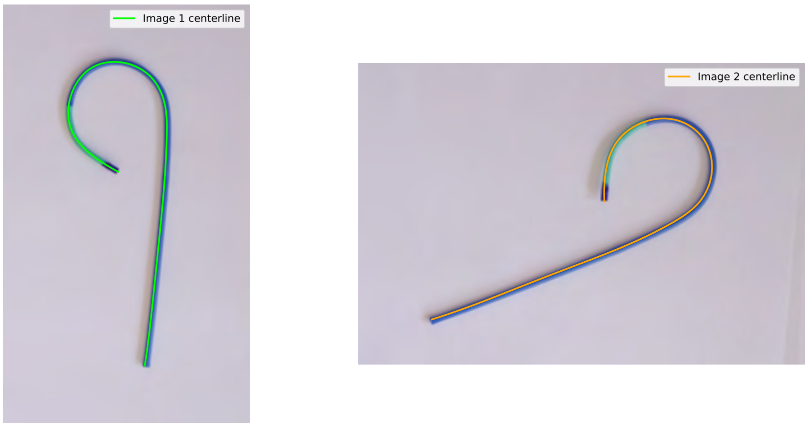


Figure 5.3: Extracted centerlines from the two images shown in Figure 4.5.

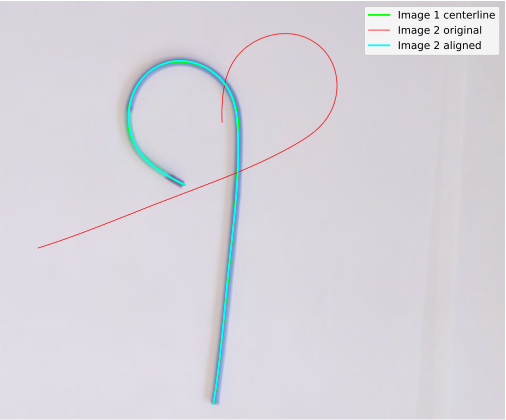


Figure 5.4: Shows the centerlines from figure 5.3, as well as the resulting aligned centerline.

5.3 Physical SIM vs. VIST

Figure 5.5 shows the resulting images of a microcatheter in both the physical SIM environment and the VIST simulation system. The comparison was performed to enable analysis of catheter behavior and positioning within the vessel geometry.


Reference image (Length: 123.8 mm)

Simulated image (Length: 123.8 mm)

Figure 5.5: Resulting images of a microcatheter in the physical SIM environment and the VIST simulation system.

Figure 5.6 shows the same images as Figure 5.5, together with their corresponding extracted centerlines.


Reference image (Length: 123.8 mm)

Simulated image (Length: 123.8 mm)

Figure 5.6: Resulting images of a microcatheter in the physical SIM environment and the VIST simulation system, together with their corresponding extracted centerlines.

Figure 5.7 shows the aligned overlay of the centerlines presented in Figure 5.6.

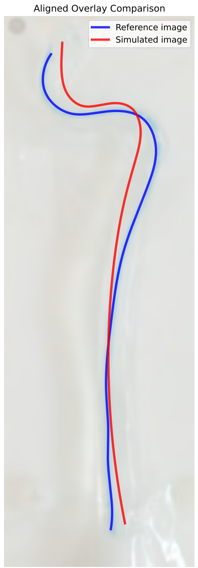


Figure 5.7: Aligned overlay of the centerlines shown in Figure 5.6.

5.4 Stiffness Sweep and Bayesian Optimization

Figure 5.8 shows the real and simulated images used for comparison of the Medtronic JR4 8F guiding catheter together with a thin and stiff guidewire. In this experiment, a guidewire stiffness configuration of $[-1, 6]$ was used, while the catheter stiffness parameters were optimized by the proposed algorithm to $[[-1, 6.0], [5.3, 6.7], [4.7, 6.2]]$ to achieve the best correspondence between the simulated and real catheter shapes.

Figure 5.8: Real and simulated images used for comparison of the JR4 8F guiding catheter from Medtronic and an accompanying thin and stiff guidewire.

Figure 5.9 shows the resulting aligned overlay of the centerlines extracted from the real and simulated images. As can be observed, the simulated catheter geometry follows the real catheter shape closely, indicating that the optimized stiffness parameters provide a good correspondence between the physical and simulated configurations.

Figure 5.9: Aligned overlay comparison of the centerlines extracted from the real and simulated images of the JR4 8F guiding catheter from Medtronic and the accompanying thin and stiff guidewire.

In Figure 5.10, the guidewire was replaced with a thicker and stiffer variant. Using the previously optimized catheter stiffness configuration, the guidewire stiffness parameters were optimized to $[-1, 6.9]$ to achieve the best correspondence between the simulated and real images.

Figure 5.11 shows the resulting aligned overlay of the centerlines extracted from the real and simulated images using the thicker and stiffer guidewire. As can be observed, the simulated catheter geometry follows the real catheter shape closely, indicating that the optimized stiffness parameters provide a good correspondence between the physical and simulated configurations.

In figure 5.12 shows the real and simulated images used for comparison of the Cordis XBLAD 4 8F guiding catheter together with a thin and stiff guidewire. In this experiment, a guidewire stiffness configuration of $[-1, 6]$ was used, while the catheter stiffness parameters were optimized by the proposed algorithm to $[[-1, 6.3], [2.6, 7.9], [1.3, 5.6]]$ to achieve the best correspondence between the simulated and real catheter shapes.

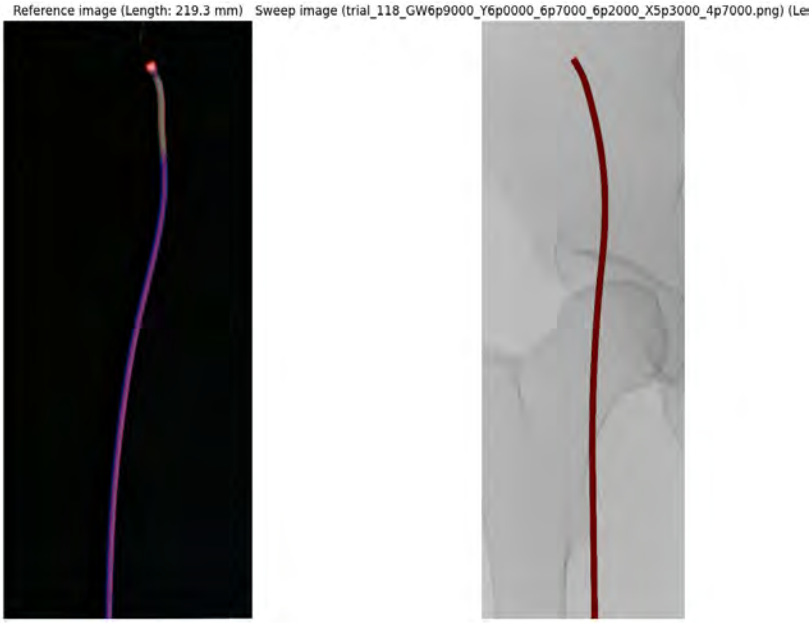


Figure 5.10: Real and simulated images used for comparison of the JR4 8F guiding catheter from Medtronic and an accompanying thick and stiff guidewire.

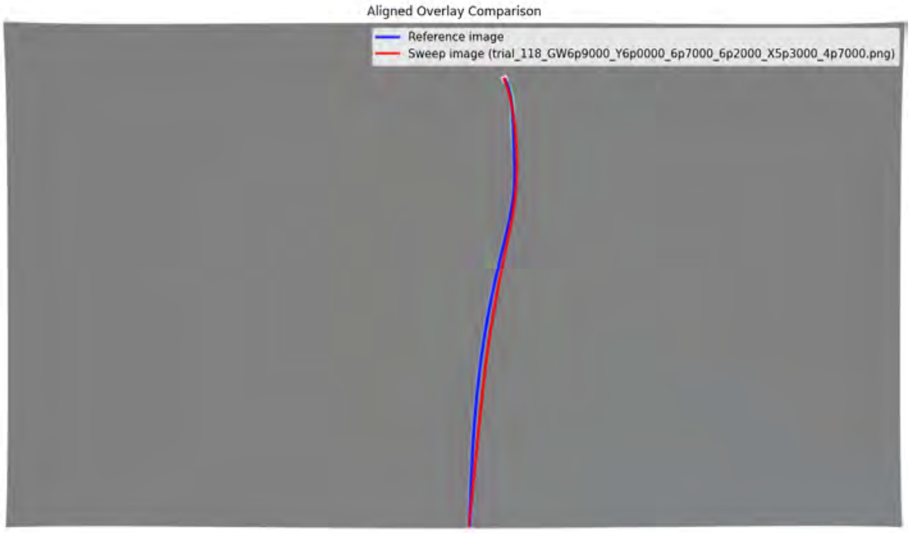


Figure 5.11: Aligned overlay comparison of the centerlines extracted from the real and simulated images of the JR4 8F guiding catheter from Medtronic and the accompanying thick and stiff guidewire.

Figure 5.13 shows the resulting aligned overlay of the centerlines extracted from the real and simulated images. As can be observed, the simulated catheter geometry follows the real catheter shape closely, indicating that the optimized stiffness parameters provide a good correspondence between the physical and simulated con-

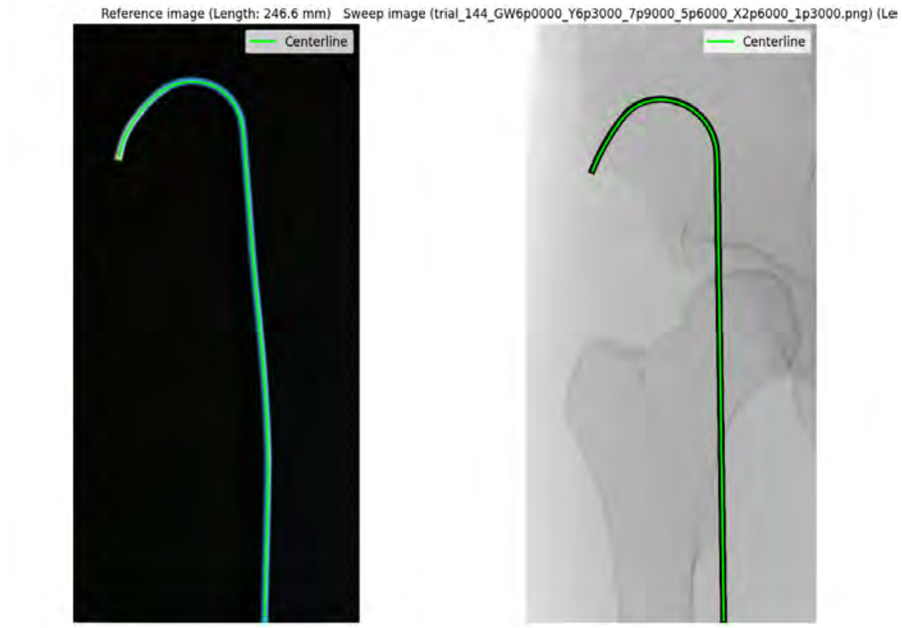


Figure 5.12: Real and simulated images used for comparison of the XBLAD 4 8F guiding catheter from Cordis and an accompanying thin and stiff guidewire.

figurations.

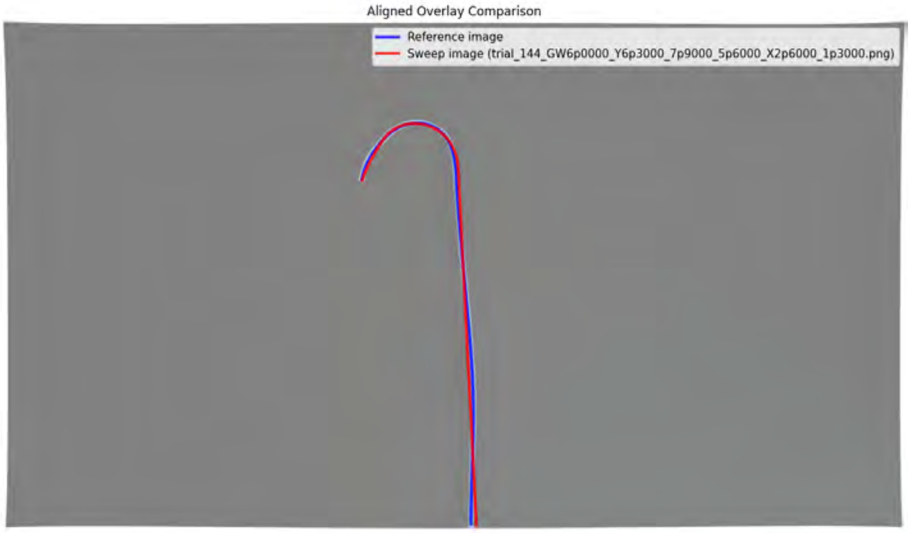


Figure 5.13: Aligned overlay comparison of the centerlines extracted from the real and simulated images of the XBLAD 4 8F guiding catheter from Cordis and the accompanying thin and stiff guidewire.

In Figure 5.14, the guidewire was replaced with a thicker and stiffer variant. The previously optimized catheter stiffness configuration $[[-1, 6.3], [2.6, 7.9], [1.3, 5.6]]$ together with the guidewire configuration $[-1, 6.9]$ was used to evaluate whether the simulation could reproduce a catheter shape similar to that observed in the real image.

Figure 5.14: Real and simulated images used for comparison of the XBLAD 4 8F guiding catheter from Cordis and an accompanying thick and stiff guidewire.

Figure 5.15 shows the resulting aligned overlay of the centerlines extracted from the real and simulated images. In contrast to the previous experiments, the simulated catheter geometry does not closely follow the shape of the real catheter. This indicates that the previously optimized stiffness parameters are less suitable for reproducing the behavior of this catheter-guidewire combination, highlighting the sensitivity of the catheter shape to variations in both catheter and guidewire properties.

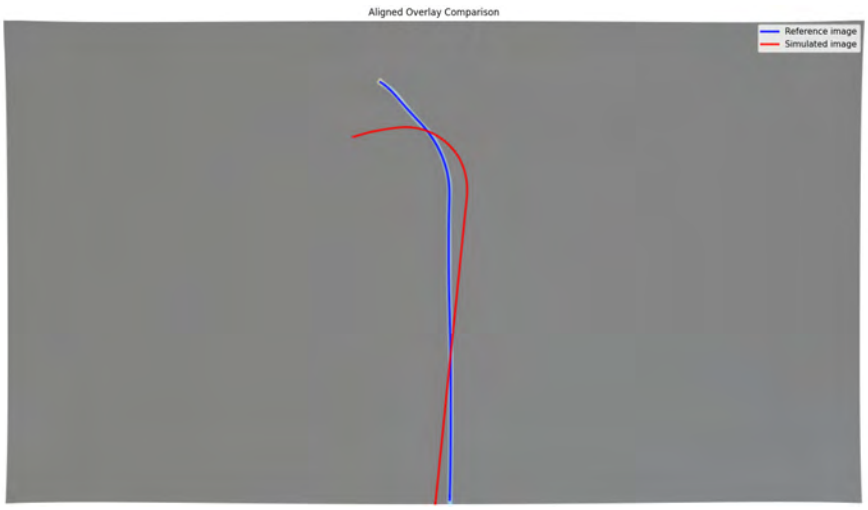


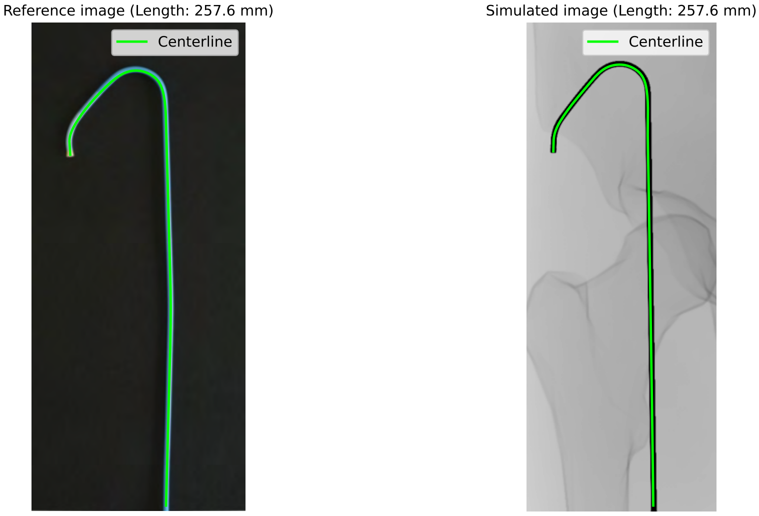
Figure 5.15: Aligned overlay comparison of the centerlines extracted from the real and simulated images of the XBLAD 4 8F guiding catheter from Cordis and the accompanying thick and stiff guidewire.

5.5 Evaluation of Suitable Metrics

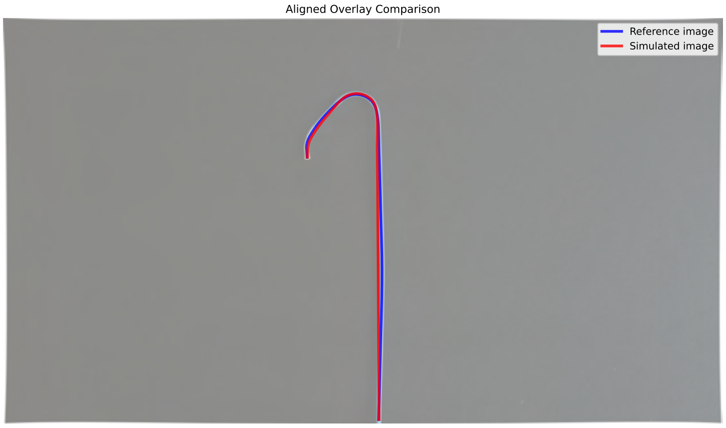
To evaluate which accuracy metrics are most suitable for comparing physical and simulated catheter configurations, two trials from the Bayesian optimization sweep are presented for the Cordis Vistabrite Tip 8F JL4 catheter - one representing the best-matching trial and one representing a poor-matching trial. By comparing the metric outputs across both cases, the discriminative ability of each metric can be assessed.

5.5.1 Optimized Trial

Figure 5.16 shows the reference and simulated images alongside the aligned centerline overlay. Visually, the two centerlines follow a similar overall trajectory, with the main discrepancy concentrated in the proximal curved region near the catheter tip.



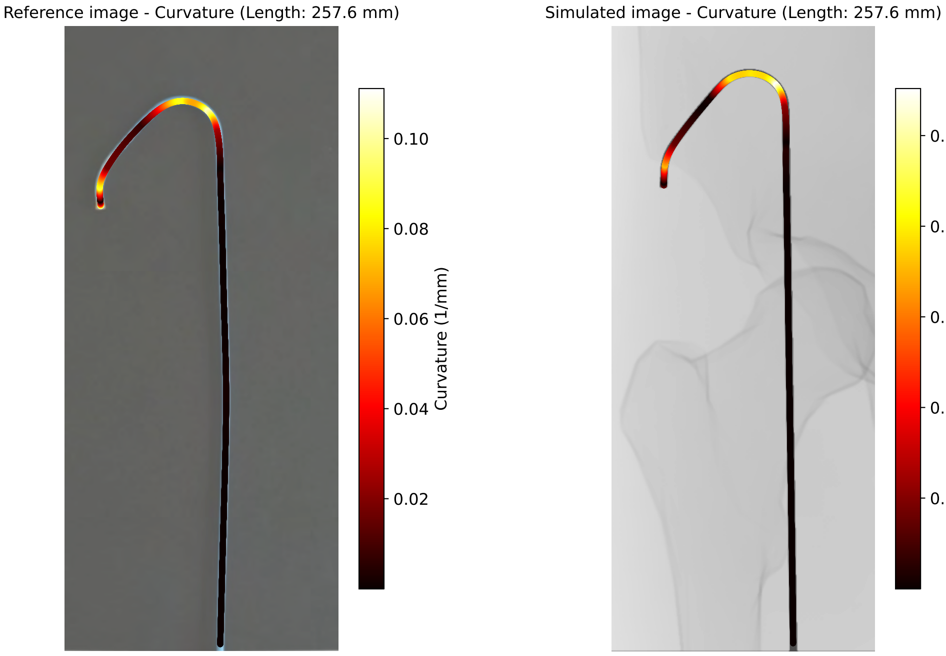
(a) Reference and simulated images.



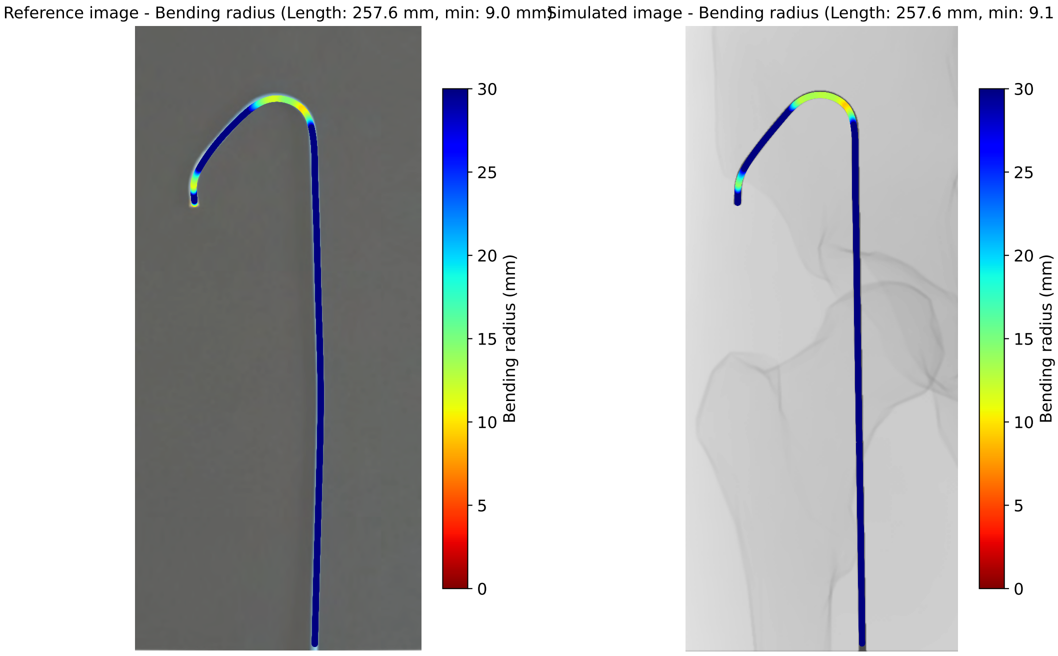
(b) Aligned overlay comparison.

Figure 5.16: Optimized trial - reference and simulated images with aligned centerline overlay.

Figure 5.17 shows the curvature and bending radius heatmaps overlaid on the reference and simulated images. The highest curvature is concentrated at the tip in both cases, with the straight shaft exhibiting near-zero curvature. The minimum bending radius was 9.0 mm for the physical catheter and 10.8 mm for the simulated one, indicating that the simulator produces a marginally less tight bend at the tip.



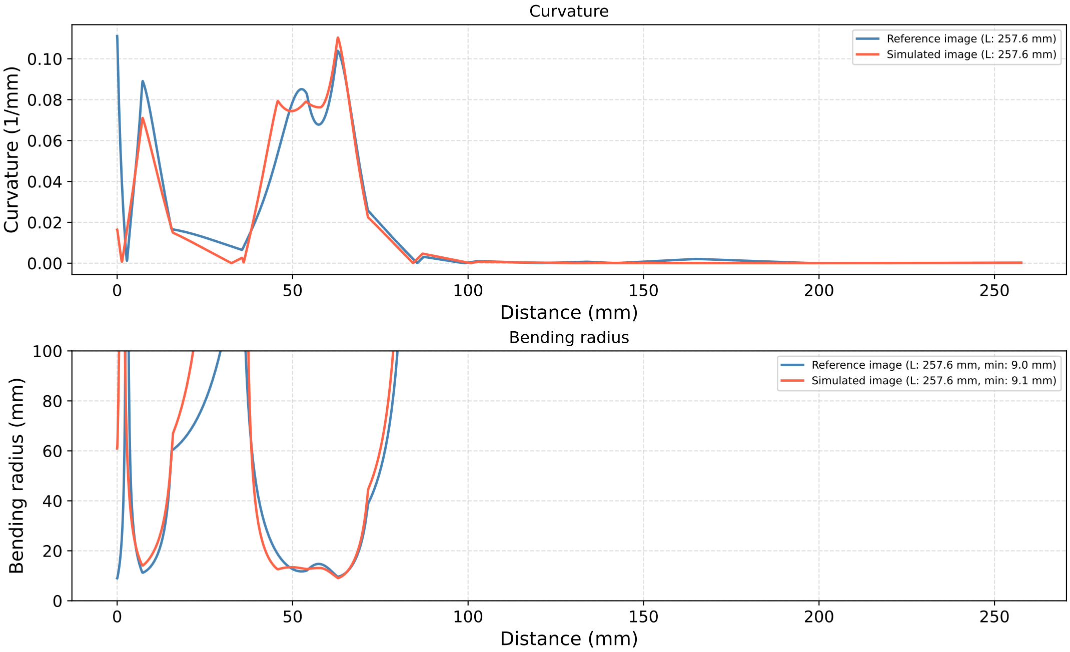
(a) Curvature heatmap.



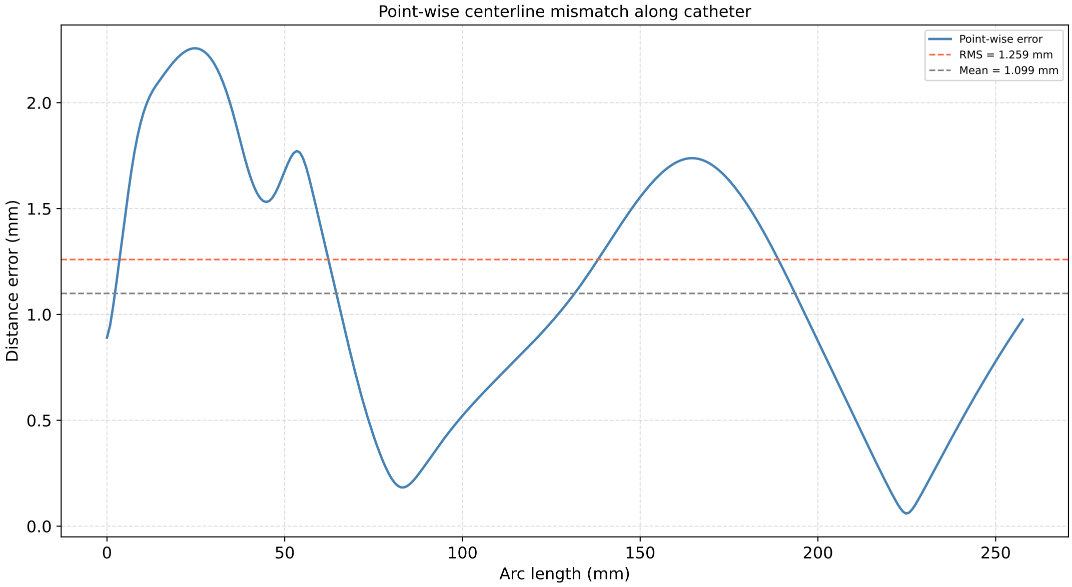
(b) Bending radius heatmap.

Figure 5.17: Optimized trial - curvature and bending radius heatmap overlays on reference and simulated images.

Figure 5.18 shows the curvature profiles and point-wise RMS error along the normalized arc-length. The curvature profiles track each other closely along the straight shaft but diverge at the proximal tip region, consistent with the bending radius difference observed above. The RMS error plot confirms this, showing the largest point-wise deviations in the same proximal region, with an overall RMS of 1.26 mm and a mean distance of 1.10 mm.



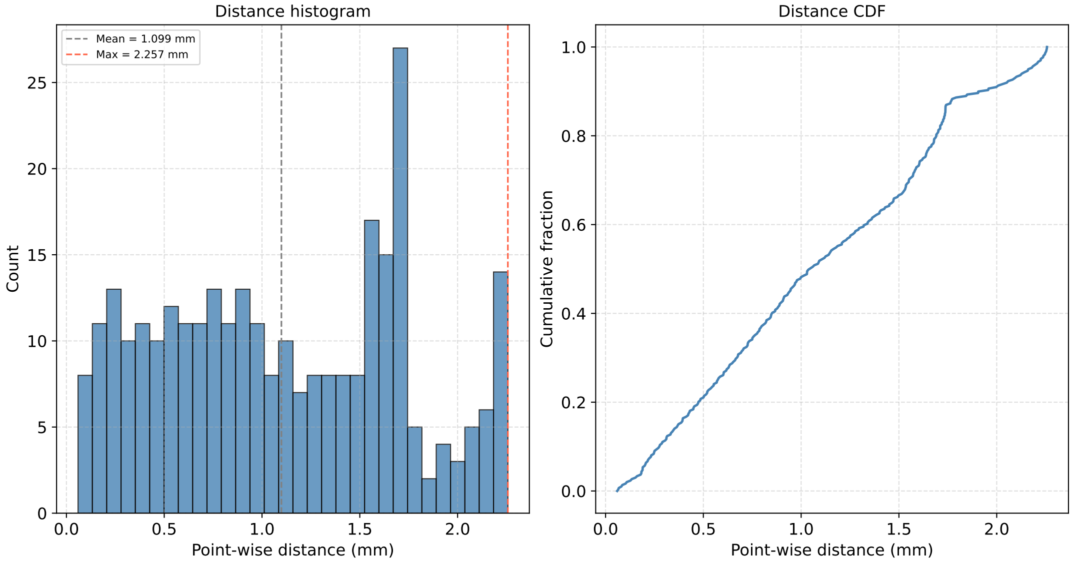
(a) Curvature profiles at matched arc-length positions.



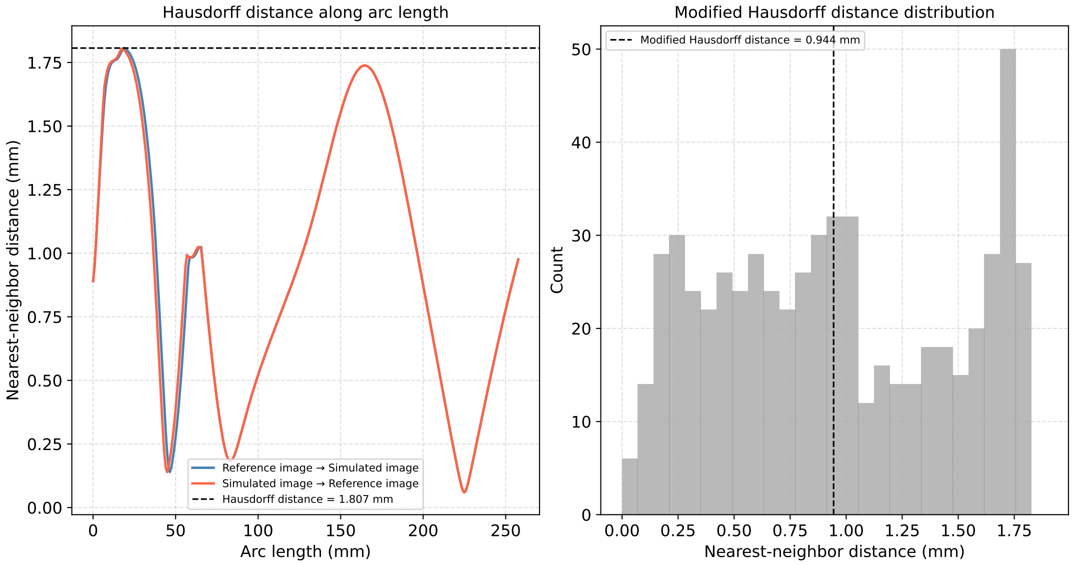
(b) Point-wise centerline mismatch along the catheter.

Figure 5.18: Optimized trial - curvature profile comparison and point-wise RMS error along the catheter.

Figure 5.19 shows the distance histogram and Hausdorff distribution. The Hausdorff distance was 1.81 mm, representing the worst-case local deviation, driven by the proximal curved region identified above.



(a) Distance histogram and CDF.



(b) Hausdorff distance distribution.

Figure 5.19: Optimized trial - distance-based metric distributions.

Figure 5.20 shows the segmentation overlap, yielding a Dice score of 56%.

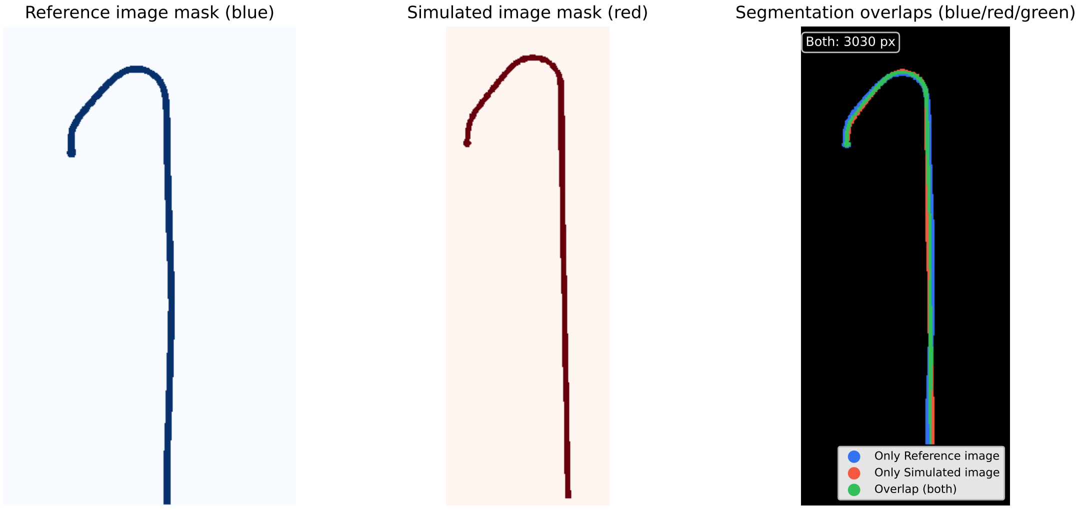
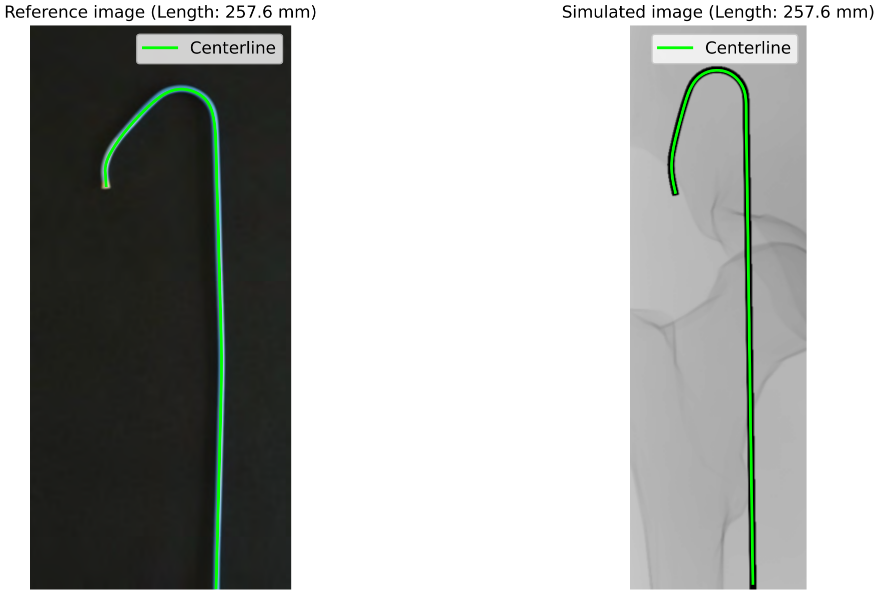


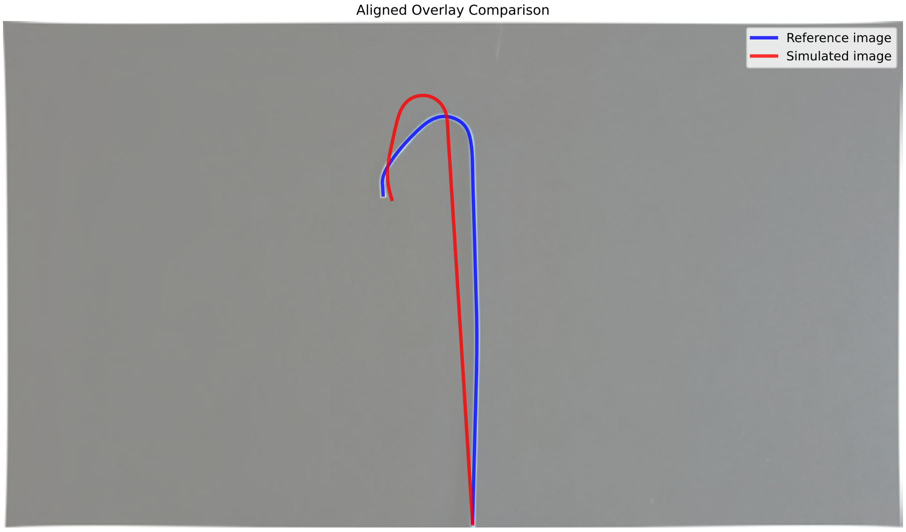
Figure 5.20: Optimized trial - segmentation overlap showing Dice similarity coefficient.

5.5.2 Non-Optimized Trial

Figure 5.21 shows the reference and simulated images for a non-optimized trial of the same catheter. The aligned centerline overlay shows a clear geometric discrepancy between the physical and simulated configurations across the full catheter length.



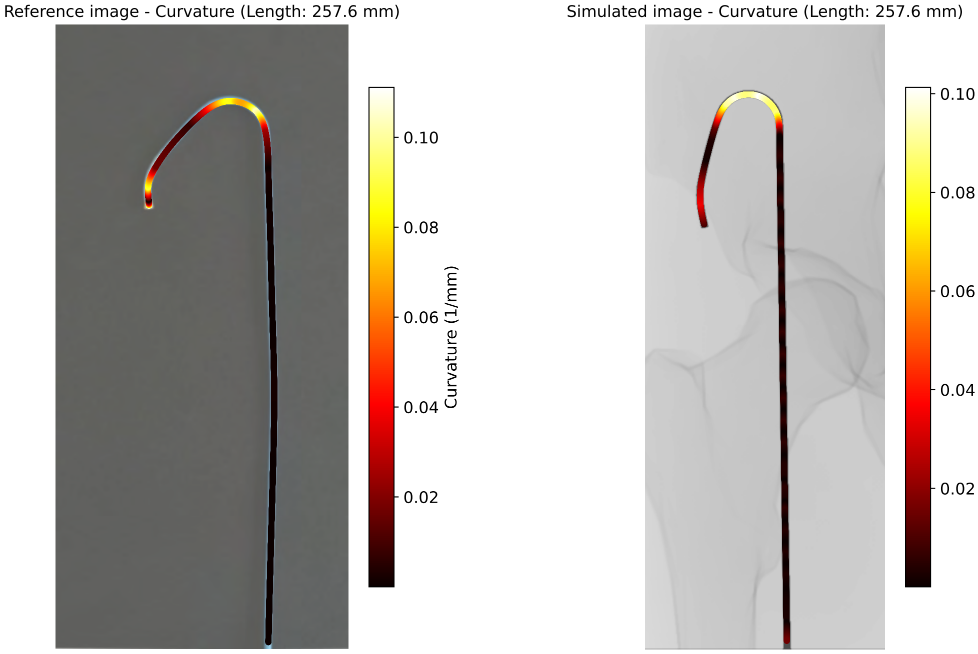
(a) Reference and simulated images.



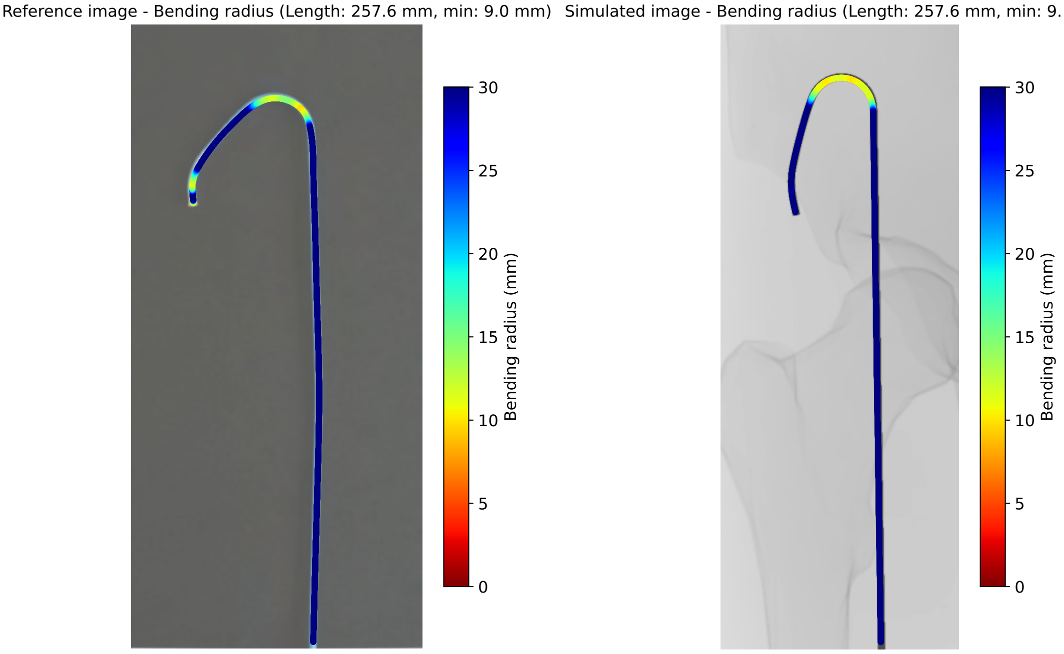
(b) Aligned overlay comparison.

Figure 5.21: Non-optimized trial - reference and simulated images with aligned centerline overlay.

Figure 5.22 shows the curvature and bending radius heatmaps for the non-optimized trial. The spatial distribution of curvature differs markedly between the physical and simulated images, reflecting the incorrect stiffness parameterization.



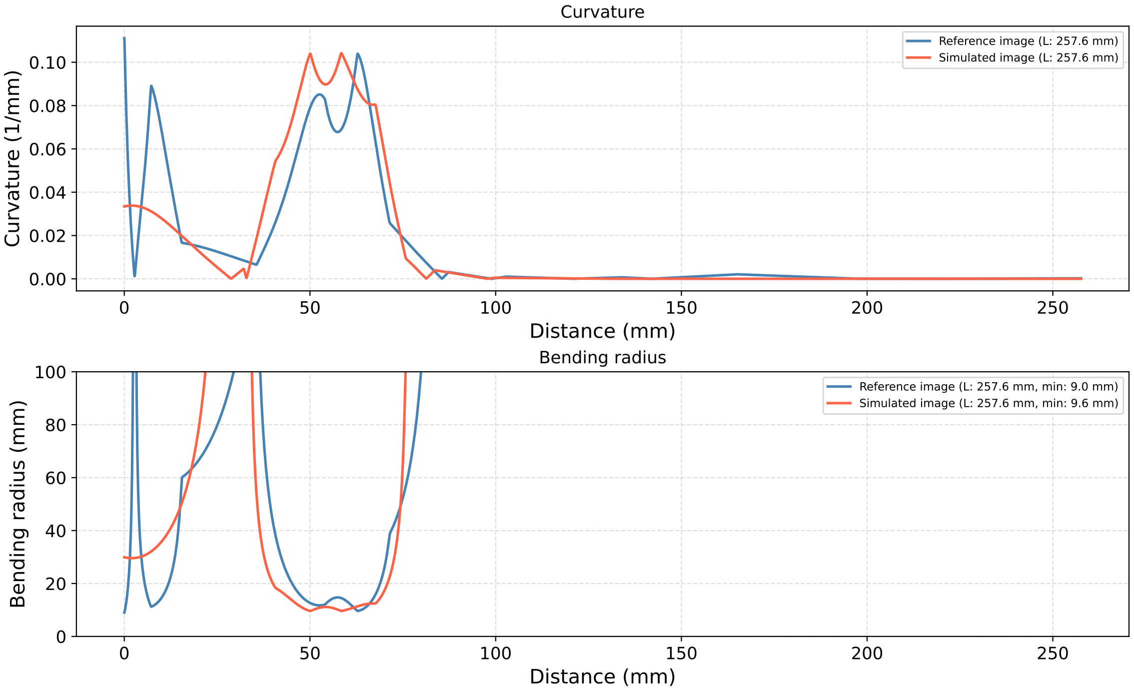
(a) Curvature heatmap.



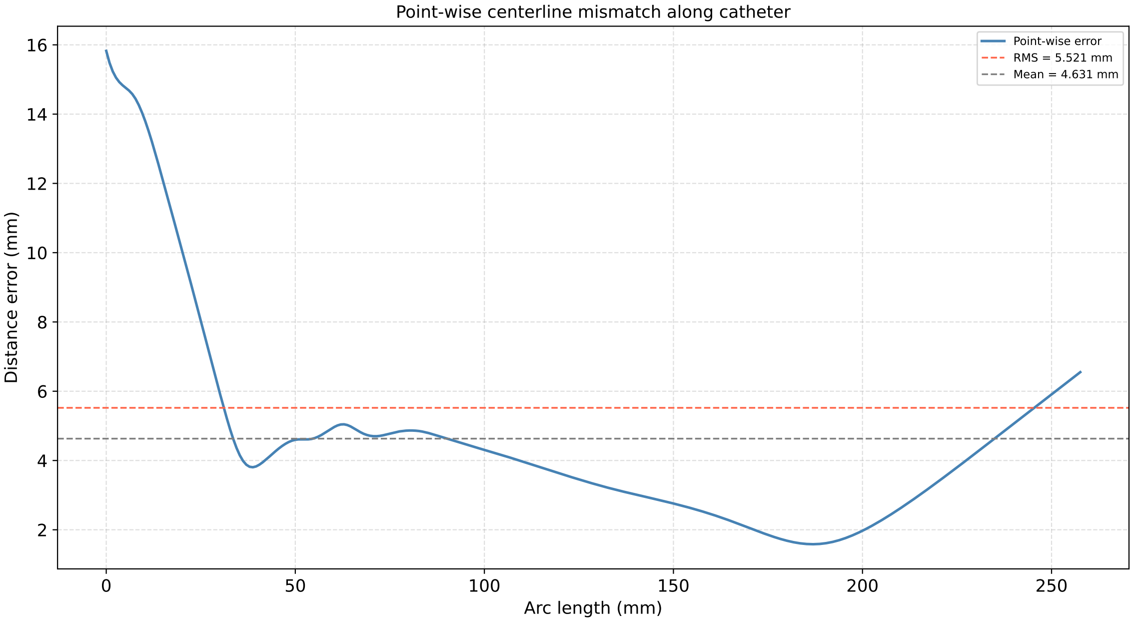
(b) Bending radius heatmap.

Figure 5.22: Non-optimized trial - curvature and bending radius heatmap overlays on reference and simulated images.

Figure 5.23 shows the curvature profiles and point-wise RMS error for the non-optimized trial. The curvature profiles diverge significantly along the full catheter length, and the RMS error plot shows substantially elevated point-wise deviations, with an overall RMS of 5.52 mm and a mean distance of 4.63 mm.



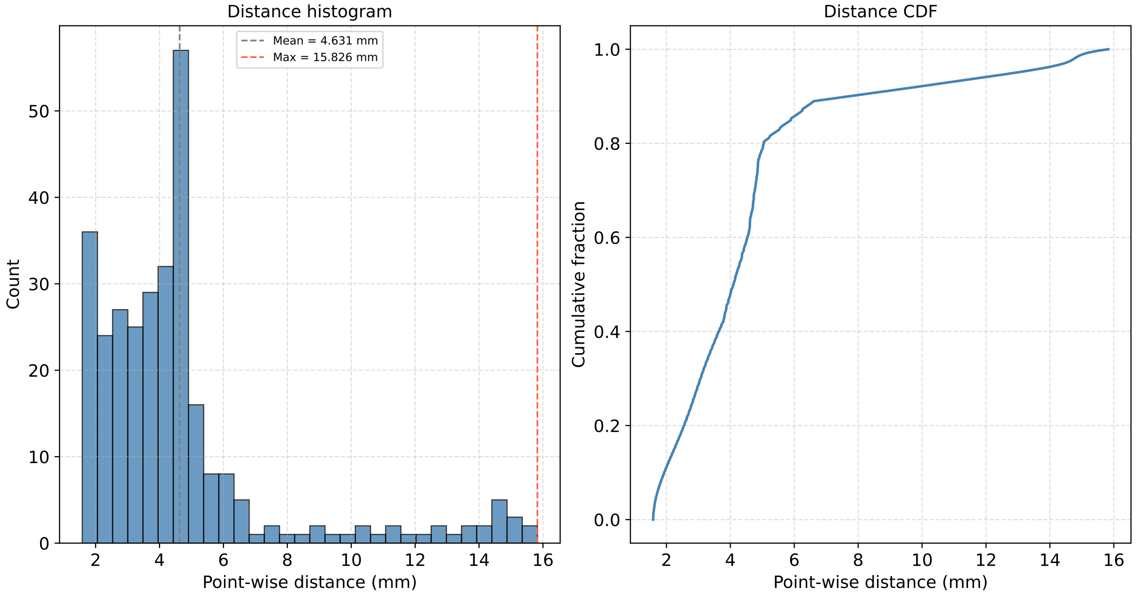
(a) Curvature profiles at matched arc-length positions.



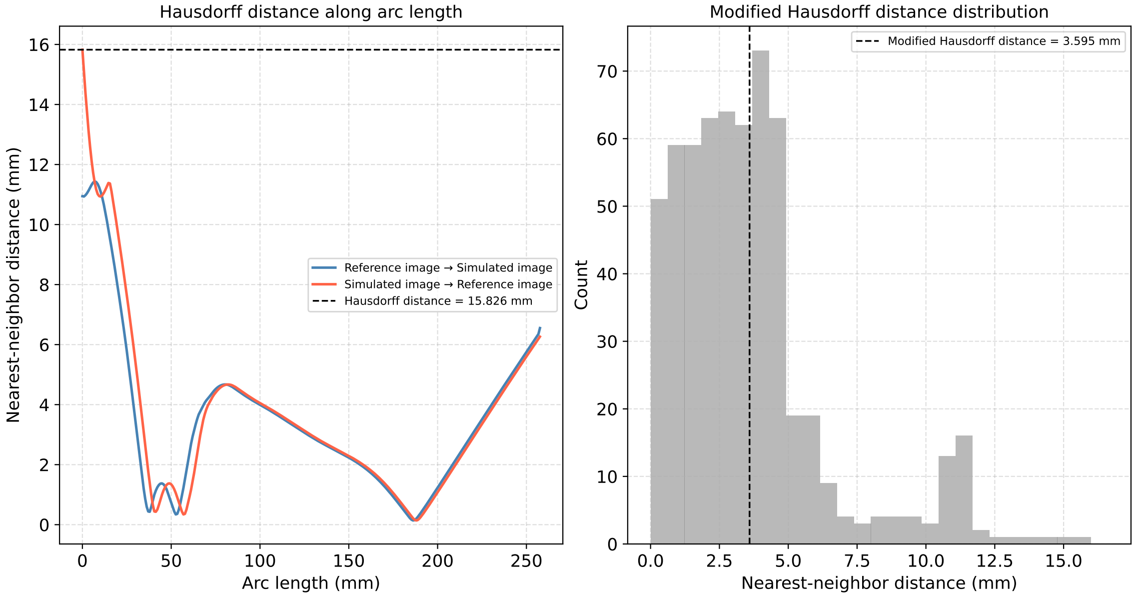
(b) Point-wise centerline mismatch along the catheter.

Figure 5.23: Non-optimized trial - curvature profile comparison and point-wise RMS error along the catheter.

Figure 5.24 shows the distance histogram and Hausdorff distribution for the non-optimized trial. The Hausdorff distance of 15.83 mm is substantially larger than in the optimized trial, reflecting the severe worst-case deviation introduced by the incorrect stiffness parameters.



(a) Distance histogram and CDF.



(b) Hausdorff distance distribution.

Figure 5.24: Non-optimized trial - distance-based metric distributions.

Figure 5.25 shows the segmentation overlap for the non-optimized trial, yielding a Dice score of 11%.

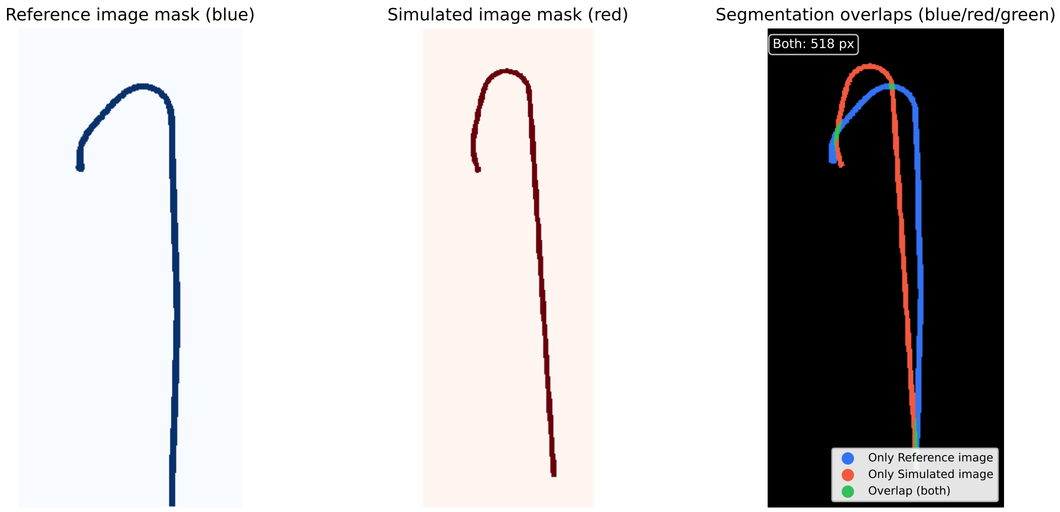


Figure 5.25: Non-optimized trial - segmentation overlap showing Dice similarity coefficient.

Table 5.4 summarizes the metric values for both trials. All metrics clearly distinguish between the two cases, with the optimized trial showing substantially lower error across all measures.

Metric	Optimized trial	Non-optimized trial
RMS (mm)	1.26	5.52
Hausdorff (mm)	1.81	15.83
Modified Hausdorff (mm)	0.94	3.60
Average Distance (mm)	1.10	4.63
Max Distance (mm)	2.26	15.83
Dice	0.56	0.11

Table 5.4: Comparison of accuracy metrics between the optimized and non-optimized trials for the Cordis Vistabrite Tip 8F JL4 catheter.

5.5.3 Metric Analysis Across All Trials

Figure 5.26 shows all six metrics plotted across the 225 trials, ordered by their composite optimization score. The distance-based metrics: RMS, Hausdorff, Modified Hausdorff, Average Distance, and Maximum Distance follow a similar trend across the ranked trials, while Dice oscillates.

Figure 5.27 shows the correlation matrix between all metrics across the 225 trails. RMS, Average Distance, and Maximum Distance are almost perfectly correlated ($r > 0.99$), while Hausdorff and Modified Hausdorff also show very high correlation with RMS ($r = 0.97$ and $r = 0.92$ respectively). The Dice score shows only a weak negative correlation with RMS ($r = -0.32$).

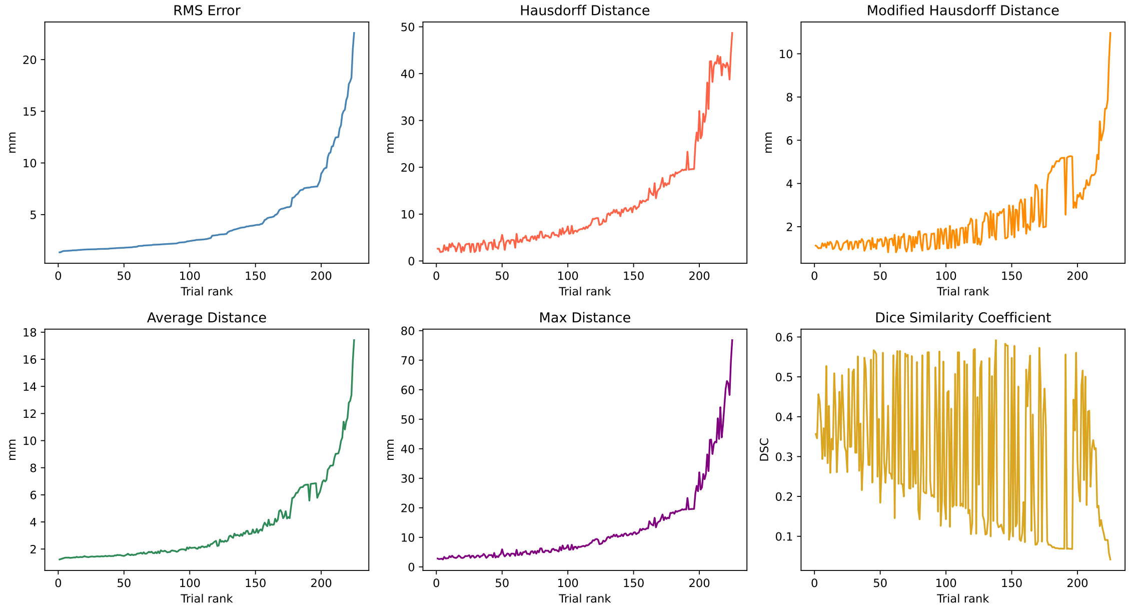
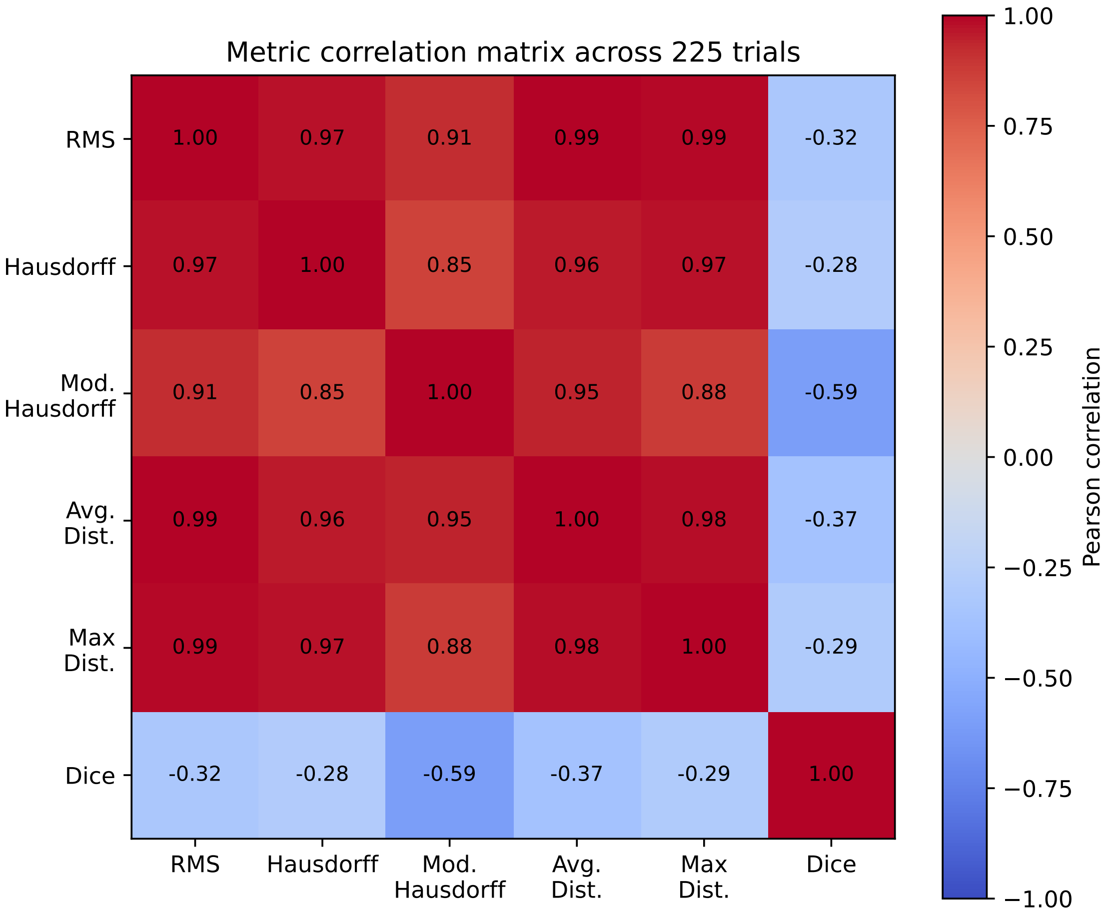


Figure 5.26: Metric values across 225 trials ordered by composite optimization score.



	RMS	Hausdorff	Mod. Hausdorff	Avg. Dist.	Max Dist.	Dice
RMS	1.00	0.97	0.91	0.99	0.99	-0.32
Hausdorff	0.97	1.00	0.85	0.96	0.97	-0.28
Mod. Hausdorff	0.91	0.85	1.00	0.95	0.88	-0.59
Avg. Dist.	0.99	0.96	0.95	1.00	0.98	-0.37
Max Dist.	0.99	0.97	0.88	0.98	1.00	-0.29
Dice	-0.32	-0.28	-0.59	-0.37	-0.29	1.00

Figure 5.27: Correlation matrix between all metrics across the 225 trails.

6

Discussion

This chapter discusses the results obtained from the proposed image-based framework and evaluates the strengths and limitations of the developed methods. The findings from the validation experiments, the comparison between the physical and virtual simulation environments, and the stiffness optimization experiments are interpreted in relation to the research questions and the overall aim of the thesis. In addition, limitations of the current framework and possible directions for future work are discussed.

6.1 Image-based parameterization of instruments

At an early stage of the project, an approach based on Scale-Invariant Feature Transform (SIFT) features was explored to identify the corresponding patterns between the captured images and the simulated catheter images. However, this approach was not pursued further, as it proved difficult to extract a sufficient number of reliable keypoints. The images primarily consisted of a uniformly colored background with a single catheter, resulting in limited texture and distinct features. Consequently, the SIFT algorithm was unable to detect enough meaningful feature points for robust matching.

The image-based parameterization approach demonstrated improved performance, allowing for reliable estimation of the catheter length in both simulated and real images.

6.2 Validation of the Algorithms

The validation results indicate that the proposed algorithms provide sufficiently accurate and consistent performance for image-based comparison of the physical and simulated catheter configurations. Crucially, these results establish that the downstream comparison metrics rest on a reliable geometric foundation. Errors introduced at the extraction and alignment stage are small enough not to confound the shape comparison results.

The length estimation algorithm achieved a mean absolute error of 0.117 cm across seventeen catheter lengths ranging from 1.0 to 25.0 cm, with $R^2 > 0.99$ against the physical reference. The error is predominantly fixed in absolute terms rather than proportional to length, a behavior consistent with pixel-level uncertainty in skeleton endpoint detection, which introduces a constant arc-length offset regardless of catheter size. As a result, relative error grows for shorter catheters: below 15 cm the fixed offset represents a non-negligible fraction of total length, while above 15 cm it becomes negligible. For the catheters relevant to this study, where shape comparison deviations are expected on the order of millimeters to centimeters, this level of length estimation error does not meaningfully affect the downstream comparison results.

The length estimates also remained consistent across the six bending configurations tested, indicating that catheter curvature and geometric deformation do not meaningfully affect the accuracy of the method. This robustness is attributed to two successive noise suppression stages in the pipeline: the Frangi vesselness filter enhances tubular structure and suppresses background noise prior to skeletonization, and the subsequent B-spline fitting regularizes the extracted centerline before arc-length is integrated. Together these stages absorb the local skeleton irregularities that curved geometries would otherwise introduce, making the length estimate largely insensitive to catheter shape.

The alignment algorithm successfully compensates for rotational and translational differences between the compared centerlines, resulting in a close geometric correspondence after alignment.

6.3 Physical SIM vs. VIST

A comparison was performed between the physical simulator (SIM) and the VIST simulator to evaluate how closely the behavior of the catheter and guidewire in the virtual environment corresponded to that observed in the physical setup. However, several known limitations affect the comparability between the two systems. The vascular structures in the physical SIM consist of relatively compliant and deformable materials, whereas the vessels in VIST are modeled as rigid structures that do not deform in response to catheter or guidewire interaction. As a result, mechanical interactions between the instruments and the vessel walls differ between the physical and virtual environments, particularly when using stiffer catheters or guidewires. Differences in frictional properties, anatomical representation, and overall instrument behavior further contribute to discrepancies between the systems. Given these fundamental environmental differences, quantitative shape comparison metrics are not reported for this section, as any numerical deviation would reflect representational mismatches between the two environments as much as genuine simulator inaccuracy. The purpose of this comparison is therefore limited to demonstrating that the proposed pipeline operates reliably across both imaging contexts.

The results confirm that the proposed algorithm successfully detects and extracts the catheter centerline in both the physical SIM environment and the VIST simulation system. The aligned overlay shown in Figure 5.7 demonstrates that the alignment procedure produces a geometrically coherent result in both cases, establishing that the framework generalizes across the two distinct imaging conditions. This cross-environment robustness is a necessary prerequisite for the subsequent stiffness optimization experiments, where the framework is applied to more controlled comparisons between physical and simulated catheter configurations under matched conditions.

6.4 Stiffness Sweep and Bayesian Optimization

The stiffness sweep and Bayesian optimization experiments were conducted to investigate whether the mechanical properties of catheters and guidewires could be approximated directly from image data. By comparing physical and simulated catheter configurations, the framework searches for stiffness parameter combinations that reproduce the observed catheter geometry as closely as possible. The results provide insight into both the feasibility and the limitations of estimating simulator parameters from static image-based observations alone.

As described previously, the stiffness sweep experiments were performed using two different catheter types together with two different guidewires, resulting in four unique catheter-guidewire combinations. The stiffness parameters were optimized using three of these combinations, while the fourth combination was evaluated using the previously optimized parameters without additional tuning. This approach made it possible to investigate not only how well the optimization could reproduce known configurations, but also how transferable the optimized parameter sets were to unseen catheter-guidewire combinations.

The results indicate that the optimization framework is capable of identifying stiffness configurations that produce simulated catheter geometries with a high degree of similarity to the corresponding physical configurations in several cases. In particular, the optimized simulations were able to reproduce the overall curvature and spatial trajectory of the catheter centerlines with relatively small geometric deviations.

However, the experiments also revealed limitations in the generalizability of the optimized parameter sets. The non-optimized catheter-guidewire combination showed noticeably larger discrepancies between the physical and simulated configurations, suggesting that stiffness parameters optimized for one setup do not necessarily transfer well to other instrument combinations. This likely reflects the complexity of catheter mechanics, where instrument behavior depends not only on stiffness values, but also on interactions between catheter geometry, guidewire properties, friction, vessel deformation, and boundary conditions within the environment.

Another important limitation is that the optimization relies solely on static image-based observations. Multiple stiffness configurations may produce visually similar catheter shapes in two-dimensional image space, making the parameter estimation problem inherently underdetermined. As a result, the optimization identifies parameter sets that reproduce the observed geometry, but not necessarily the true physical material properties of the instruments.

Despite these limitations, the results demonstrate that the proposed framework provides a practical method for quantitative tuning and evaluation of simulator parameters. The combination of image-based comparison and Bayesian optimization offers a systematic alternative to manual parameter tuning and may provide useful support for future simulator calibration and validation workflows.

6.5 Evaluation of Suitable Metrics

The evaluation of suitable metrics was performed to investigate which quantitative measures are most appropriate for assessing the accuracy and realism of endovascular simulations, since no standardized evaluation methodology currently exists for this domain. The comparison between the optimized and non-optimized trials provided a controlled basis for this analysis, as the two cases represent clearly distinct levels of geometric correspondence and therefore allow the discriminative ability of each metric to be assessed directly.

Among the distance-based metrics, RMS error proved the most informative for characterizing overall geometric correspondence. As shown in Table 5.4, RMS error increased from 1.26 mm in the optimized trial to 5.52 mm in the non-optimized trial, clearly reflecting the difference in simulator parameterization. The point-wise error profiles further demonstrated that RMS error is effective at localizing spatial deviations along the catheter shaft, making it useful not only as a global summary statistic but also as a diagnostic tool for identifying where along the catheter the simulation diverges from the physical reference. However, RMS error alone does not capture differences in local bending behavior, since two centerlines may achieve similar point-wise distances while differing in their curvature distribution.

Curvature-based analysis provided complementary information precisely in these cases. When comparing configurations affected by different guidewire stiffnesses, the curvature heatmaps revealed mechanical differences in local bending behavior that were not fully reflected in the distance-based metrics alone. This makes curvature analysis particularly valuable for evaluating whether the simulator reproduces the mechanical response of the instrument, rather than just its global shape.

The remaining distance-based metrics: Hausdorff distance, Modified Hausdorff distance, Average Distance, and Maximum Distance, all demonstrated strong discriminative ability, clearly separating the optimized and non-optimized trials. However, as shown in Figure 5.26, all distance-based metrics follow a near-identical trend across the 225 trials ranked by composite optimization score, while the Dice coefficient oscillates independently of the others. This redundancy is further confirmed by the correlation matrix in Figure 5.27, where all distance-based metrics are strongly intercorrelated with RMS error ($r > 0.91$ in all cases), indicating that they carry largely redundant information. Among these, Hausdorff distance warrants particular attention: as a worst-case measure it is disproportionately sensitive to isolated outliers or segmentation artifacts, which limits its reliability as a general summary statistic. Modified Hausdorff distance partially addresses this by averaging over the nearest-neighbor distribution rather than taking the maximum, but remains strongly correlated with RMS ($r = 0.92$) and does not offer meaningfully independent information. Given this redundancy, RMS error is recommended as the primary distance-based metric, as it is the most widely used, interpretable, and representative measure of overall geometric correspondence.

The Dice similarity coefficient, by contrast, proved the least suitable metric for this application. Despite the optimized trial showing strong geometric correspondence by all distance-based measures, the Dice score reached only 56%, which is a value that would typically suggest poor overlap. This behavior stems from the sensitivity of Dice to segmentation mask width rather than centerline position: small lateral offsets between thin, curvilinear structures produce disproportionately large reductions in pixel overlap, making the metric poorly suited for catheter shape comparison. The correlation analysis across 225 trials confirmed this, showing only a weak negative correlation between Dice and RMS error ($r = -0.32$), while all distance-based metrics were strongly intercorrelated ($r > 0.91$).

Taken together, the results indicate that no single metric is sufficient for comprehensive evaluation of endovascular simulation accuracy. A combination of RMS error for spatial localization and curvature analysis for mechanical shape characterization provides the most complete picture of simulator fidelity, and is recommended as the basis for future evaluation workflows.

6.6 Societal, ethical and ecological aspects

The societal impact of this project lies in its potential to improve patient safety, as more accurate simulations can reduce the risk of errors during real procedures by enhancing clinician training prior to patient contact. The use of objective accuracy metrics can also contribute to the standardization of simulator quality across institutions, reducing the reliance on subjective expert evaluation. However, there is a potential risk that clinicians may place excessive trust in simulators if accuracy metrics are misinterpreted or overstated, highlighting the importance of clearly communicating the scope and limitations of the proposed framework.

From an ethical perspective, the project uses exclusively simulator data, without involving patient data. This approach avoids privacy concerns and minimizes ethical approval requirements. Furthermore, the focus on objective, data-driven evaluation helps reduce subjective bias in simulator assessment, supporting fair and transparent evaluation practices.

With respect to ecological considerations, this project is not expected to have a significant environmental impact, as it is limited to offline image processing of existing simulation and experimental data. No additional hardware or energy-intensive processes are required beyond standard computational resources.

6.7 Future work

For future work, several directions can be explored to further improve and extend the proposed framework. Evaluating a broader range of catheter and guidewire types, including different diameters, stiffness profiles, and tip configurations, would enable a more comprehensive assessment of the robustness and generalizability of the system across different clinical scenarios. Expanding the experimental setup in this way would also improve statistical reliability of the results and allow stronger conclusions to be drawn regarding the applicability of the proposed methods.

Another important extension would be to incorporate three-dimensional analysis. The current framework operates entirely in two-dimensional image space, which reflects the projection-based nature of fluoroscopic imaging but inherently loses spatial depth information. Extending the framework to three-dimensional data would allow a more realistic representation of anatomical structures and instrument behavior, enabling more advanced geometric comparison and potentially improving the accuracy of alignment and shape analysis.

In addition, incorporating temporal information through the analysis of video sequences represents a promising direction for future development. The current evaluation is based on static image comparisons, whereas catheter navigation is inherently dynamic. By analyzing continuous image sequences, the framework could be extended to evaluate motion consistency, temporal stability, and dynamic deformation behavior during instrument manipulation. This could also enable continuous tracking of catheter motion and provide more detailed insight into transient mechanical effects and interaction forces.

Another limitation of the current study is the relatively simple physical phantom model used during validation. The phantom contains a limited vascular structure and therefore restricts the range of catheter configurations and navigation scenarios that can be evaluated. Employing a larger and more anatomically realistic phantom with more extensive vascular branching would enable testing under conditions that more closely resemble clinical procedures. This would also allow evaluation of a wider variety of catheter types, guidewire combinations, and procedural paths, thereby strengthening the clinical relevance of the framework.

Finally, future studies could investigate the integration of the proposed framework into automated simulator calibration workflows. The current optimization process demonstrates that simulator stiffness parameters can be tuned to reproduce physical catheter behavior. Building upon this, future systems could automatically calibrate simulator parameters against large physical reference datasets, enabling continuous validation and refinement of endovascular simulation models during simulator development.

7

Conclusion

This thesis presented an image-based framework for quantitative evaluation of endovascular simulation systems. The proposed framework was developed to address the lack of standardized and objective methods for assessing the technical accuracy and mechanical realism of simulated catheter behavior. By combining camera-based image acquisition, centerline extraction, geometric alignment, and quantitative comparison metrics, the framework enables systematic comparison between physical and simulated catheter configurations.

The validation experiments demonstrated that the proposed algorithms provide accurate and consistent performance for catheter analysis in two-dimensional image space. The catheter length estimation method achieved a mean absolute error of 0.117 cm and showed strong agreement with physical reference measurements across a wide range of catheter lengths. The alignment procedure further demonstrated robust compensation for geometric transformations between compared centerlines, enabling reliable comparison of physical and simulated catheter shapes.

The comparison between the physical simulator and the Mentice VIST simulator showed that the framework is capable of detecting and quantifying geometric differences between the systems. Although close correspondence could be achieved in several experiments, the results also highlighted important limitations related to differences in vessel compliance, frictional properties, and anatomical representation between the physical and virtual environments.

In addition, the proposed optimization framework demonstrated that simulator stiffness parameters can be tuned to reproduce physical catheter configurations with relatively high geometric similarity. The experiments further showed that Bayesian optimization provides an efficient alternative to manual parameter tuning when searching high-dimensional stiffness parameter spaces. However, the results also revealed limitations in the transferability of optimized parameter sets between different catheter-guidewire combinations, illustrating the complexity of catheter mechanics and the challenges associated with estimating physical properties from static image observations alone.

Among the evaluated comparison metrics, RMS error and curvature analysis were found to provide the most informative measures of geometric correspondence and mechanical behavior. RMS error was particularly useful for localizing spatial deviations between centerlines, while curvature analysis provided complementary information regarding local bending characteristics and catheter mechanics.

Overall, the results demonstrate that the proposed framework enables quantitative, repeatable, and objective evaluation of endovascular simulator realism. The framework provides a foundation for future research on automated simulator validation and calibration and may contribute to the development of more standardized evaluation methodologies for endovascular simulation systems.

Bibliography

[1] T. Peters and K. Cleary, Eds., *Image-Guided Interventions*, 1st ed. New York, NY, USA: Springer, 2008.

[2] J. Qiu, J. Ming, C. Qian, Y. Chen, D. Gao, S. Zhou, G. zhao, S. Tang, and L. Xing, “Guidewire simulation of endovascular intervention: A systematic review,” *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 18, no. 6, p. e2444, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rcs.2444

[3] H. Sharei, T. Alderliesten, J. J. van den Dobbelsteen, and J. Dankelman, “Navigation of guidewires and catheters in the body during intervention procedures: a review of computer-based models,” *Journal of Medical Imaging*, vol. 5, no. 1, p. 010902, 2018. [Online]. Available: https://doi.org/10.1117/1.JMI.5.1.010902

[4] D. L. Dawson, J. Meyer, E. S. Lee, and W. C. Pevec, “Training with simulation improves residents’ endovascular procedure skills,” *Journal of Vascular Surgery*, vol. 45, no. 1, pp. 149–154, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0741521406016272

[5] A. Amin, J. Salsamendi, and T. Sullivan, “High-fidelity endovascular simulation,” *Techniques in Vascular and Interventional Radiology*, vol. 22, no. 1, pp. 7–13, 2019, simulation in Interventional Radiology. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1089251618300751

[6] G. M. Fried, L. S. Feldman, M. C. Vassiliou, S. A. Fraser, D. Stanbridge, G. Ghitulescu, and C. G. Andrew, “Proving the value of simulation in laparoscopic surgery,” *Annals of Surgery*, vol. 240, no. 3, pp. 518–528, Sep. 2004.

[7] L. E. Davis, T. D. Storjohann, J. J. Spiegel, K. M. Beiber, and J. F. Barletta, “High-fidelity simulation for advanced cardiac life support training,” *American Journal of Pharmaceutical Education*, vol. 77, no. 3, p. 59, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0002945923029145

[8] N. J. Maran and R. J. Glavin, “Low-to high-fidelity simulation–a continuum of medical education?” *Medical education*, vol. 37, pp. 22–28, 2003.

[9] D. A. Gould and J. A. Reekers, “The role of simulation in training endovascular interventions,” *Cardiovascular and Interventional Radiology*, vol. 29, no. 1, pp. 3–9, 2006.

[10] K. See, K. Chui, W. Chan, K. Wong, and Y. Chan, “Evidence for endovascular simulation training: A systematic review,” *European Journal of Vascular and Endovascular Surgery*, vol. 51, no. 3, pp. 441–451, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1078588415007601

[11] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[12] W. Tang, T. R. Wan, D. A. Gould, T. How, and N. W. John, “A stable and real-time nonlinear elastic approach to simulating guidewire and catheter insertions based on cosserat rod,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 8, pp. 2211–2218, 2012.

[13] J. Spillmann and M. Teschner, “Corde: Cosserat rod elements for the dynamic simulation of one-dimensional elastic objects,” in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA).* Aire-la-Ville, Switzerland: Eurographics Association, 2007, pp. 63–72. [Online]. Available: https://cg.informatik.uni-freiburg.de/publications/2007_SCA_ropes.pdf

[14] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Cham, Switzerland: Springer, 2022.

[15] K. Sadekar. (2020, Oct.) Understanding lens distortion (image). Accessed: 2026-05-03. [Online]. Available: https://learnopencv.com/understanding-lens-distortion/

[16] D. C. Brown, “Decentering distortion of lenses,” *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, May 1966.

[17] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[18] OpenCV Team, “OpenCV: Open source computer vision library,” 2015, version 4.x. [Online]. Available: https://opencv.org

[19] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[20] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[21] A. Longo, S. Morscher, J. M. Najafababdi, D. Jüstel, C. Zakian, and V. Ntziachristos, “Assessment of hessian-based frangi vesselness filter in optoacoustic imaging,” *Photoacoustics*, vol. 20, p. 100200, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213597920300409

[22] T. Y. Zhang and C. Y. Suen, “A fast parallel algorithm for thinning digital patterns,” *Communications of the ACM*, vol. 27, no. 3, pp. 236–239, Mar. 1984.

[23] L. Piegl and W. Tiller, *The NURBS Book*, 2nd ed. Springer Science & Business Media, 1997.

[24] S. Umeyama, “Least-squares estimation of transformation parameters between two point patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.

[25] V. Luboz, J. Zhai, P. Littler, T. Odetoyinbo, D. Gould, T. How, and F. Bello, “Endovascular guidewire flexibility simulation,” in *Biomedical Simulation*, F. Bello and S. Cotin, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 171–180.

[26] Raspberry Pi Ltd, “Raspberry pi documentation: Camera modules,” 2026, accessed: 2026-03-30. [Online]. Available: https://www.raspberrypi.com/documentation/accessories/camera.html#about-the-camera-modules

[27] R. P. Ltd, “Raspberry pi documentation: Introduction,” 2026, accessed: 2026-03-30. [Online]. Available: https://www.raspberrypi.com/documentation/computers/raspberry-pi.html#introduction

[28] C. de Boor, *A Practical Guide to Splines*, revised ed., ser. Springer tates in Applied Mathematics. Springer-Verlag New York, 2001.

[29] A. Blake and M. Isard, *Active Contours.* Springer Science & Business Media, 1998.

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY