

Aus dem Deutschen Krebsforschungszentrum  
(Vorstand: Prof. Dr. Michael Baumann, Ursula Weyrich)  
Abteilung für Medizinische Bildverarbeitung  
(Leiter: Prof. Dr. rer. nat. Klaus H. Maier-Hein)

# ROBUST DEEP LEARNING FOR COMPUTER-ASSISTED SPINAL SURGERY

Inauguraldissertation  
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)  
an der  
Medizinischen Fakultät Heidelberg  
der  
Ruprecht-Karls-Universität

vorgelegt von  
Lisa Kausch  
aus  
Wolfenbüttel

2021



Dekan: Prof. Dr. med. Hans-Georg Kräusslich  
Doktorvater: Prof. Dr. rer. nat. Klaus H. Maier-Hein



# Contents

<b>Acronyms</b>	<b>i</b>
<b>List of Figures</b>	<b>iii</b>
<b>List of Tables</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Objectives and Contributions . . . . .	4
1.3 Outline . . . . .	7
<b>2 Background</b>	<b>9</b>
2.1 Imaging Background . . . . .	9
2.1.1 Fluoroscopy . . . . .	9
2.1.2 Computed tomography . . . . .	12
2.1.3 Conclusion . . . . .	14
2.2 Medical background . . . . .	15
2.2.1 Spinal anatomy . . . . .	15
2.2.2 Pedicle screw planning for CT-navigated spinal surgery . . . . .	21
2.2.3 C-arm positioning for fluoroscopy-guided spinal surgery . . . . .	24
2.2.4 Conclusion . . . . .	27
<b>3 State of the Art</b>	<b>29</b>
3.1 Pedicle screw planning for CT-navigated spinal surgery . . . . .	29

3.1.1	Pedicle screw planning . . . . .	29
3.1.2	Conclusion . . . . .	32
3.2	C-arm positioning for fluoroscopy-guided spinal surgery . . . . .	33
3.2.1	Data simulation . . . . .	33
3.2.2	Data analysis . . . . .	37
3.2.3	C-arm positioning . . . . .	42
3.2.4	Conclusion . . . . .	45
<b>4</b>	<b>Materials and Methods</b>	<b>47</b>
4.1	Pedicle screw planning for CT-navigated spinal surgery . . . . .	47
4.1.1	Retrospective planning data . . . . .	48
4.1.2	Baseline approach . . . . .	50
4.1.3	Proposed approach . . . . .	51
4.1.4	Prototype implementation . . . . .	53
4.2	C-arm positioning for fluoroscopy-guided spinal surgery . . . . .	54
4.2.1	Training data generation . . . . .	55
4.2.2	Acquisition of validation data . . . . .	64
4.2.3	Assessment of manual C-arm positioning accuracy and efficiency . . . . .	65
4.2.4	Baseline approach . . . . .	68
4.2.5	2-step intensity-based C-arm pose regression . . . . .	70
4.2.6	1-step sequential C-arm pose regression . . . . .	70
<b>5</b>	<b>Experiments and Results</b>	<b>73</b>
5.1	Pedicle screw planning for CT-navigated spinal surgery . . . . .	73
5.1.1	Manual pedicle screw planning accuracy . . . . .	74
5.1.2	Automatic pedicle screw planning accuracy . . . . .	75
5.1.3	Clinical acceptability . . . . .	82
5.1.4	Post-operative screw accuracy . . . . .	83
5.2	C-arm positioning for fluoroscopy-guided spinal surgery . . . . .	85
5.2.1	Training data accuracy . . . . .	86
5.2.2	Validation data accuracy . . . . .	87
5.2.3	Manual C-arm positioning accuracy and efficiency . . . . .	88
5.2.4	Simulated experiments . . . . .	95
5.2.5	Specimens experiments . . . . .	97
5.2.6	Systematic evaluation of influencing factors . . . . .	101
<b>6</b>	<b>Discussion</b>	<b>111</b>
6.1	Pedicle screw planning for CT-navigated spinal surgery . . . . .	111
6.1.1	Manual pedicle screw planning . . . . .	112

6.1.2	Automatic pedicle screw planning . . . . .	112
6.1.3	Conclusion . . . . .	116
6.1.4	Limitations and future work . . . . .	117
6.2	C-arm positioning for fluoroscopy-guided spinal surgery . . . . .	119
6.2.1	Manual C-arm positioning . . . . .	120
6.2.2	Automatic C-arm positioning . . . . .	122
6.2.3	Conclusion . . . . .	127
6.2.4	Limitations and future work . . . . .	128
<b>7</b>	<b>Summary</b>	<b>131</b>
<b>8</b>	<b>Zusammenfassung</b>	<b>133</b>
<b>A</b>	<b>Appendix</b>	<b>137</b>
	<b>Bibliography</b>	<b>137</b>
	<b>Own Contributions</b>	<b>159</b>
	Own share in data acquisition and data analysis . . . . .	159
	Own Publications . . . . .	161
	<b>Curriculum Vitae</b>	<b>162</b>
	<b>Acknowledgement</b>	<b>165</b>
	<b>Statutory Declaration</b>	<b>167</b>



## Acronyms

<b>AEC</b>	automatic exposure control
<b>Adam</b>	adaptive moment estimation algorithm
<b>a.-p.</b>	anterior-posterior
<b>BMD</b>	bone mineral density
<b>CAD</b>	computer-aided design
<b>CBCT</b>	cone-beam computed tomography
<b>cGAN</b>	conditional generative adversarial network
<b>CNN</b>	convolutional neural network
<b>CT</b>	computed tomography
<b>dof</b>	degrees-of-freedom
<b>DR</b>	domain randomization
<b>DRR</b>	digitally reconstructed radiograph
<b>FoV</b>	field-of-view
<b>FPD</b>	flat panel detectors
<b>GAN</b>	generative adversarial network
<b>GR</b>	Gertzbein-Robbins
<b>HU</b>	Hounsfield units
<b>L1</b>	first lumbar vertebra
<b>L2</b>	second lumbar vertebra

## Acronyms

---

<b>L3</b>	third lumbar vertebra
<b>L4</b>	fourth lumbar vertebra
<b>L5</b>	fifth lumbar vertebra
<b>MC</b>	Monte Carlo
<b>MIC</b>	Medical Image Computing
<b>MIS</b>	minimally invasive surgery
<b>MITK</b>	Medical Imaging Interaction Toolkit
<b>MRI</b>	magnetic resonance imaging
<b>NCC</b>	normalized cross-correlation
<b>OR</b>	operating room
<b>PCA</b>	principal component analysis
<b>ReLU</b>	rectified linear unit
<b>RQ</b>	research questions
<b>ROI</b>	region-of-interest
<b>S1</b>	first sacral vertebra
<b>Th1</b>	first thoracic vertebra
<b>VerSe</b>	large scale vertebrae segmentation challenge

## List of Figures

1.1	Clinical workflow. . . . .	2
2.1	Medical imaging devices. . . . .	10
2.2	C-arm out-of-plane rotations. . . . .	11
2.3	Medical terminology. . . . .	15
2.4	Anatomy of the spine. . . . .	16
2.5	Standard projections of the spine (L4). . . . .	19
2.6	CT-navigated spinal instrumentation workflow. . . . .	22
2.7	Workflow of percutaneous pedicle screw placement. . . . .	24
2.8	Variations in terminology for C-arm positioning. . . . .	26
3.1	Classification overview of state of the art methods for pedicle screw planning. . . . .	30
3.2	Data simulation: Synthetic X-rays. . . . .	34
3.3	<i>DeepDRR</i> simulation framework. . . . .	35
3.4	Data analysis: View-independent landmark detection and segmentation. . . . .	38
3.5	Classification overview of state of the art methods for C-arm positioning. . . . .	42
4.1	Vertebra level-specific screw planning dataset distribution. . . . .	48
4.2	Screw parameters and data preprocessing. . . . .	49
4.3	Instance-based screw planning based on direct CNN landmark regression. . . . .	50
4.4	Workflow for automatic pedicle screw placement. . . . .	52
4.5	<i>MITK Screw Planning Plugin</i> for pedicle screw planning. . . . .	53
4.6	<i>MITK Standard Plane Plugin</i> with DRR Preview. . . . .	56
4.7	DRR simulation framework. . . . .	57

## List of Figures

---

4.8	Preprocessing of X-ray images. . . . .	58
4.9	Generation of ground truth 3D vertebra segmentations and landmarks. . . . .	59
4.10	Generation of ground truth 2D annotations. . . . .	62
4.11	Preparation for specimens study. . . . .	64
4.12	Workflow for validation data acquisition in specimens study. . . . .	64
4.13	Accuracy measures for standard projections. . . . .	66
4.14	<i>Swipe App</i> : standard plane clinical accuracy web tool. . . . .	67
4.15	Baseline approach for automatic C-arm positioning. . . . .	68
4.16	2-step intensity-based C-arm pose regression. . . . .	69
4.17	1-step pose regression pipeline. . . . .	71
5.1	Manual pedicle screw planning accuracy. . . . .	74
5.2	Quantitative results of automatic pedicle screw planning. . . . .	76
5.3	Visual results of automatic pedicle screw planning. . . . .	77
5.4	Performance comparison: direct regression vs. segmentation. . . . .	78
5.5	Comparison of automatic vs. manual screw planning. . . . .	79
5.6	Vertebra-level specific Dice evaluation for manual/automatic screw planning. . . . .	81
5.7	Comparison of automatic vs. manual screw planning efficiency . . . . .	82
5.8	Gertzbein-Robbins classification for pedicle screw placement accuracy. . . . .	83
5.9	Example visualization of <i>navi</i> and <i>control</i> screws. . . . .	84
5.10	Comparison of post-operative screw placement to intraoperative planning. . . . .	85
5.11	Training data accuracy for standard projections. . . . .	86
5.12	Validation data accuracy after manual C-arm positioning. . . . .	88
5.13	Reference standard projections of 18 acquired validation specimens. . . . .	89
5.14	<i>Swipe App</i> : Measuring clinically acceptable deviations of standard projections. . . . .	90
5.15	Inter-/Intra-rater study to assess manual C-arm positioning accuracy. . . . .	93
5.16	Efficiency of manual C-arm positioning. . . . .	94
5.17	Quantitative evaluation of proposed pose regression approach (simulated data). . . . .	96
5.18	Qualitative evaluation of proposed pose regression approach (simulated data). . . . .	96
5.19	Quantitative evaluation of proposed pose regression approach (real X-rays without metal). . . . .	97
5.20	Exemplary standard DRRs complemented with proposed implant simulations. . . . .	98
5.21	Quantitative evaluation of proposed pose regression approach (real X-rays with metal). . . . .	99
5.22	Qualitative evaluation of proposed pose regression approach (real X-rays). . . . .	100

5.23	Overview of analyzed influencing factors. . . . .	101
5.24	Patient-specific evaluation of proposed approach on test DRRs. . . . .	103
5.25	Simulated reference standards. . . . .	103
5.26	View dependence of proposed approach on test DRRs. . . . .	104
5.27	Importance of landmark channels for pose regression performance. . . . .	104
5.28	Performance under ideal conditions (simulated data). . . . .	105
5.29	Performance under ideal conditions (real X-rays). . . . .	106
5.30	Variants of the pose regression pipeline. . . . .	106
5.31	Quantitative comparison of C-arm pose regression pipeline variants. . . . .	108
5.32	Ranking stability of proposed variants. . . . .	108
5.33	Performance of 1-step pose regression after training on validation specimens. . . . .	109
5.34	Importance of proposed implant simulation strategies . . . . .	110
A.1	Vertebra-level specific evaluation of manual C-arm positioning efficiency. . . . .	138
A.2	Quantitative evaluation of proposed pose regression approach (real X-rays). . . . .	139

*List of Figures*

---

## List of Tables

2.1	Quality criteria for spinal standard projections. . . . .	20
5.1	Statistical significance evaluation of automatic vs. manual screw planning. . .	80
5.2	<i>Swipe App</i> : Number of examples assessed for each standard projection. . .	91
5.3	<i>Swipe App</i> : Angular deviations ( $d\alpha, d\beta$ ) for each standard projection. . . .	91
5.4	Comparison of mean positioning accuracies across different studies. . . .	95
A.1	Specimen-specific quantitative evaluation of <i>inter-/intra-rater study</i> . . . . .	140

*List of Tables*

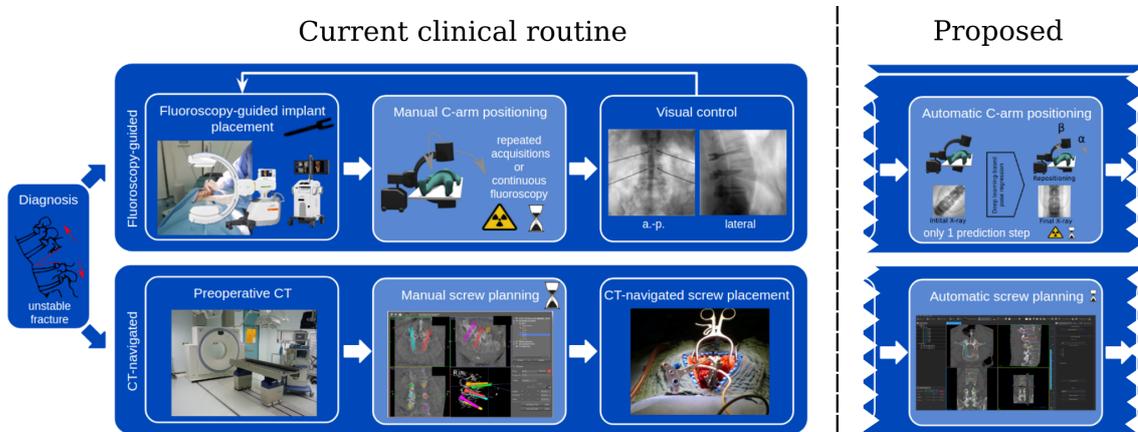
---

## 1.1 Motivation

Spinal fusion surgery is the standard technique to treat spinal instabilities. Spinal fixation typically involves bilateral pedicle screw insertion via a posterior approach, where the spine is assessed from dorsal. The screw heads serve as anchors for connecting plates or rods, which form a rigid stabilization of the spine.

Spinal instabilities can result from degenerative diseases, spinal tumors, or spinal fractures (Attar et al., 2001). A spinal fracture affects the vertebrae of the spine. It is typically caused by high-energy trauma such as motor vehicle accidents or falls from height, or low-energy osteoporotic trauma (Bhatia and Bowen, 2007). High-energy spinal fractures account for 3-6% of skeletal fractures and occur most often in humans between 20-50 years, with a four-fold higher increased risk for men (Walther, 2017). In contrast, females carry a higher risk for low-energy osteoporotic spinal fractures. Bone fragments resulting from a fracture or dislocation of a vertebra bear the risk of damaging the spinal nerves or the spinal cord. Consequently, all trauma patients with suspected spinal injury require clarification. The initial screening method in most hospitals is conventional radiography (i.e., the acquisition of anterior-posterior (a-p.) and lateral standard projections). The radiologic diagnosis can additionally include computed tomography (CT) or magnetic resonance imaging (MRI) examinations. The incidence of a secondary injury of the spine is 20%; thus, for every significant spine injury, all vertebra levels need to be examined clinically and radiologically (Rüedi et al., 2008).

Due to the spine's proximity to the spinal cord and the risk of neurological disorders or paralysis in case of co-injury, spinal surgeries are challenging and of high risk. Traumatic treatment of spine fractures commonly involves (1) fracture reduction, (2) decompression, (3) fixation by fusion, (4) reconstruction, (5) inspection and documentation of the result with X-ray imaging. Pedicle screw placement for spinal fusion is challenging due to the fracture itself, variations in pedicle morphology and spinal deformities. The major challenge is the accurate placement of the screws within the pedicle without cortical breach, which can otherwise result in nerve or spinal cord damage and fixation instability (Gelalis et al., 2012). Advances in medical imaging enable detailed visualization of the



**Figure 1.1: Clinical workflow:** Spinal fusion surgery can be guided by different image-assisted procedures: fluoroscopy guidance (top row), CT-navigated (bottom row). Both guidance techniques involve manual intervention steps that are highly expertise-dependent and time-consuming. The computer-assisted methods proposed in this thesis offer support for the surgeon and reduce manual interaction, expertise-dependence of results, radiation exposure, and procedural time. The methods can be seamlessly integrated into the clinical workflow.

internal anatomy and allow for less invasive surgical approaches and higher surgical accuracy. Imaging assists the surgeon in monitoring, guidance, and controlling the quality of interventions. Different image-assisted procedures have been introduced to facilitate pedicle screw insertion, e.g., fluoroscopy guidance or CT navigation (Patel et al., 2010).

An overview of both image-assisted procedures is given in Fig. 1.1.

Fluoroscopy is the standard procedure for guiding spinal interventions and evaluating the surgical result. It is an imaging technique that allows real-time 2D or 3D imaging of the interior of a patient. Fluoroscopy-guided spinal surgery involves the repeated acquisition of a-p. and lateral X-rays with an intraoperative mobile C-arm device. The a-p. and lateral X-rays are anatomical standard projections that allow for fracture reduction and implant positioning assessment in two orthogonal planes. To achieve a standard projection, a specific individual C-arm pose is required to adjust for the patient's positioning on the operating room (OR) table. C-arm positioning is usually performed manually using a trial-and-error approach involving repeated fluoroscopy (fluoro hunting), at the cost of radiation exposure and intervention time that largely depends on the experience of the operator (Blatter et al., 2004). It is challenging to obtain the desired view due to intraoperative positioning requirements, patient-related factors, and because the internal

anatomy is not visible from the outside.

Spinal surgery can also be performed under CT navigation, provided that a costly navigation system is accessible. The surgeon can assess the positions of surgical instruments in 3D images in real-time. CT navigation involves the acquisition of a pre- or intraoperative CT scan, time-consuming screw trajectory planning prior to surgery, and registration with the navigation system. The intervention is guided by computer navigation upon these pre-planned CT scans. The explicit 3D information has shown to increase screw placement accuracy, which is particularly beneficial in case of difficult anatomical conditions, e.g., in trauma and revision surgeries, when relevant anatomical landmarks are distorted or in deformities with altered pedicles (Gelalis et al., 2012; Mason et al., 2014; Kosmopoulos and Schizas, 2007; Rajasekaran and Shetty, 2021). However, the necessity of a CT volume comes at the cost of additional radiation for the patient. Surgeon's radiation was significantly reduced by navigation, requiring less or even no fluoroscopic acquisitions during implantation (Villard et al., 2014).

Spinal interventions constitute a multi-stage process in the OR that involves imaging as a key tool for guidance and quality assurance during the procedure. Both guidance techniques require manual intervention: For the CT navigation technique, this comprises the pedicle screw planning and, for fluoroscopy guidance, the positioning of the C-arm for standard projections. These processes require a high level of expertise and in-depth knowledge of anatomy and the anatomical orientation of structures. The anatomical understanding is additionally challenged by intraoperative specific requirements caused by varying patient positioning, patient-related factors (obesity, fractures, implants, limitations in joint movement, degenerative alterations of anatomy such as spondylosis), and varying surgical procedures which hinders the establishment of standardization. Many recently proposed state of the art approaches to replace or assist the manual steps make specific prior assumptions (Soltow and Rosenhahn, 2015; De Silva et al., 2017) or require a specific hardware setup (Bott et al., 2011; Gong et al., 2014; Bodart et al., 2018), that interferes with the clinical workflow. Others rely on manually defined geometric and structural modeling with explicitly defined constraints (Knez et al., 2016; Xiaozhao et al., 2016; Li et al., 2019). Methods that rely on modeling techniques that are targeted to represent the mean anatomical shape were proposed by, e.g., Shao et al. (2014); Goerres et al. (2017); Vijayan et al. (2019). Those approaches may be limited in application due to altered, deformed, degenerated, or fractured bones. Due to the mentioned methodological limitations, state-of-the-art assistance systems exhibit limited clinical applicability until today.

For automatic C-arm positioning, no commercial system is currently available. For pedicle screw planning, Brainlab (Riem, Germany) released a commercial software based on atlas registration. While leading to acceptable results in most cases, it suffers from the atlas-related limitations in case of complicated anatomical conditions (fractured, degenerated, or deformed vertebrae) since the atlas is targeted to represent the healthy anatomical shape.

Both techniques offer potential regarding the support of the surgeon. Despite their different setups, a recommended computer-assisted system could improve both techniques in terms of reducing manual interaction, expertise-dependence of results, radiation exposure, and procedural time. Specifically, this thesis proposes an image-based pedicle screw planning method for CT-navigated spinal surgery and an automatic C-arm positioning method for standard projections for fluoroscopy-guided spinal surgery.

The proposed methods are targeted to facilitate the clinical workflow integration without requiring additional technical equipment. In contrast to atlas- or model-based approaches, the anatomical variation is learned implicitly with deep learning techniques from retrospective CT datasets. Robust deep learning is enabled by developing and combining various methods that address aspects like data scarcity, intraoperative confounding factors, anatomical variation, including learning from simulations, advanced data augmentation, and integration of task-specific expert knowledge derived from surgical routine. The robust deep learning techniques could replace time-consuming and radiation-intensive manual processing steps in spinal interventions, thereby reducing the human intervention and radiation. Using automated deep learning tools could induce objective and standardized results learned from expert annotations and thus promote independence of personal training and experience.

## **1.2 Objectives and Contributions**

The main objective of this work was to develop and to evaluate methods for computer-assisted spinal surgery fulfilling the following requirements:

- **Clinical applicability:** Seamless integration of the proposed methods into the clinical workflow, without requiring external tracking hardware, additional equipment, or additional prior knowledge assumptions.
- **Performance requirements:** The developed methods should be as precise as possible and fulfill clinically relevant accuracy bounds.

- Benefits for the patient: The proposed methods should improve the quality of the procedure by reducing the radiation dose and intervention time.
- Benefits for the surgeon: The developed methods should assist or automate manually performed steps in spinal surgery techniques that currently require a high level of expertise. This ensures experience-level and surgeon-independent results. Further, the methods should reduce the radiation dose and intervention time.
- Representation of anatomical variability: The developed methods are based on deep learning techniques that are capable of capturing anatomical variations from retrospective computed tomography (CT) datasets. Robust learning is further ensured by utilizing simulations, advanced data augmentation, and integration of domain knowledge.
- Addressing intraoperative requirements: The proposed methods should tackle intraoperative requirements i.e., the presence of implants or surgical tools in the field-of-view (FoV) that may partly overlay the anatomy should be addressed.
- Generalization: The concepts should be transferable across vertebra levels or standard projections.

Considering the mentioned requirements, this thesis proposes computer-assisted methods to assist the surgeon during CT-navigated and fluoroscopy-guided spinal surgeries. The contributions of this work are summarized below, separated for the two spinal surgery techniques.

### 1. Pedicle screw planning for CT-navigated spinal surgery

**Method development:** To automate the time-consuming manual step of screw trajectory planning in preoperative CT scans, an image-based pedicle screw planning method is proposed and developed (Kausch et al., 2021a) (Sec. 4.1.3). The anatomical variation was learned implicitly with deep learning techniques utilizing a sizeable surgical planning dataset. The screw planning task was interpreted as a segmentation task, resulting in superior performance compared to direct convolutional neural network (CNN) landmark regression (Sec. 5.1.2). Screw locations were simulated in retrospective CT datasets, and potential screw locations were learned based on the context of the images. The proposed method is transferable between different vertebra levels. Automating the pedicle screw planning can significantly decrease the manual intervention, the expertise-dependence of results, and the preparation time needed for CT-navigated spinal surgery.

**Intraoperative clinical evaluation:** To evaluate the clinical applicability of the proposed approach for intraoperative planning, a software prototype for automatic pedicle screw planning was developed (Sec. 4.1.4). Quantitative evaluation of screw accuracy was performed by comparing automatically planned screws to corresponding manually planned screws. Additionally, intra- and inter-rater variance of manual screw planning was assessed, serving as a measure for clinically acceptable screw deviations (Sec. 5.1.2). Screw planning efficiency is compared between manual and automatic planning. For qualitative evaluation, all automatically planned screws were reviewed by experts in spine surgery and rated according to the Gertzbein-Robbins (GR) classification (Sec. 5.1.3).

**Postoperative clinical evaluation:** To assess the accuracy of the navigation procedure, post-operative screw trajectories are compared to the corresponding planned screw trajectories (Sec. 5.1.4).

## 2. C-arm positioning for fluoroscopy-guided spinal surgery (Sec. 4.2,5.2)

**Method development:** The problem of C-arm positioning is tackled without requiring patient-individual prior information like preoperative CT scans, without the need of registration and without additional technical burden, directly working on the 2D projection images just as the operator. The necessary anatomical hints for efficient C-arm positioning are learned from *in silico* simulations, leveraging 3D CTs. The generated dataset uniquely provides ground truth pose labels for arbitrary many training examples given by the different viewing directions defined during simulation. A baseline approach employing direct intensity-based CNN regression (Sec. 4.2.5) is proposed (Kausch et al., 2020) and compared to the proposed sequential approach, which leverages domain knowledge to guide robust decision-making (Kausch et al., 2021b) (Sec. 4.2.6). The developed methods were trained specifically for the fourth lumbar vertebra in a.-p. and lateral standard projections, but are generalizable to different anatomical regions and standard projections. The developed method promises to reduce radiation dose and time for patients and clinical staff and offer assistance especially beneficial for inexperienced surgeons.

**Domain requirements:** Intraoperative X-rays may contain metal implants like k-wires or screws. The domain gap between training from simulations and application on intraoperative X-rays was addressed by complementing simulated images with k-wire and screw simulations (Sec. 4.2.1). To further successfully achieve a transfer from synthetic to real X-rays and also to increase the interpretability of results, the pipeline was designed by mimicking clinical decision-making processes of spinal neurosurgeons.

**Validation:** The proposed methods were evaluated on synthetic X-rays not seen during training (Sec. 5.2.4). To evaluate the developed methods on real X-rays, a large specimens study with and without implants (i.e., k-wires and screws) was performed, simulating different stages of a real clinical procedure (Sec. 4.2.2). C-arm positioning accuracies were assessed separately for the selected standard projections and specimens (Sec. 5.2.5).

**Clinical context:** To derive a measure for the clinically relevant accuracy of C-arm positioning for specific standard projections, an inter-/intra-rater study was performed. Two experts positioned the C-arm manually to acquire the a.-p. and lateral standard projections across different vertebra levels for two specimens. The number of necessary X-rays for positioning, the time, and the dose was evaluated, as well as the variance in C-arm positioning (Sec. 5.2.3).

### 1.3 Outline

The thesis is structured as follows: In Chapter 2 the relevant information on the imaging and medical background is provided. Chapter 3 reviews first the state of the art on computer-assisted methods for pedicle screw planning for computed tomography (CT)-navigated spinal surgery. Thereafter, the state of the art techniques for C-arm positioning for fluoroscopy-guided spinal surgery are summarized, including strategies for data simulation (sec. 2.2.1), data analysis, i.e. view-independent segmentation and landmark detection (sec. 3.2.2), and computer-assisted methods for C-arm positioning (sec. 2.2.3).

All acquired datasets and developed methods in the scope of this thesis are presented in Chapter 4. The chapter is divided into two parts: Section 4.1 focuses on pedicle screw planning for CT-navigated spinal surgery. First, the retrospective planning data is introduced (Sec. 4.1.1). Subsequently, the baseline approach based on direct convolutional neural network (CNN) landmark regression (Sec. 4.1.2) and the proposed approach based on segmentation (Sec. 4.1.3) are presented. The developed software prototype for automatic pedicle screw planning used for clinical evaluation is described in Section 4.1.4. Section 4.2 focuses on C-arm positioning for fluoroscopy-guided spinal surgery. First, the training data generation (Sec. 4.2.1) and validation data acquisition (Sec. 4.2.2) is described. In Section 4.2.3 methods to assess the manual C-arm positioning accuracy and efficiency are introduced. Subsequently, the proposed intensity-based and sequential C-arm pose regression approaches are presented (Sec. 4.2.4-4.2.6).

Experiments and results are presented in Chapter 5 equally divided into two parts as before: Section 5.1 focuses on pedicle screw planning for CT-navigated spinal surgery. The

manual pedicle screw planning accuracy is evaluated in Section 5.2.3. Subsequently, the results for automatic pedicle screw planning are presented and related to the accuracies and efficiency of manual planning (Sec. 5.1.2). All automatically planned screws were reviewed by experts in spine surgery and results are summarized in Section 5.1.3. Section 5.1.4 evaluates the accuracy resulting from navigated screw placement in comparison to the preoperative planning. Section 5.2 focuses on C-arm positioning for fluoroscopy-guided spinal surgery. The training and validation data accuracy is evaluated in Section 5.2.1-5.2.2. A measure for the clinical relevant accuracy for specific standard projections derived from manual C-arm positioning is given in Section 5.2.3. Subsequently, qualitative and quantitative results for the proposed C-arm positioning methods are presented for synthetic and real X-rays in Section 5.2.4 and 5.2.5 respectively. Section 5.2.6 analyzes the influence and importance of design choices for the overall performance.

Chapter 6 discusses the developed methods based on the obtained results in the context of the current clinical practice and the state of the art and gives directions for future work.

The contributions and results are summarized in Chapter 7.

The following chapter describes the imaging and medical background. The goal is to give an insight into the imaging procedures, the anatomy of interest, and the different techniques for guidance during spinal surgeries. Section 2.1 introduces the imaging techniques utilized in this thesis. Section 2.2 provides the medical background on the spine anatomy and different image-based procedures for the assistance of spinal fusion surgeries.

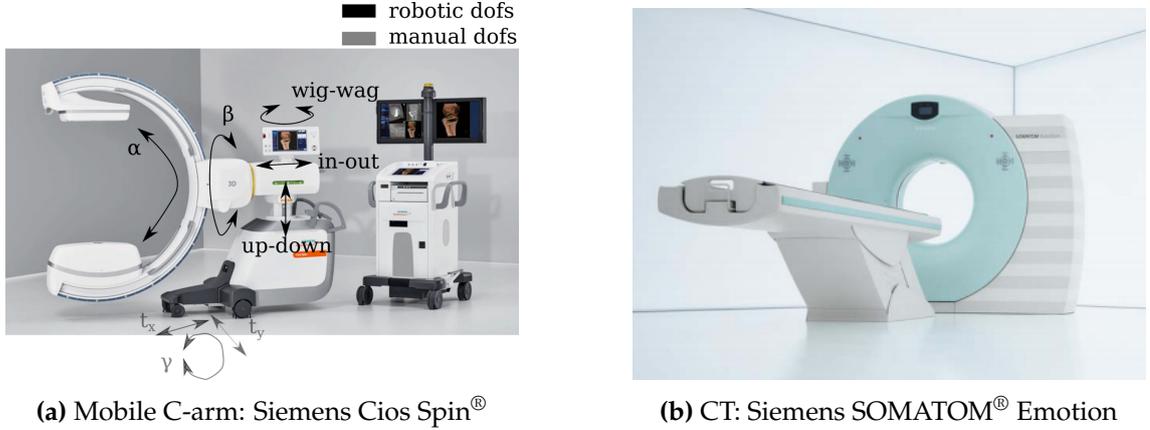
## **2.1 Imaging Background**

This section introduces the imaging techniques 2D fluoroscopy (sec. 2.1.1) and 3D computed tomography (CT) (sec. 2.1.2).

### **2.1.1 Fluoroscopy**

The discovery of X-rays by Wilhelm Conrad Röntgen in 1895 and the introduction of the first intraoperative C-arm in 1955 by Hugo Rost and Lothar Diethelm (Diethelm et al., 1956) allowed for intraoperative imaging through fluoroscopy. Fluoroscopy is a radiologic imaging technique that enables real-time visualization of the internal anatomy. It allows assessing fracture reduction and implant positioning intraoperatively. The increasing availability of intraoperative image guidance has triggered further developments of surgical techniques such as minimally invasive surgery (MIS) in trauma and orthopedic surgery like, e.g., percutaneous pedicle screw placement (Keil et al., 2018). MIS comes with the benefit of limiting the extent of muscles and healthy tissue disruption compared to open surgery. In contrast to conventional radiography, where static images are acquired pre- or post-operatively for diagnosis, fluoroscopy describes interventional radiography acquired intraoperatively by a C-arm (static or mobile) that delivers real-time images for monitoring, guidance, and quality control. While the first C-arm system was solely intended for surgical use, nowadays, the C-arm is used in different disciplines, e.g., orthopedics, trauma surgery, vascular surgery, and cardiology.

The mobile X-ray device, called C-arm, is a C-shaped imaging system with an X-ray source and a detector at its ends (Fig. 2.1a). The X-ray source emits X-rays that pass



**Figure 2.1: Medical imaging devices used in this thesis.** While the mobile C-arm 2.1a (adapted from Siemens AG (2019)) is used for intraoperative guidance and validation, the CT scanner 2.1b (Siemens AG, 2013) is employed preoperatively for diagnosis and planning.

through the body and hit the detector, where transmitted photons are converted into a visible image displayed on the C-arm monitor. X-ray photons are either absorbed, scattered, or transmitted without interaction while penetrating the patient (Berger et al., 2018). Assuming a mono-energetic X-ray beam with base intensity  $p_0$  and a homogeneous object with absorption coefficient  $\mu$ , the observed intensity  $p(\mathbf{u})$  at the detector pixel  $\mathbf{u}$ , according to Lambert-Beer's law, is related to the intersection length  $x$  of the object and the ray by

$$p(\mathbf{u}) = p_0 \cdot \exp(-\mu x). \quad (2.1)$$

However, in practical setups, emitted X-ray photons consist of poly-energetic energy spectra, and rays may pass through inhomogeneous materials. For an X-ray spectral density  $p_0(E)$  and an material- and energy-dependent linear attenuation coefficient  $\mu(\mathbf{x}, E)$  at each point  $\mathbf{x}$  on the ray path, the formula can be generalized to

$$p(\mathbf{u}) = \int p_0(E) \exp\left(-\int \mu(\mathbf{x}, E) d\mathbf{x}\right) dE. \quad (2.2)$$

The fluoroscopic image is a 2D projection of the attenuating properties of the different tissues penetrated by the X-rays (Aichinger et al., 2012). It is mainly used for fracture examination and implant positioning assessment. Since bones have a high absorption coefficient compared to the surrounding tissue, the bony structures are visualized with high contrast, making them easily distinguishable from the neighborhood. Very dense material such as metal is represented as white in the X-ray intensity domain. Soft tissue



**Figure 2.2: C-arm out-of-plane rotations**

can absorb a significant amount of radiation in the abdominal region and lead to contrast reduction. Especially in obese patients, this can impede the examination of bony structures (Keil et al., 2018). Since the attenuations are summed up along the ray, resulting X-rays are missing the depth information, and anatomical structures may overlap depending on the projection direction. For correct interpretation, the acquisition of anatomy-specific standard projections is essential. The most frequently acquired standard projections for the spine are the a.-p. and lateral standard projection (Rüedi et al., 2008). A key task of the surgeon is to position the C-arm to acquire the necessary standard projections for the operation.

A mobile C-arm has six degrees-of-freedom (dof), three translational, and three rotational parameters. The primary orbital rotation is denoted by  $\alpha$  and describes the rotation within the plane of the C-arm gantry (Fig. 2.2 (a)). The secondary angular rotation is denoted by  $\beta$  and describes the rotation perpendicular to the plane of the C-arm gantry (Fig. 2.2 (b)). The in-plane rotation in the detector plane is denoted by  $\gamma$ . The spatial point about which the C-arm is rotated is called isocenter. This setup enables the acquisition of X-rays from arbitrary projection angles, only limited by the patient's anatomy. A standard projection corresponds to a specific orientation of the C-arm that depends on the patient's position.

The C-arm X-ray source is a point source emitting a cone-beam with X-rays traveling in straight lines. Structures further away from the detector will be magnified in the detector plane. Magnification is reduced by moving the source further away from the object. At the same time, this reduces the radiation dose for the patient's skin according to the inverse square law (International Atomic Energy Agency, 2021). For accurate measuring, structures should be placed as close as possible to the detector.

Flat panel detectors (FPD) became the state of the art technology for X-ray detectors (Berger et al., 2018). Compared to the previous detector generation of image intensifiers,

the main advantage of FPD is the direct conversion of X-ray radiation into a digital signal, which makes data storage and patient handling more efficient and allows for separate photon detection and image processing. Additional benefits are the production of distortion-free images, lower patient dose, improved image quality, contrast, and spatial resolution (Berger et al., 2018; Aichinger et al., 2012; Keil et al., 2018). Further, FPD facilitates a more compact and lighter design, which benefits the rotational speed of the C-arm and simplifies the handling in an operating room (OR) with limited space (Berger et al., 2018). However, the C-arm field-of-view (FoV) is limited by the distance between both ends of the detector. On the one hand, larger radii facilitate the acquisition of bigger objects, but on the other hand, they may hamper the integration into the clinical workflow. FPDs pose higher acquisition costs than image intensifiers.

Intraoperative C-arm systems have a lower power supply than stationary radiography devices. This implies lower radiation doses, causing lower image contrast and resolution (Keil et al., 2018). Fluoroscopic devices contain an automatic exposure control (AEC) mode which regulates the exposure factors such that the X-ray radiation at the detector is above a certain threshold that allows for the generation of clinically useful X-rays.

Mobile C-arms are mandatory to immediately and continuously control fracture reduction and implant position intraoperatively in trauma surgery (Keil et al., 2018). X-ray imaging allows for non-invasive and painless visualization of the internal anatomy. However, the potential hazard of radiation needs to be weighed against the benefit of the examination.

### 2.1.2 Computed tomography

Computed tomography (CT) is the most widely available and used image modality in radiology and constitutes a standard procedure for fast 3D imaging in trauma surgery (Buzug, 2011). While 2D X-rays are summation images with inherent depth ambiguities, resulting in poor contrast and challenging interpretation, CTs are reconstructed from a set of X-rays non-invasively acquired from different directions. The reconstructed cross-sectional views do not suffer from superposition and exhibit higher image contrast with differentiable soft tissue (Buzug, 2011).

The mathematical theory for 3D image reconstruction dates back to Johann Radon, who introduced the filtered backprojection in 1917. Due to its computational complexity, the first commercial scanner for human imaging was only introduced in 1971, after the development of computer technology, by Godfrey Newbold Hounsfield and Allan McLeod

Cormack (Taubmann et al., 2018). The most common methods for image reconstruction are the analytic filtered backprojection and the algebraic reconstruction (Taubmann et al., 2018).

After reconstruction, the CT values represent physical attenuation values  $\mu$  that are commonly converted to Hounsfield units (HU), defined by

$$CT - value = \frac{\mu - \mu_{water}}{\mu_{water}} \cdot 1000 [HU].$$

On this scale, air is mapped to the CT value of -1000 HU and water to the CT value of 0 HU.

A conventional CT setup is static, consisting of an X-ray source and a detector hidden in the rotating gantry and a patient table, as depicted in Fig. 2.1b. Since the introduction of CT scanners, different scanner generations have emerged driven by the goals of lower acquisition times, lower radiation exposure, and decreased cost (Buzug, 2011). They differ in the construction of X-ray tubes (pencil beam, fan beam, cone beam) and detectors and their movement around the table (Buzug, 2011). The intervention of helical acquisition paths by Willi Kalender in 1990 allowed the imaging of large objects under continuous rotation. The helical path is achieved by moving the patient table through the gantry while the X-ray source rotates in the x-y plane. Standard reconstruction methods can be applied by interpolating circular acquisition trajectories from neighboring helical segments (Taubmann et al., 2018).

The image quality is influenced by several factors: the acquisition geometry (tube, detector, movement), the tube voltage (kV), the tube current (mA), the rotation time, the beam filtration, the focal spot size, speed of table movement (pitch), number of projections, detector element spacing, beam collimation, slice increment, and the reconstruction algorithm. Tube current and tube voltage define the quantity and energy of the X-ray quanta emitted from the source (Riemer, 2017).

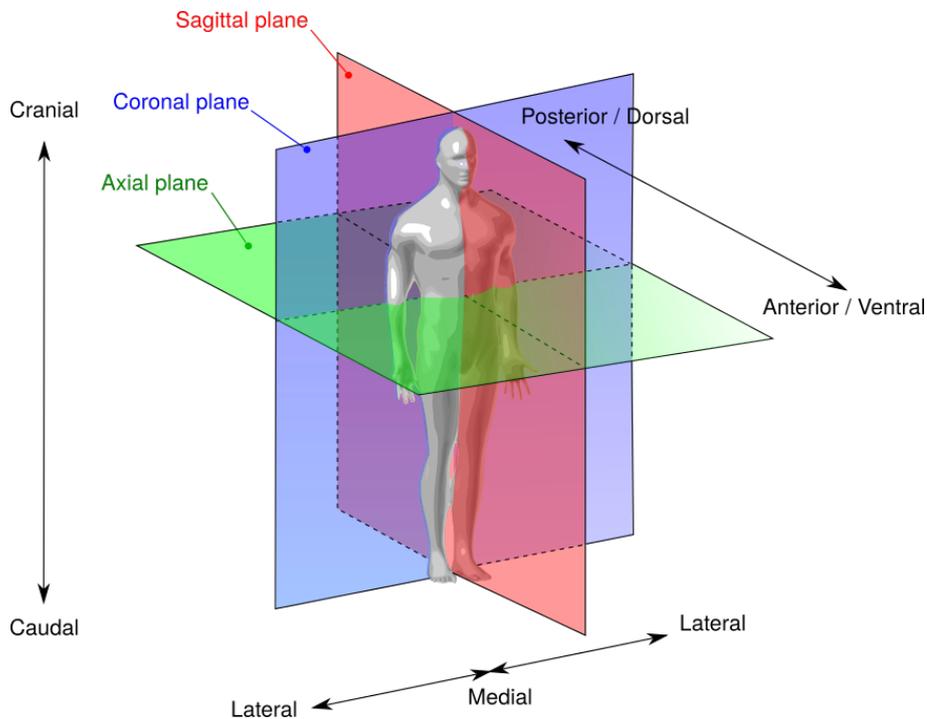
Image quality is additionally influenced by image artifacts that reduce the diagnostic value. The partial volume effect is caused by limited detector resolutions that result in a volume averaging across tissue boundaries projecting in the same detector element. It becomes particularly noticeable at sharply contrasted boundaries that appear blurred in the projection image. Beam hardening artifacts are caused by energy shifts in the poly-energetic X-ray spectrum along with the patient penetration. Energy shifts result from the property that low-energy X-ray beams are more strongly absorbed than high-energy beams (beam hardening). However, only the integral intensity over all wavelengths is

measured at the detector. Consequently, inconsistent individual projections are measured from different directions, which do not complement each other in the filtered backpropagation and result in streak artifacts. With filtered backpropagation, those 2D artifacts can spread across the entire image (Buzug, 2011). The energy dependence of the attenuation can be corrected by filtering the soft radiation before penetration. Motion artifacts resulting from breathing, patient motion, or contrast medium can lead to blurring and streak artifacts. It can be tackled by the acceleration of the data acquisition. Compton scattering can cause X-rays photons to change direction and energy, which is particularly noticeable if the photon is scattered into detector elements with few photons, e.g., behind metal objects. Anti-scatter grids in front of the detector can reduce this effect. Metal artifacts cause beam hardening, scatter, and photon starvation, leading to noisy projections.

With cone-beam computed tomography (CBCT) devices, the conventional static CT devices have been transferred to the OR setting. In contrast to typical CT scanners with line-shaped detectors, CBCT systems emit cone-beam shaped X-ray beams to an area detector. The image sequence for reconstruction is acquired in a single motorized rotation about the patient. CBCT uses low-power continuous fluoroscopy throughout the scan, is quicker in motion than traditional spiral motion, and involves less radiation. The central beam passes through the center of rotation, and patient movement is not required. Consequently, the acquisition volume is constrained by the detector FoV ( $30\text{ cm}^2$  flat detector,  $16\text{ cm}^3$  reconstructed 3D volume for the Siemens Cios Spin<sup>®</sup>, Fig. 2.1a) and the C-arm geometry ( $116.4\text{ cm}$  source-to-detector distance for the Cios Spin).

### 2.1.3 Conclusion

To sum up, conventional CT is a static 3D imaging modality that is widely used for preoperative diagnosis and planning and post-operative control. Intraoperative CTs also exist but are still rarely available due to their expensive acquisition costs. Mobile systems allow flexible positioning of the device relative to the patient and are suitable for intraoperative use. The flexibility comes at a cost of an unknown system to device orientation. Fluoroscopy delivers intraoperative X-rays in real-time that allow assessing reduction and implant placement. However, interpretation of the summation images requires expertise to compensate for the missing depth information. In contrast, 3D cross-sectional CT imaging explicitly delivers this information at the cost of increased radiation dose and computation time. Both imaging techniques, 2D fluoroscopy and 3D CT, involve ionizing radiation; thus, the potential risk needs to be weighted against the benefit of the examination. Particularly, the iterative radiation during manual C-arm positioning for standard projections affects the patient as well as the clinical staff, while only the final



**Figure 2.3: Medical terminology:** Anatomical viewing planes and directional relationships extended from Mrabet (2012).

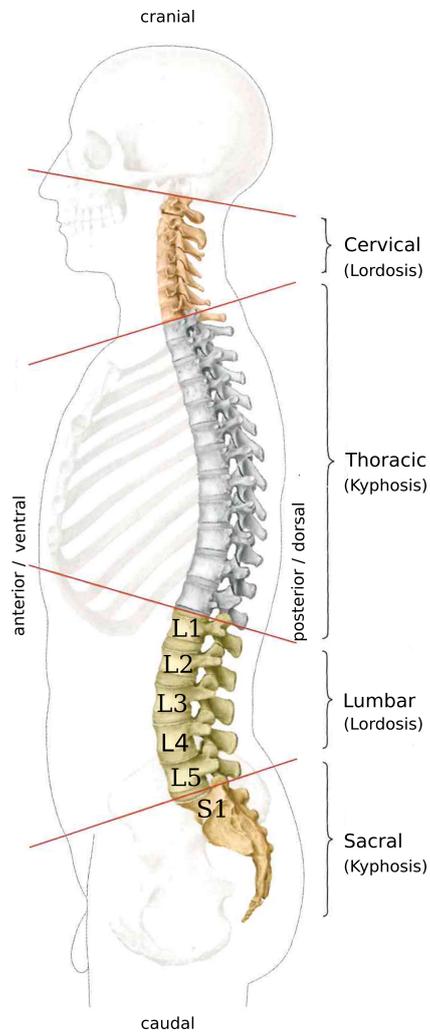
standard projection is employed for verification and guidance purposes. Therefore, one aim of this thesis comprises the development of computer-assisted methods for automatic C-arm positioning.

## 2.2 Medical background

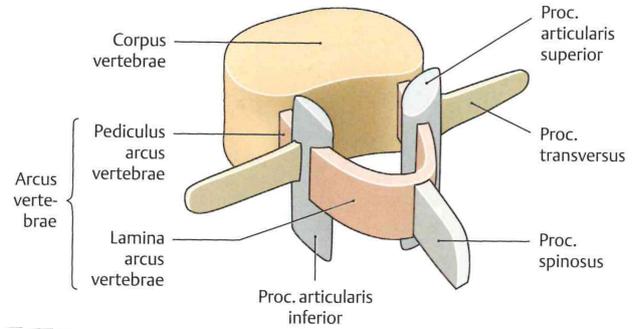
This section provides information about the spinal anatomy, spinal fractures, and treatment (sec. 2.2.1). Further, CT-navigated (sec. 2.2.2) and fluoroscopy-guided (sec. 2.2.3) spinal surgery are introduced. Fig. 2.3 introduces the relevant medical terminology for the following section.

### 2.2.1 Spinal anatomy

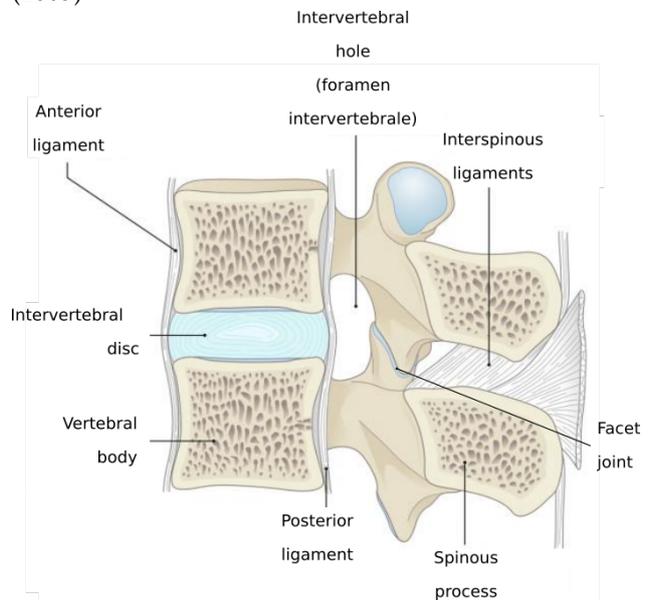
The spine is the central weight-bearing element of the human skeleton. It supports and stabilizes the body, enables the upright posture, and facilitates movements in varying directions. Additionally, it directly or indirectly connects all skeletal bones. At the top, the spine is connected through the head joint with the skull, and at the bottom, through the



(a) Spine anatomy with labeled vertebrae for the lumbosacral segment adapted from Schünke et al. (2005)



(b) Vertebra structure adapted from Schünke et al. (2005)



(c) Vertebra arrangement adapted from Bühren and Josten (2012)

Figure 2.4: Anatomy of the spine.

Os sacrum-Os ilium joint with the pelvis. The spinal canal, formed by the vertebra holes of the stacked vertebrae, surrounds the spinal cord and protects it against injuries.

**Spinal structure:** Typically, the spine consists of 24 free vertebrae, which are moveable, connected via intervertebral discs and various ligaments. From cranial to caudal, the spine can be divided into four segments (Fig. 2.4a): the cervical spine (C1-C7), the thoracic spine (Th1-Th12), the lumbar spine (L1-L5), and the sacral spine (Os sacrum, Os coccyges). Occasionally, vertebrae in the transition areas can be shaped atypically. Most frequently, this is observed for the transition between the lumbar and sacral segments. For example, if the first sacral vertebra is not combined with the Os sacrum, resulting in one additional lumbar vertebra, this is called lumbalisation. In contrast, if the fifth lumbar vertebra is fused with the Os sacrum, reducing the number of lumbar vertebrae by one, this is called sacralisation (Schünke et al., 2005). From the sagittal view, the spine resembles a double S shape (Fig. 2.4a): the cervical and lumbar spine segment have the curvature ventral, while the thoracic and sacral spine segment bend dorsal. While the kyphosis in the thoracic segment allows for a good lung extension, the cervical and lumbar lordosis are critical to ensure an even distribution of bodyweight (Schünke et al., 2005). In addition, the double S-shaped curvature limits the disruptions generated by walking to a minimum in the head region containing the brain. In coronal view, the spine is approximately line-shaped. Lateral curvature indicates scoliosis. The transition areas between the different segments are predilection sites for spinal injuries, like fractures or ligament injuries (Schünke et al., 2005).

**Vertebra structure:** Vertebra shape and size differ across different spinal regions (Schünke et al., 2005). Since the spine needs to carry most body weight, the vertebra size increases from cranial to caudal. Additionally, the shaping of the vertebral arch and the processes vary (Schünke et al., 2005). However, the vertebra structure is similar across all vertebra levels (1st and 2nd cervical vertebra excluded) as depicted in the model view in Fig. 2.4b. It consists of: a vertebral body (corpus vertebrae), a vertebral arch (arcus vertebrae), a processus spinosus, two transverse processes, and four articular processes. The processes serve as anchor points for muscles and ligaments and specifically for the thoracic segment as rib-vertebra-joints. The spinal canal is built by the union of all vertebra holes. An intervertebral hole is formed bilaterally between the vertebral arches of two neighboring vertebrae. It is called foramen intervertebrale and is the exit point for spinal nerves originating from the spinal column.

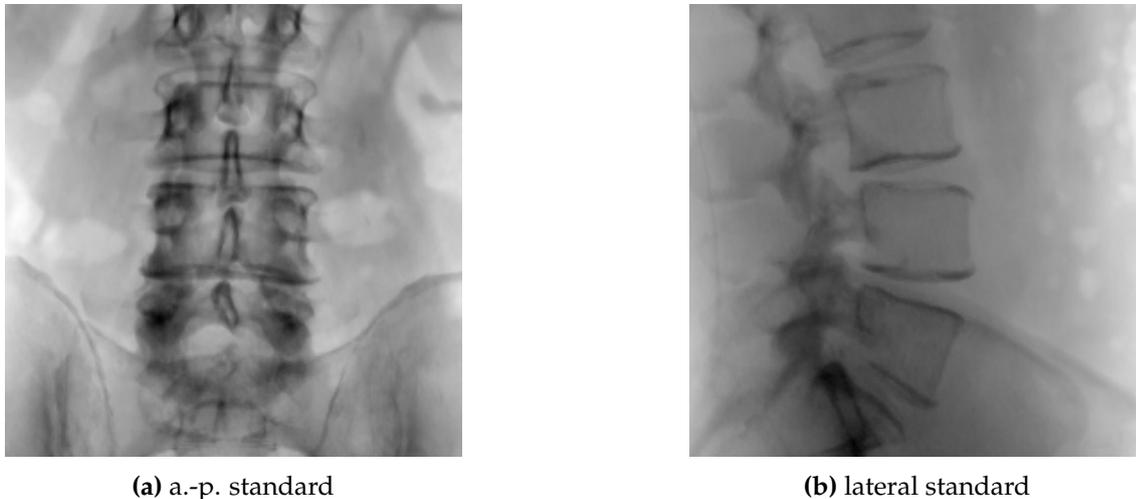
**Incidence of spinal fractures:** Approximately 50% of spinal fractures affect the thoracolumbar transition area, 40% are located in the thoracic spine, and 10% in the lumbar

spine (Tscherne and Blauth, 1998). Several anatomical factors cause the increased fracture incidence of the thoracolumbar transition area: the thoracolumbar spine is (1) a transition between kyphosis (Th11-L2) and lordosis (L3-L5), (2) a transition between the relatively rigid thoracic spine and the mobile lumbar spine (Payer et al., 2020). Spinal fractures account for 3-6% of skeletal fractures and have a four-fold increased incidence for males, usually between 20-50 years (Walthers, 2017). High energy traumas are typically caused by traffic accidents or fall from height (Tscherne and Blauth, 1998). Low energy trauma fractures most often affect older women with osteoporosis (Walthers, 2017). If a spinal injury is identified, the entire spine requires examination because the incidence of a secondary vertebra injury is 20% (Rüedi et al., 2008).

**Fracture types:** Most commonly observed fracture patterns in the thoracolumbar segment include compression fractures, burst fractures, flexion-distraction injuries, and fracture-dislocations (Vaccaro et al., 2003). Though numerous classification systems for thoracolumbar fractures have been presented, spinal injuries' great variability and complexity still pose a major challenge in spinal fracture assessment. Consensus exists, however, that the assessment of spinal stability is essential to determine adequate treatment of spinal fractures. Thereby, stability refers to neurologic as well as mechanical stability (Patel et al., 2010). Unstable spinal fractures typically require surgical intervention, while stable fractures without neurologic symptoms can be managed conservatively.

**Standard projections:** The acquisition of a.-p. and lateral X-rays form the basis of each radiological examination of the spine (Rüedi et al., 2008). Further, the acquisition of spinal standard projections is essential for interventional guidance during fluoroscopy-guided spinal surgery. The standard projections are defined to minimize projection artifacts, thereby enabling a reliable assessment of the relevant bony structures. Depending on clinical indications, additional standard projections can be required. The spinal standard projections are vertebra-level specific. Fluoroscopic projections are summation images, and incorrect projection poses can result in overlays and ambiguities caused by the projective simplification. The correct acquisition of anatomy-specific standard projections is essential to evaluate the surgical results including fracture reduction and implant positioning and to avoid overlooked errors (Rikli et al., 2018). The most frequently acquired spinal standard projections are shown in Fig. 2.5. While the a.-p. X-ray allows for the assessment of coronal and rotational alignment, the lateral X-ray enables the evaluation of the sagittal alignment and vertebra body height (Patel et al., 2010).

In the following, criteria for the intraoperative acquisition of a.-p. and lateral standard projections are summarized. During the intervention via the posterior approach, the patient lies in prone position. For the a.-p. standard projection, the X-rays travel from



**Figure 2.5: Standard projections of the spine (fourth lumbar vertebra (L4)).**

anterior to posterior. For the lateral projection, X-rays travel from side to side. To reduce radiation dose during a.-p. acquisitions, the detector should be positioned above the table, and during lateral acquisitions, the clinical staff should be located on the side of the detector (Keil et al., 2018). According to the quality criteria for X-ray diagnostic examinations released by the German Medical Organization, spinal standard projections should fulfill the image criteria given in Tab. 2.1 (Bundesärztekammer, 2007; Harris et al., 2008). For each standard projection, the vertebra of interest should be centered in the central beam area.

**Treatment:** Surgical treatment aims at maximizing neurological recovery, restoring spinal stability, facilitating fracture healing, enabling early mobilization and rehabilitation (Patel et al., 2010). Spinal fracture treatment involves (1) reposition, (2) decompression, (3) stabilization, and (4) reconstruction.

**(1) Reposition:** Fracture reduction concerns the alignment of bone structures back to their correct position such that the spinal shape is reconstructed. Exact repositioning is a requirement for effective stabilization and fracture healing (Bühren and Josten, 2012).

**(2) Decompression:** Sometimes, repositioning of bone and ligament structures can already reduce the stress on the spinal cord and nerves. If required, intervertebral disc or bone fragments are removed from the spinal canal. Further, consideration of direct surgical decompression techniques may be required, e.g., discectomy, laminectomy, corpectomy. Discectomy includes the removal of intervertebral disc material

a.-p. standard	lateral standard
<ul style="list-style-type: none"> <li>• contours of vertebrae should appear sharp without doubled end-plate shadow</li> <li>• gaps between vertebrae should be clearly visible, vertebrae should project symmetrically</li> <li>• spinous processes should be centered between the pedicles</li> </ul>	<ul style="list-style-type: none"> <li>• base and cover plates of vertebrae should appear line-shaped in the central beam area</li> <li>• gaps between vertebrae should be clearly visible</li> <li>• the lumbar foramina intervertebral are maximally extended (Fig. 2.4c)</li> <li>• pedicles should be superimposed</li> </ul>

**Table 2.1: Quality criteria for spinal standard projections.**

to decompress affected nerves. In laminectomy, the lamina (Fig. 2.4b) is removed to widen the spinal canal for decompression. Corpectomy refers to removing parts of the vertebra bodies to release stress on the spinal cord and nerves in case of spinal canal narrowing due to degenerate vertebra changes like bone spurs (overgrowth). All named techniques can cause spinal destabilization (Bühren and Josten, 2012).

**(3) Stabilization:** The gold standard treatment to maintain the stability of unstable spinal fractures is spinal fusion (Mobbs et al., 2011). Further, spinal fusion surgery is used to treat spinal deformity such as scoliosis or hyperkyphosis, degenerative conditions (spondylolisthesis, disc degeneration, spinal stenosis), or spinal instabilities resulting from tumors (Attar et al., 2001). Regarding thoracolumbar fixation, pedicle screw instrumentation via a posterior approach, i.e., where the spine is assessed from dorsal, is the most common technique (Patel et al., 2010). Spinal instrumentation involves the bilateral placement of screws through the pedicle, which is the strongest point of the vertebra (Molinari, 2016). The screw heads serve as anchor points for connecting plates and rods that rigidly stabilize the spine, restore the spinal alignment and enable successful fusion. Insertion of screw instrumentation into the pedicle area requires extensive anatomical knowledge. It is of high risk due to its proximity to the spinal cord or exiting nerve roots and the potential complications, e.g., neurological disorders or paralysis, in case of co-injury caused by inaccurate screw placement within the pedicle (Gelalis et al., 2012).

**(4) Reconstruction:** Destroyed bony structures can require anatomical reconstruction. Reconstruction is most commonly performed in the anterior spine column by inserting

artificial vertebra body implants (cages) or bone grafting.

**(5) Inspection and documentation:** Imaging is employed to validate individual procedural steps and to document the correct instrumentation.

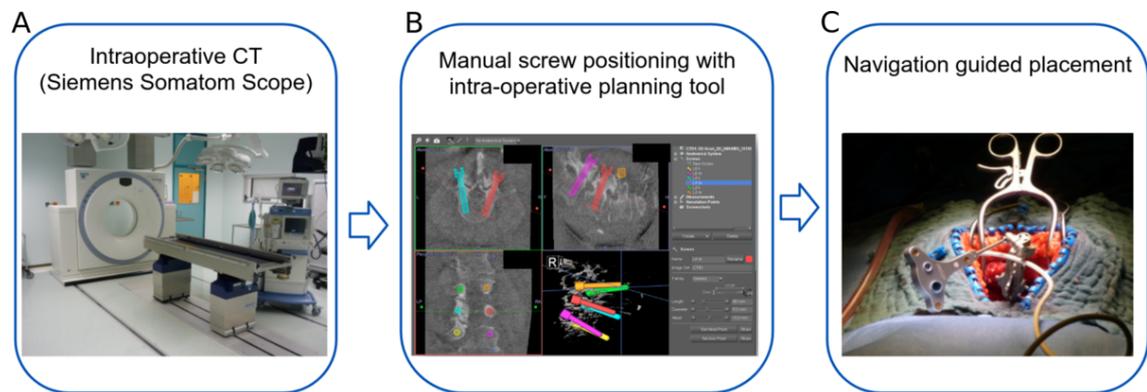
Spinal instrumentation can be performed with a standard midline incision posterior approach or minimally invasive. Percutaneous pedicle screw insertion entails benefits for the patient like a limited disruption of muscles and healthy tissue, diminished postoperative pain, less blood loss, and improved cosmetic result (Rodts, 2003). In contrast to open techniques, where important landmarks for optimal screw placement are visible, minimally invasive approaches rely much stronger on intraoperative image guidance. Spinal fractures that require a fracture reduction or direct decompression of nerve structures can often not be treated with minimally invasive surgery (MIS). Also, the open technique is recommended if the distance from the skin to the screw insertion point is larger than 9 cm, as can be observed in obese patients (Peloza, 2016).

Advances in imaging technologies and computer-assisted methods offer different guidance techniques to facilitate pedicle screw insertion, e.g., fluoroscopy or CT navigation. Fluoroscopy is the standard procedure for guiding spinal interventions due to its widespread availability. However, anatomical distortions or patient's obesity can complicate the interpretability and reduce the image quality of the 2D fluoroscopic summation images and thereby compromise fluoroscopic guidance. In complex anatomical scenarios, CT navigation, which offers explicit 3D information, can foster the accurate placement of pedicle screws. Both techniques are introduced in the following.

### 2.2.2 Pedicle screw planning for CT-navigated spinal surgery

In CT-navigated spinal surgery, the intervention is guided by computer navigation upon pre- or intraoperative planned screw trajectories on a computed tomography (CT) scan. Additionally, CT-guidance enables the visualization of instrument and implant positioning in 3D.

The surgical routine in Heidelberg for CT-navigated spinal instrumentation since 2008 is depicted in Fig. 2.6. The process consists of three steps: an intraoperative CT containing the vertebrae of interest is acquired in prone position with an intraoperative CT scanner or a cone-beam computed tomography (CBCT) device (A), screw trajectories and dimensions are planned in the acquired CT using an intraoperative planning tool (SpineMap 3D software, Stryker) (B), the plans are used for navigation guidance (C). To transfer the knowledge from CT to the patient, a 3D-3D registration is required. Therefore,



**Figure 2.6: CT-navigated spinal instrumentation workflow.** Surgical routine in Heidelberg for CT-navigated spinal instrumentation since 2008 consisting of the three steps: intraoperative CT acquisition (A), manual screw trajectory planning (SpineMap 3D software, Stryker) (B), navigation-guided screw placement (C). Image courtesy of Department of Neurosurgery, Heidelberg University Hospital.

small screw markers are placed on the vertebrae before the CT is acquired. By pointing to these markers with traceable instruments, the point-to-point registration is automatically initialized (Ishak et al., 2019).

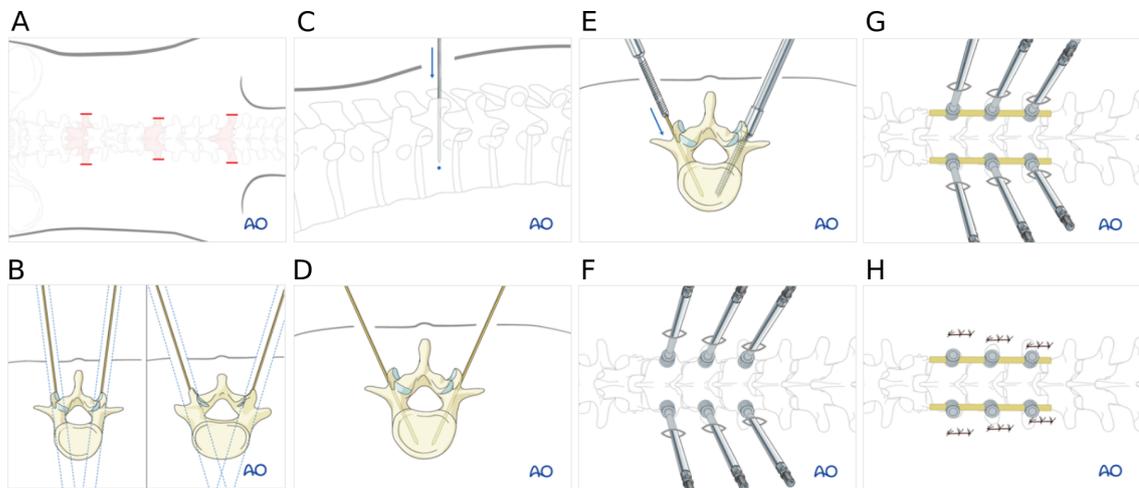
A systematic literature review by Gelalis et al. (2012) revealed that only 28 – 85% of screws were correctly placed in the pedicle with traditional fluoroscopic guidance. This has triggered the increasing use of navigation systems in spinal instrumentation, which has been shown to lead to a significant reduction of misplaced screws (89-100% accuracy with CT-navigation, 81-92% accuracy with fluoroscopy-based navigation) (Gelalis et al., 2012). CT-navigation can support implant placement, particularly in anatomically critical areas, when pedicles are small, disparate, malformed, or variable angulated, and when anatomical landmarks are distorted in trauma or revision surgeries (Rajasekaran and Shetty, 2021). Mason et al. (2014) reported an increased screw accuracy of 95.5% across all spine segments with CT-navigation, compared to 84.3% with 2D fluoroscopic navigation and 68.1% with conventional fluoroscopy. Other studies, however, reported no difference between robot-assisted and conventional fluoroscopy-guided pedicle screw placement (Schatlo et al., 2014).

Navigation guidance has shown to reduce the operative time, especially in complicated scenarios (Sasso and Garrido, 2007). Further, it enables minimally invasive screw placement because the visibility of anatomical structures is not required. That implies less blood loss, smaller wounds, less pain, and faster patient recovery.

CT navigation facilitates precise instrument tracking and consequently requires fewer verification scans. Thus, the intraoperative X-ray radiation is reduced. Bratschitsch et al. (2019) compared image-assisted guidance techniques in terms of radiation exposure. They showed that patient's radiation dose is approximately increased by 3-fold during navigation compared to fluoroscopy. The increased patient's radiation results from the necessity of a CT scan. In contrast, Villard et al. (2014) reported increased patient's radiation, however, without statistical significance. The variable patients' radiations reported by different studies can be attributed to different underlying patient groups, where the strengths of navigation are most notable in cases of trauma, revision surgeries, and spinal deformities (Rajasekaran and Shetty, 2021). Nonetheless, studies agree that spinal surgeons' radiation dose was decreased up to 10-fold by the use of navigation (Bratschitsch et al., 2019; Villard et al., 2014), which is of high relevance since spinal neurosurgeons exhibit the highest amount of radiation dose among all orthopedic surgeons (Matityahu et al., 2017).

The integration of navigation systems into the clinical workflow requires surgical training. If the workflow is not smooth, the intervention time is extended, which reduces the acceptance among surgeons. The navigation system purchase is expensive, while the financial returns for navigated and image-guided interventions are the same. However, costs for surgical correction of potentially misplaced screws have to be balanced against the initial investment for a navigation system, not to mention the increased patient safety that can be achieved through CT navigation.

CT-navigated spinal instrumentation can improve safety and accuracy by providing precise 3D anatomical information, especially in degenerated anatomical conditions and anatomically critical areas with many superpositioning bone and tissue structures, e.g., the upper cervical, upper thoracic region, or pelvic region (Bühren and Josten, 2012). To gain the full potential of CT navigation, dedicated planning of screw trajectories in CT volumes is mandatory. However, the process is error-prone, time-consuming, and expertise-dependent. The manually created plans are used for intraoperative surgical navigation. Even though no gold standard for pedicle screw placement exists, special care has to be taken during the positioning of such screws not to perforate the pedicle wall or vertebra body. The insertion *direction* of optimally positioned screws should be aligned to the pedicle angle, and the *head* should not cross vertebra joints. The screw *length* should reach well into the vertebral body while not crossing the anterior wall, and the screw *diameter* should fill the pedicle region. Automatic screw planning in CT volumes can assist the surgeon in screw planning, reduce errors, standardize the quality, significantly



**Figure 2.7: Workflow of percutaneous pedicle screw placement** with posterior approach taken from Vaccaro et al. (2021): (A) incision localization, (B) Needle orientation, (C) Needle insertion, (D) K-wire insertion, (E) Screw insertion, (F) Repetition for selected vertebra levels, (G) Rod placement, (H) Wound closure.

improve time efficiency, reduce manual intervention, and facilitate the clinical workflow when integrated into navigation systems.

### 2.2.3 C-arm positioning for fluoroscopy-guided spinal surgery

Fluoroscopic images are the most commonly used assisted technique for spine surgeries because C-arm devices are easily accessible in most operating rooms (Perna et al., 2016). In contrast to CT-navigated spinal surgery that requires preoperative preparation like CT scanning, planning, data transfer, pre- to intraoperative registration, and the accessibility of a navigation system, fluoroscopy guidance delivers intraoperative projection images in real-time. Fluoroscopic guidance involves the repeated acquisitions of a.-p. and lateral X-rays with a mobile C-arm. The acquired standard projections allow to assess reduction and evaluate implant placement intraoperatively. Corrections can be performed immediately, which reduces surgical revisions.

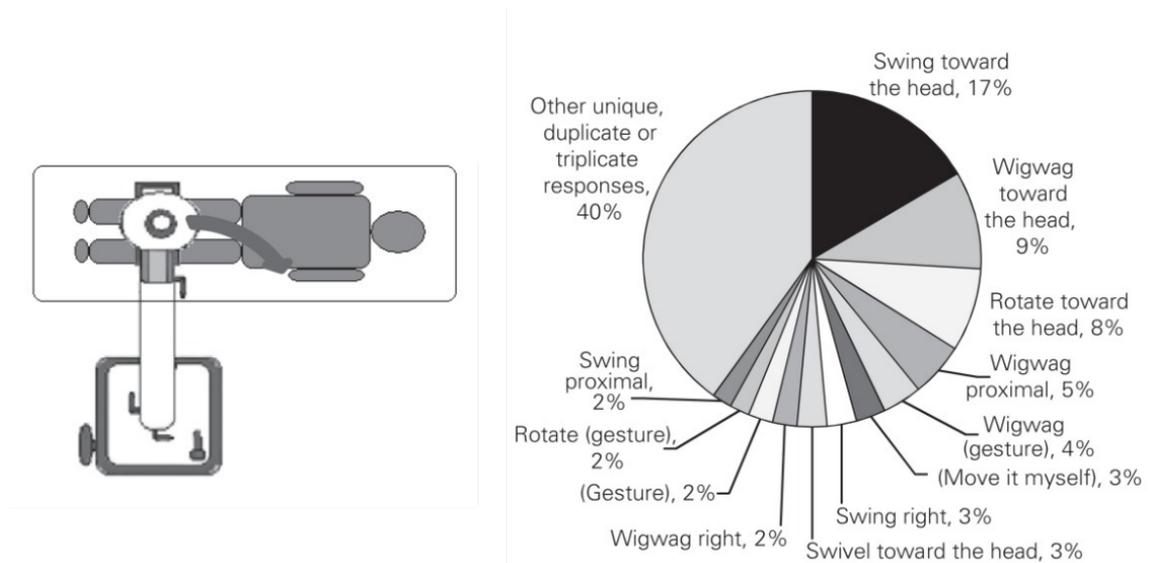
The percutaneous screw placement procedure consists of several steps as depicted in Fig. 2.7 that require fluoroscopic confirmation: The correct level of skin incision is confirmed with fluoroscopy (A). Blunt dissection of the subcutaneous tissue, fascia, and muscles is performed using scissors. The incision point is more lateral for the lumbar vertebra than for the middle thoracic spine (B). Cannulated needles (Jamshidi) are inserted through the pedicle corridor (C). The entry point is confirmed with a.-p. X-ray guidance

and the insertion is performed under lateral X-ray guidance. K-wires are inserted through the cannulated needles, and the needles are removed (D). The cannulated pedicle screw is inserted via the k-wires (E). Lateral X-ray projections confirm the insertion depth. The process is repeated for each vertebra level (F). The final instrumentation is verified by fluoroscopy. The rod is placed percutaneously, connecting the screw heads (G). The wound is closed (H).

During spinal fusion surgeries, the k-wires must be guided through narrow pedicle corridors. Thereby, the correct tool trajectories are monitored by X-ray projections from different poses. For intraoperative monitoring, guidance, and verification, the accurate acquisition of anatomy-specific standard projections is essential. Otherwise, the spinal instruments may be placed inaccurately. It is challenging to obtain the desired view due to variable patient placement and because the internal anatomy is not visible from the outside. In fluoroscopic-based orthopedic procedures, good reduction, optimal implant placement, and correct fluoroscopic visualization are closely related (Norris et al., 1999). A correct fluoroscopic view requires a good reduction, whereas a good reduction is a pre-requirement for optimal implant placement. Incorrect projections can result in overlays of anatomical structures and ambiguities due to the effects of projective simplification, thus increasing the risk of overlooked errors (Rikli et al., 2018). Examples of critical errors include the malunion of fractures, leading to functional impairment and, in the worst case, requiring a subsequent intervention at increased rates of complication. Even if correct standard projections are acquired, the assessment of optimal implant placement is still limited since fluoroscopic images are summation images. Exact placement control is only achievable with cross-sectional imaging; however, an experienced surgeon might be able to compensate for the reduction in spatial information. A study performed by Samdani et al. (2010) confirmed that fluoroscopy-guided screw placement accuracy is significantly improved for more experienced spine surgeons, indicating the steep learning curve.

C-arm positioning for standard projections is usually performed manually using a trial-and-error approach of iterative fluoroscopy, called fluoro hunting, at the cost of radiation exposure and intervention time that largely depends on the operator's experience (Blatter et al., 2004). The radiation exposure is also influenced by additional factors like anatomical variability, patient size, type of disease, bone density, prior imaging studies, and surgical approach to the spine (Reiser et al., 2017).

Typically, the C-arm positioning is performed by an operating room (OR) assistant guided by verbal commands of the surgeon. Pally and Kreder (2013) conducted a survey of terminology used for the intraoperative direction of C-arm. They discovered that



**Figure 2.8: Variations in terminology for C-arm positioning.** Summary of responses to the survey question: "How would you ask the technician to perform the movement in the right illustration?" adapted from Pally and Kreder (2013). The survey was performed among 261 members of the Canadian Orthopedic Association in Dec. 2009.

no standard terminology is employed (Fig. 2.8), which results in longer OR times and increased radiation exposure. Further, their survey revealed that many radiation technologists were inexperienced in operating the C-arm.

Reiser et al. (2017) evaluated across 240 thoracolumbar/lumbar minimally invasive spine procedures that 25% of the total radiation occurs during preoperative surgical planning and positioning. According to a study by Matthews et al. (2007), only 20% of all acquired fluoroscopic images are necessary for the procedure, while 80% accumulate during the iterative positioning process. Consequently, patients and clinical staff are exposed to unnecessary radiation. A systematic review of literature on radiation exposure of orthopedic surgeons is given by Matityahu et al. (2017). The highest exposure and subsequent equivalent doses were reported from spinal surgery (up to 4.8 mSv of equivalent dose to the hand). Radiation dose is harmful, and strategies to reduce exposure need to be considered.

The positioning radiation can be reduced by assisting the C-arm positioning procedure with simulations. Simulated guidance involves tacking of the C-arm movement and registration of a reference CT with the patient for digitally reconstructed radiograph (DRR)

simulation. Touchette et al. (2021) analyzed the effect of artificial X-rays on C-arm positioning performance. They found a significant 53% decrease in the number of acquired X-rays and an improved lateral C-arm positioning procedure of 10 – 26% for the pelvic anatomy. Incorporating the assistance system in the positioning procedure did not increase the total positioning time. Due to the inherent prior assumption like the availability of a preoperative CT or external tracking hardware, state-of-the-art assistance systems have so far limited broad clinical applicability. However, guiding and automating this procedure can reduce time and radiation exposure for patients and clinical staff and avoid fracture reduction or implant placement errors.

### 2.2.4 Conclusion

Different image-based procedures can assist spinal surgeries. Intraoperative C-arms are accessible in most operating rooms, and fluoroscopic guidance is a routine and fast procedure for guidance and control. It delivers intraoperative imaging in real-time. In contrast, CT-based navigation requires investment in a navigation system and preoperative preparation like CT scanning, trajectory planning, data transfer, and patient registration. Particularly in easy anatomical conditions and with experienced surgeons, fluoroscopic guidance can reduce the patients' doses, surgical time, and surgical costs. In contrast, the explicit 3D information provided by CT navigation may be advantageous for heavy volumes and complex surgical cases with degenerate anatomies where particularly high accuracy and safety are required, which otherwise would result in long fluoroscopic times. Despite different available techniques for pedicle screw placement, surgeons should choose the technique based on their experience and training (Perna et al., 2016).

Spinal surgery constitutes a challenging, multi-stage process that requires image guidance. Existing image guidance techniques require manual intervention that is highly expertise-dependent and impedes the clinical workflow. For the CT navigation technique, this comprises the pedicle screw planning and, for fluoroscopy guidance, the positioning of the C-arm for standard projections. These processes require a thorough knowledge of the spine anatomy and the anatomical orientation of structures, particularly the pedicles. The spatial orientation is additionally challenged by intraoperative requirements, i.e., variable patient positioning on the operating table, confounding factors resulting from surgical implants, patient-related factors (fracture types, obesity). Sasso and Garrido (2007) showed that both spinal surgery techniques require surgical training and are intensely learning curve dependent. If the workflow is not smooth, the intervention time gets extended. Therefore, both image guidance techniques offer potential regarding the support for the surgeon to decrease expertise-dependence of results, radiation exposure, and procedural

## *Chapter 2. Background*

---

time while at the same time facilitating the clinical workflow by requiring less manual intervention.

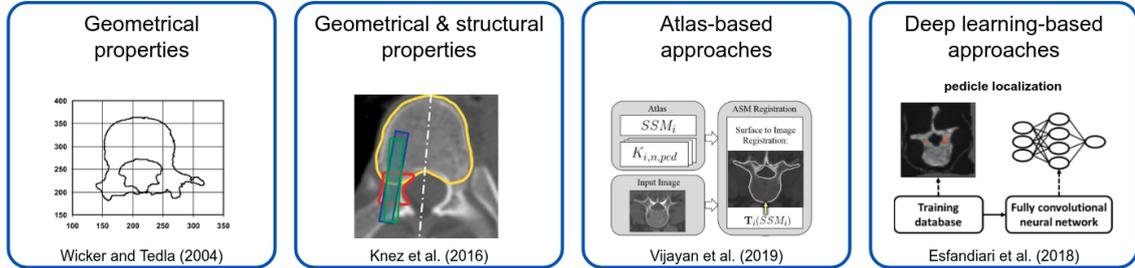
This chapter gives an overview of the related state of the art for different image-based assistance techniques during spinal surgeries. Spinal surgeries can be conducted with navigation systems or guided by 2D fluoroscopy. The choice of technique depends on a variety of factors, e.g., the experience of the surgeon, familiarity with the technique, accessibility of the technique, the complexity of the fracture and anatomy. Navigation can be performed based on a pre-operative 3D CT scan, where screw trajectories are determined prior to surgery. The state of the art methods for pedicle screw planning for CT-navigated spinal surgery are presented in section 3.1. Fluoroscopy-guided spinal surgeries involve the repeated acquisition of spinal standard projections for monitoring, guidance, and quality control. Standard projections correspond to a specific pose of the C-arm relative to the patient's position. Section 3.2 introduces the state of the art for C-arm positioning for fluoroscopy-guided spinal surgery. For further details on the presented methods, one can follow the cited papers.

### 3.1 Pedicle screw planning for CT-navigated spinal surgery

Various methods for pedicle screw planning have been reported that differ in their modeling assumptions (Fig. 3.1). While early methods employed only geometrical properties of vertebral structures (i.e., shape) to constrain the screw trajectory and dimensions, later methods combine this with structural properties (i.e., appearance) such as bone mineral density (BMD), and incorporate additional clinical parameters (e.g., screw diameter, length, patient age, spine curvature). Recently proposed methods for pedicle screw planning are mostly based on atlas registration. The approaches are presented individually in the following.

#### 3.1.1 Pedicle screw planning

The first approach to pedicle screw planning was presented by Wicker and Tedla (2004) who determined the screw trajectory and dimensions by fitting a line to the pedicle shape. The alignment is based only on 2D images and geometrical properties of segmented vertebra slices. Their method was derived in detail; however, it does not consider



**Figure 3.1: Classification overview of state of the art methods for pedicle screw planning.** Deep learning-based methods were only considered for pedicle localization and not for automatic pedicle screw planning.

spatial characteristics of vertebrae and lacks experimental results, and was not tested in clinical data. Subsequently, Lee et al. (2012) extended the geometric approach to 3D vertebra representations. Focusing on the lumbar segment, their approach for pedicle segmentation was successful in 93.3% of all cases and required a high computational time of 14 – 15 mins. The approach was not further extended to pedicle screw planning. Daemi et al. (2015) additionally considered the planning requisites of a safety margin and maximizing the insertion depth, which is correlated with the screw fixation strength (Azwan and Abd Rahim, 2011). Automatic planning results were only assessed in terms of safety margin and insertion depths due to a lack of reference screw trajectories.

In contrast to purely geometric modeling, Knez et al. (2016) combined 3D geometrical with structural properties and additionally constrained screw trajectories by an operational safety margin. Geometric properties of vertebrae were extracted by 3D parametric modeling of vertebral bodies and pedicles, which imposes limitations on the pedicle screw trajectories and dimensions. The shape constraint was combined with appearance properties by maximizing the screw fixation strength in an image-based concept. Therefore, the underlying BMD which is directly related to the fixation strength (Lehman Jr et al., 2003) was measured in Hounsfield units (HU) inspired by (Schreiber et al., 2011). They reported qualitative results on 81 thoracic pedicles of 11 cases with a mean absolute deviation and standard deviation ( $\mu \pm \sigma$ ) of  $0.4 \pm 0.4$  mm for screw diameter,  $5.8 \pm 4.2$  mm for the screw length,  $2.0 \pm 1.4$  mm for the pedicle cross point, and  $7.6 \pm 5.8^\circ$  for the insertion angle, indicating a relatively high agreement between manual and automatic screw plans. Automatic planning resulted, on average, in significantly shorter and thinner screws, which indicates that the fixed manual selected safety margin might not be optimal for all cases. Automatic planning led to increased screw fastening strength in 99% of all cases. Parametric vertebra and pedicle modeling has the advantage of not requiring any training data or manual

---

### 3.1. Pedicle screw planning for CT-navigated spinal surgery

annotations; however, the generalized shape description limits the application to local anatomical deformations. Further, each 3D parametric vertebra model requires manual initialization, followed by structure modeling at a runtime of 2 mins, and introducing a 3D modeling error of 0.5 mm, which serves as initialization for the subsequent pedicle screw planning. In additional work, Knez et al. (2018) extended this approach to address specific clinical aspects by constraining the pedicle screw entry points to follow the spinal curvature or to follow specific surgical insertion techniques. They reported a statistically significant improvement in screw planning accuracy compared to the original method while still suffering from the limitations related to parametric modeling. The parametric computer-assisted pedicle screw planning was clinically evaluated on 25 patients (204 screws) covering the thoracic and lumbar spine. It achieved a Gertzbein-Robbins (GR) classification Grade A in 88.7% (Knez et al., 2019).

Similarly, Xiaozhao et al. (2016) combined geometric and structural properties for screw trajectory planning. At first, vertebra bodies and pedicles were automatically segmented for screw trajectory initialization. Subsequently, the screw trajectory and dimensions were optimized so that the overlapped HU are maximized. Due to vertebra shape assumptions, the method is limited to lumbar vertebrae.

Recently, different atlas-based approaches for screw trajectory planning were introduced. Li et al. (2019) proposed to combine spine landmark detection with multi-atlas vertebra segmentation. Based on the segmentation, a distance map was created, representing the minimal distance of each voxel to the vertebral edge. Pedicle regions were determined based on the distance map, and safety regions within the pedicles were derived. Additional clinical parameters (screw diameter, length, patient age) can be incorporated into the optimization algorithm while the screw trajectory was constrained to pass through the pedicle safety region. However, the large number of mainly qualitative clinical parameters complicates the planning pipeline, and the importance of different parameters is difficult to assess and may be surgeon-dependent. A multi-atlas registration approach was presented by Goerres et al. (2017). An atlas consists of a vertebral surface with ideal reference trajectory annotations. The reference trajectories were transferred to the computed tomography (CT) using non-rigid surface registration and probabilistically aggregated to derive trajectory plans. In a leave-one-out study, their method demonstrated close agreement with expert reference plans on selected vertebra levels T7, T8, L3, where the atlas comprised only a limited number of 60 screw trajectories. However, the method may be limited in application to altered, deformed, degenerated, or fractured bones since atlas-based approaches are targeted to represent the mean anatomical shape. Vijayan et al. (2019) built on this approach and replaced the vertebra surface multi-atlas with a vertebra level-specific statistical shape model. The method was evaluated in a leave-one-out study

on 200 patients for the lumbar spine; 9 cases were excluded in the analysis because the initialization of the anatomical atlas failed. On the remaining 191 cases (382 screws), they reported an accuracy of  $\mu \pm \sigma$  of  $2.4 \pm 1.6 \text{ mm}$  for the entry point and  $3.6 \pm 0.8^\circ$  for the insertion angle. The statistical shape model registration is sensitive to initialization and requires separate training for each vertebra level.

So far, deep learning-based methods have not been applied for the task of pedicle screw planning. For pedicle localization Esfandiari et al. (2018) presented an approach targeted to initialize the 2D-3D registration of pre- and intraoperative images. The method works on 2D CT image slices and can handle one vertebra at a time. It was trained and tested on CT scans acquired from synthetic lumbar bone models. Thus, it lacks real data evaluation and does not take 3D structural properties into account. Further, the pedicle locations are automatically identified; however, automatic screw trajectory planning was not addressed.

### 3.1.2 Conclusion

Many state of the art approaches for pedicle screw planning rely on manually defined geometric and structural entities with explicitly defined clinical constraints. The insertion trajectories are optimized with respect to pedicle geometry or bone density employing parametric modeling and defined assumptions. While parametric modeling is targeted to represent the general anatomical shape, its generalization to locally deformed vertebrae may be limited. Further, experience-based assumptions, e.g., that the pedicle center line determines the optimal screw trajectory or defining a fixed safety margin that might not be ideal depending on the vertebra shape. Instead of relying on manually defined properties, subsequently proposed atlas-based techniques derive the geometric constraints statistically from a set of vertebra shapes with expert annotations. Since the atlas represents normal anatomical vertebra shapes, its application may be limited in deformed, degenerated, or fractured vertebrae. Furthermore, the statistical shape model-based atlas is vertebra level-specific, requires retraining, and is sensible towards the initialization. Deep learning-based applications have proven success in many medical imaging applications; however, they were not yet applied in the context of 3D pedicle screw planning. In this work, deep learning is employed to learn the screw trajectories directly from clinical data, leveraging a large intraoperative screw planning dataset. The screw placement is implicitly learned from a diverse surgical planning dataset that takes individual anatomical conditions into account, as well as the surgeon's preference for screw placement. The problem of screw trajectory planning is tackled in a vertebra instance-based approach, implicitly taking geometrical and structural properties into account.

## 3.2 C-arm positioning for fluoroscopy-guided spinal surgery

A general problem related to the task of automatic C-arm positioning is given by the lack of annotated fluoroscopic data. Therefore, this section introduces state of the art simulation techniques for X-ray images (Sec. 3.2.1). Specific methods for fluoroscopic image understanding have been proposed and are revised in Sec. 3.2.2. This includes methods for view-independent fluoroscopic segmentation and landmark detection. Sec. 3.2.3 gives an overview of how these methods were applied in the context of C-arm positioning so far.

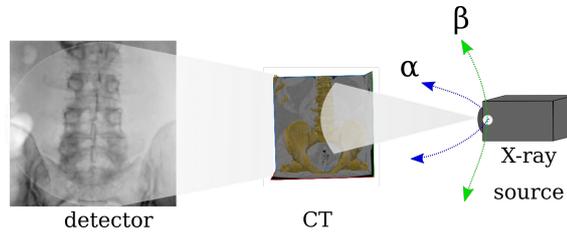
### 3.2.1 Data simulation

Deep learning methods showed excellent performance for various medical image processing tasks (Maier et al., 2019). However, their success strongly depends on the amount and quality of the training data, which limits their application in interventional imaging, particularly 2D fluoroscopy-guided procedures, due to several factors:

- **Lack of archiving:** Many intraoperative X-rays are acquired for monitoring, guidance, and quality control. However, only very few X-rays that document the procedural outcome are archived.
- **Lack of annotations:** Learning targets are not well established or defined, and manual annotation is associated with a high workload given the high amount of data.
- **Multiple view points:** The patient to C-arm alignment is arbitrary, resulting in varying 2D views onto the anatomy.
- **Individual variability:** Confounding factors caused by surgical implants or instruments, pathologies, or operating table present in the projection images increase the data variability, which is challenging to address with augmentation strategies.

An alternative to using real clinical X-rays is *in silico* fluoroscopy simulation from 3D diagnostic CT scans (Fig. 3.2). The synthetic X-ray is referred to as digitally reconstructed radiograph (DRR). *In silico* simulation has the advantage that it uniquely provides ground truth labels for arbitrary many training examples. Since the reference system is known, it enables the annotation of pose labels. Moreover, the annotation process is simplified and accelerated because it can be performed directly in 3D.

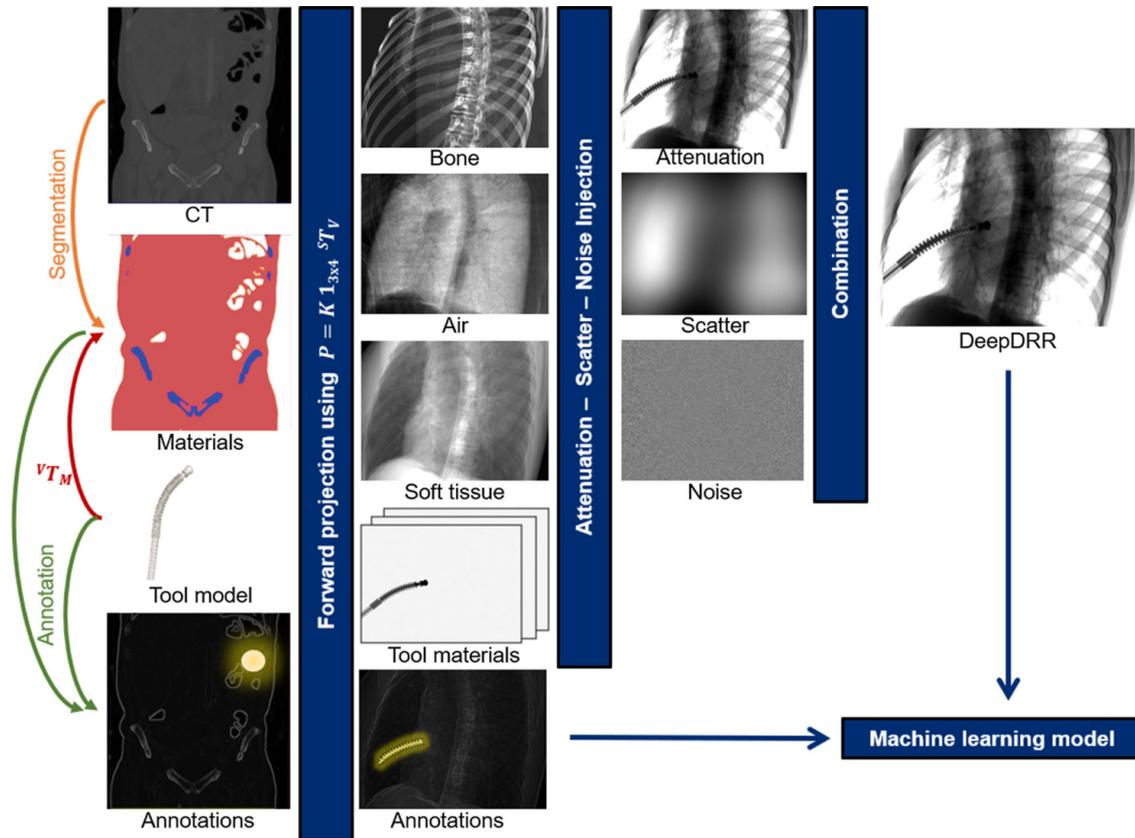
The DRR simulation task requires determining detector responses for a given imaging geometry according to Lambert-Beer's law. Approaches can be classified into analytic



**Figure 3.2: Data simulation: Synthetic X-rays** (Digitally reconstructed radiographs) can be simulated from a computed tomography (CT) scan for arbitrary poses.

ray tracing and statistical Monte Carlo (MC) methods. Analytic approaches compute the detector responses by applying attenuations that are summed up along a 3D line to all photons emitted in that direction (Ghafurian et al., 2016; Russakoff et al., 2005). Traditional ray tracing models assume mono-energetic X-ray sources that pass through single-material objects, which is violated by clinical X-ray imaging. Consequently, synthetic images do not reflect typical artifacts such as beam hardening and scattering noise. In contrast, MC simulations determine single-photon attenuations from a sequence of probabilistic material- and energy-dependent photon-matter interactions (Badal and Badano, 2009). Though the simulations are very realistic, MC simulations are not suitable to generate large machine learning training sets due to the long computation times (Unberath et al., 2018b).

An issue related to learning from simulations is the effect of the so-called domain shift, i.e., a shift between training (synthetic DRRs) and testing (clinical X-rays) data distributions. Machine learning methods trained on traditional ray-tracing simulations often restrict their evaluations to synthetic images generated by the same simulation technique (Albarqouni et al., 2017; Kügler et al., 2020; Terunuma et al., 2018; Würfl et al., 2018). Generalization to clinical X-rays is limited due to inaccurate X-ray image formation simulation assumptions (Zhang et al., 2018). One approach to tackle this problem was introduced by Zhu et al. (2017) who proposed unsupervised domain adaptation by unpaired conditional image-to-image style transfer networks. They addressed the issue of translating an image from a source domain to a target domain in the absence of aligned image pairs for both domains. Zhang et al. (2018) showed that machine learning methods trained with traditional ray tracing simulations can be applied to clinical X-rays if those are preprocessed with a style transfer network to mimic the appearance of DRRs. However, compared to supervised training on labeled clinical X-rays, performance degradation can still be observed. Dhont et al. (2020) proposed a DRR simulation framework that combines a forward projector with a conditional generative adversarial network (cGAN) to translate



**Figure 3.3:** Schematic overview of *DeepDRR* realistic simulation pipeline taken from Unberath et al. (2019).

the synthetic forward projection to a realistic DRR. However, this approach requires sufficient data of the target domain. If the samples do not represent the target domain, the method will not generalize. This poses a significant challenge to medical imaging, especially intraoperative imaging, where available clinical data is scarce. Kazemina et al. (2020) identified several drawbacks of generative adversarial network (GAN) for medical imaging. Firstly, the trustability of the generated data is limited since the generation process is based on deep neural networks, which are still not sufficiently understood, and image intensities are not directly associated with a specific material characteristic. Secondly, GAN training is unstable, and evaluation with unlabeled data is challenging. Thirdly, GAN suffer from the uninterpretability, which impedes their application in medical environments.

Recently, Unberath et al. (2018b, 2019) proposed a framework for fast and realistic

forward projection and noise injection combined with 3D material decomposition and 2D scatter estimation. Fig. 3.3 gives a schematic overview of the simulation pipeline. It consists of four steps:

1. CT volumes are decomposed into materials  $M = \{\text{air, soft tissue, bone}\}$  using thresholding or a U-Net model for segmentation (Ronneberger et al., 2015).
2. Forward projections are generated by a material- and energy-dependent ray-tracing projector.
3. A neural network-based Rayleigh scatter estimation is added.
4. Quantum and electronic readout noise is injected.

Given the X-ray spectral density  $p_0(E)$  and a projection matrix  $\mathbf{P} \in \mathbb{R}^{3 \times 4}$ , the total X-ray attenuation in detector pixel  $\mathbf{u}$  can be computed by

$$\begin{aligned} p(\mathbf{u}) &= \int p(E, \mathbf{u}) dE \\ &= \int p_0(E) \exp \left( \sum_{m \in M} \delta(m, M(\mathbf{x})) (\mu/\rho)_m(E) \int \rho(\mathbf{x}) d\mathbf{l}_{\mathbf{u}} \right) dE, \end{aligned}$$

with  $\delta(\cdot, \cdot)$  denoting the Kronecker delta,  $(\mu/\rho)_m(E)$  denoting the material- and energy-dependent mass attenuation coefficient (Hubbell and Seltzer, 1995),  $\rho(\mathbf{x})$  denoting the material density derived from Hounsfield units (HU),  $d\mathbf{l}_{\mathbf{u}}$  representing the ray passing from the source to the detector pixel  $\mathbf{u}$ , specified by the projection matrix  $\mathbf{P}$ . Given the projection domain image  $p(\mathbf{u})$ , the Rayleigh scatter is estimated using a convolutional neural network (CNN) trained on MC simulations. Subsequently, quantum noise is modeled analytically by a Poisson noise model, and electronic noise is modeled as additive Gaussian noise, resulting in a realistic DRR simulation. Unberath et al. (2019) extended the original framework to enable the integration of tools or surgical implants. At positions where the 3D CT is overlapping with the surgical object, HU were omitted to model insertion.

Unberath et al. (2019) validated in an ablation study that more accurate DRR simulation models improved the performance of CNNs on real data. The performance boost can mainly be attributed to the energy- and material-dependent ray-tracing modeling, while the addition of scatter did not significantly improve the performance. Bier et al. (2018b) demonstrated generalization to real X-rays without re-training or domain adaptation for the task of pelvic landmark detection. In contrast, Unberath et al. (2019) observed compromised performance on real data for the task of landmark detection and robotic

end-effector localization. They argued that the degradation results from several factors such as inaccurate reference annotations, vendor-specific preprocessing of the clinical X-rays, or simulation from low-resolution CTs.

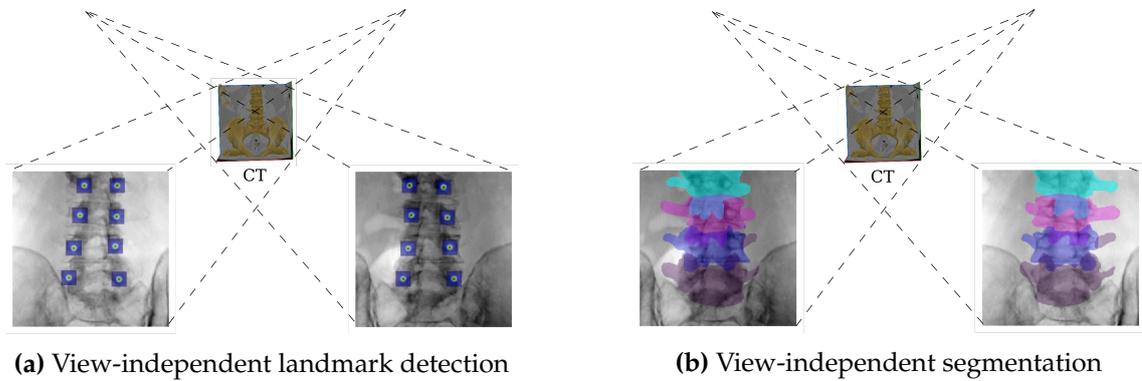
Learning from simulations leads to the challenge of transferring to real-world applications. Already presented approaches were based on transfer learning (Zhang et al., 2018) or realistic tool modeling (Kügler et al., 2020; Unberath et al., 2019). In contrast, Kordon et al. (2020) proposed a novel data augmentation approach for simulation of surgical instruments, tools, and implants that is based on augmentations with transformed characters, numerals, and abstract symbols from open-source fonts. Additionally, they introduced coarse random pixel dropout and showed it increases robustness towards superimposed surgical objects. Toth et al. (2019) proposed another approach for domain adaptation based on domain randomization (DR) which does not require data of the target domain. DR adds simulated random perturbations to the training data, e.g., collimation on X-rays or random HU before projection. The transformations create such a large variety of simulated training data that real images become just another variation. They observed more consistent results on the task of cardiac image registration.

To sum up, learning in 2D fluoroscopy-guided procedures highly depends on simulated data. Due to the long computation times of MC simulations, most models are trained on ray-tracing simulations. Learning from simulation leads to the effect of the so-called domain shift. Approaches that address this limitation can be classified into those that aim for a more realistic ray-tracing simulation and those that combine traditional ray-tracing with transfer learning to address the domain shift. Furthermore, DR can address the domain gap without requiring data of the target domain. The high variation of intraoperative X-rays caused by confounding factors like superimposed surgical tools poses an additional challenge. Realistic and random tool modeling techniques have been proposed to represent the variety in the training data for specific applications.

#### 3.2.2 Data analysis

Fluoroscopy images are summation images that result from the projection of the 3D anatomy on the detector plane. Specific challenges can be attributed to the imaging device and the intraoperative environment:

- **Overlapping anatomical structures:** Due to the projective imaging, anatomical structures overlap in fluoroscopy images, and the depth information is missing.
- **Confounding factors:** Surgical tools or equipment can superimpose the projection images.



**Figure 3.4: Data analysis: View-independent landmark detection and segmentation.**

- **Changing viewing directions:** Different C-arm to patient poses lead to varying appearances of anatomical structures.

Overlaps of the anatomy, superimposed surgical objects, and varying viewing directions complicate the understanding of fluoroscopic images. Providing implicit 3D information through annotations in the projection domain can support the 2D-3D mental mapping and benefit intraoperative decision-making. Additional contextual information can be provided through segmentation masks or anatomical landmarks extracted from the X-ray. While labeling 2D annotations in X-rays is challenging and time-consuming due to large overlaps in the anatomy and the vast amount of data, 3D annotations are easier to annotate due to sharper boundaries and non-overlapping structures. Thus, combining the DRR simulation with automatic 3D annotations reduces the labeling effort drastically and, at the same time, leads to consistent annotations in the projection domain. Most methods presented in the literature restrict their representation to a specific viewing direction. Recently, some approaches that generalize to arbitrary views were introduced and are revised in the following.

**View-independent landmark detection:** View-independent landmark detection is particularly challenging in X-ray images because landmarks may appear differently depending on the viewing direction (Fig. 3.4a). Among others, Laina et al. (2017) proposed to model landmark locations as heatmaps where every pixel represents confidence proportional to its proximity to the landmark location. Heatmap representation was shown to increase performance compared to explicit regression of 2D landmark coordinates. By modeling the landmark locations spatially, global context was exploited for localization. In contrast to direct regression, the heatmap modeling can handle an arbitrary number of

landmarks resulting from occlusion or field-of-view (FoV) variation. Further, heatmaps can represent annotation and prediction uncertainties.

Bier et al. (2019) presented the first approach to detect anatomical landmarks of the knee anatomy in X-ray images independent of their viewing direction. The method is based on CNN prediction consisting of multiple stages Wei et al. (2016). The network architecture was initially developed for human joint pose estimation in RGB images. They showed the accuracy is sufficient for global initialization of 2D-3D registration tasks. Training data was simulated from 20 CTs, and evaluation was performed on synthetic DRRs generated from one CT and real X-rays from five specimens. For the synthetic evaluation, an average landmark error of  $9.1 \pm 7.4$  mm in the detector plane was reported, while for the real X-rays, the error was in the range of 12 – 24 mm in the detector plane. Partly this can be attributed to unrepresented scenarios such as tools and fractures during training. Also, they observed that the accuracy from views at the viewport border is slightly inferior compared to frontal views. This can result from a higher amount of anatomical overlaps from these directions and can potentially be addressed by increasing the viewport during training. Further, the sequential CNN architecture yields downsampled predictions, which limits the accuracy of detection. To preserve the input resolution, the network could be replaced by encoder-decoder architectures with skip connections. Esteban et al. (2019) extended this work by incorporating a patient-specific refinement scheme generated by back-projecting predicted landmarks that are used for network re-training after refinement. No analysis on real X-rays was conducted.

Bier et al. (2018a) transferred the CNN-based sequential prediction framework for landmark detection to X-rays from arbitrary viewing directions from the knee anatomy. Leave-one-out training was based on simulations from 16 CTs. On synthetic data, a mean landmark error of  $8.4 \pm 8.2$  pixel was achieved at a resolution of 0.6 mm. In contrast, validation on real clinical data was limited due to practical issues like a second leg in the FoV and was not evaluated quantitatively. Further, the authors attribute limited accuracy to the downsampling inherited by the network architecture, resulting in a factor of around eight times.

Liao et al. (2019) perform 2D-3D registration via landmark detection. They randomly sampled points in the 3D pre-intervention volume to train a set of view-specific 2D descriptors that learn to establish 2D point-to-point correspondences between DRRs and X-rays that are particularly robust to translation offsets. The 2D tracked points were subsequently triangulated in an end-to-end fashion. The method showed promising pose estimation results for the task of 2D-3D registration. However, the capture range

was limited to  $10^\circ$  rotation offset. Due to the use of triangulation, the method requires X-rays from at least two views. Since this approach is designed for 2D-3D registration, a preoperative CT scan is required.

Liu et al. (2020) proposed a CNN-based landmark prediction, which additionally models relationships among the anatomical landmarks by incorporating an extra relation loss in a multi-task learning strategy. The relationships were defined by edges according to the clinical significance to model their relationships. The method was evaluated on anterior-posterior (a.-p.) hip X-rays under changing hip appearances and showed the highest accuracy boost when combined with the Hourglass network architecture (Newell et al., 2016).

To reflect uncertainties in landmark annotations, Payer et al. (2020) proposed to learn the Gaussian covariances of target heatmaps. Landmark location distributions were obtained by fitting Gaussian functions to predicted heatmaps.

**View-independent segmentation:** Segmentation algorithms for fluoroscopic images have to deal with overlapping classes since multiple anatomical structures can overlap in one pixel (Fig. 3.4b).

Viewport independent segmentation for fluoroscopic images was first proposed by Thomas et al. (2019) for the ankle anatomy. They addressed the problem of segmenting overlapping bones in fluoroscopic images using a U-Net approach. Training was conducted on clinical fluoroscopic images of the ankle joint. Incorporating the predicted segmentation masks as additional statistical shape model refinement information led to more robust and accurate model representations.

Concurrent segmentation and landmark localization was first introduced by Laina et al. (2017) for tracking of surgical instruments. They reformulated the 2D instrument pose estimation as a heatmap regression and simultaneously optimized the two tasks in one deep learning architecture. The approach was evaluated on retinal microsurgery data and endoscopic data. The interdependent modeling led to significantly improved performance compared to direct landmark regression. Gao et al. (2019) built on this to perform concurrent segmentation and landmark detection of continuum dexterous manipulators in X-ray images of the femoral head. It was employed for the robust initialization of 2D-3D registration. Training was based on synthetic X-rays generated from four CTs complemented with 2D segmentation masks and key landmarks represented as Gaussian distributions. Their architecture is based on a U-Net, where the final feature

### 3.2. C-arm positioning for fluoroscopy-guided spinal surgery

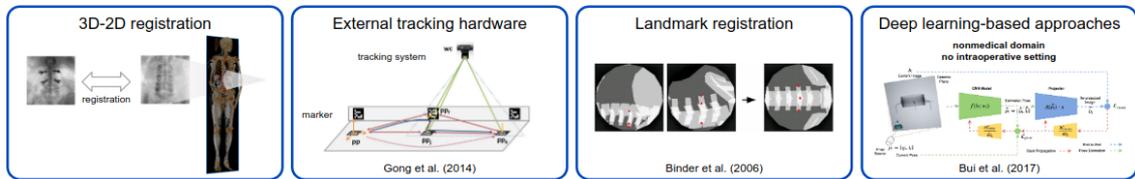
---

maps are shared across the segmentation and landmark detection path. The final predicted segmentation was concatenated with the shared feature maps to boost landmark detection. Evaluation was performed on synthetic X-rays generated from one CT resulting in a mean DICE score of  $0.996 \pm 0.001$  and a mean landmark error of  $0.365 \pm 0.345$  mm. Evaluation on 87 real X-rays result in a mean DICE score of  $0.915 \pm 0.063$  and  $2.54 \pm 0.95$  mm. The performance drop was caused by tool configurations never seen during training.

Grupp et al. (2020) proposed a method for iterative intensity-based 2D-3D registration regularized by anatomical annotations of the hip anatomy. Annotations were automatically produced by a CNN, which was trained for simultaneous anatomy segmentation and landmark detection. Therefore, an image was processed by a U-Net, and the segmentation was computed subsequently. Simultaneously, the segmentation features concatenated with the feature map output of the U-Net were used to estimate the landmark heatmaps. Predicted segmentation maps were employed to restrict candidate landmark locations, avoiding false alarms. Training was conducted on clinical data resulting from six specimens in a leave-one-out study. A unique class of anatomy was assigned for each 2D pixel during training data generation according to a precedence ranking. On real X-rays of the intact hip anatomy, they reported an average landmark localization error of  $5.0 \pm 5.2$  mm in the detector plane. Further, semantic labeling of fluoroscopy can be used to restrict the FoV during 3D-2D registration and limit the DRR simulation to the relevant range.

To sum up, many recent approaches address the scarcity problem of fluoroscopic X-rays with semantic annotations by training on synthetic X-rays. 3D annotations are easier to establish, reduce the labeling effort, and increase 2D consistency. However, the domain shift often results in a performance drop when applying algorithms trained on synthetic X-rays to clinical X-rays. These performance drops can primarily be attributed to scenarios not seen during training. This issue can be addressed by increasing the for the most approaches limited number of CT volumes during training, thereby representing more anatomical variation. Further, a crucial factor for generalization to intraoperative X-rays is suitable modeling of confounding factors like surgical tools, implants, or operating table.

For the task of view-independent landmark detection, the heatmap representation has shown to be superior to direct coordinate regression. Different architectures were proposed, revealing that the encoder-decoder design is superior to a multi-stage sequential CNN since the input resolution is preserved, enabling more accurate localization. Liu et al. (2020) compared different encoder-decoder architectures like U-Net (Ronneberger et al., 2015), Hourglass (Newell et al., 2016), and HRNet (Sun et al., 2019) indicating no significant performance difference. Recently, different approaches were presented that



**Figure 3.5: Classification overview of state of the art methods for C-arm positioning.**

combine segmentation with localization and learn shared feature maps in order to exploit contextual information. This is shown to benefit training, especially for scarce datasets like annotated real fluoroscopic X-rays.

Semantic information extracted from X-ray images has shown to benefit computer-assistance systems in fluoroscopy-guided procedures across various tasks.

### 3.2.3 C-arm positioning

Many state of the art assistance systems for C-arm positioning either require preoperative CT scans, external tracking hardware, or manual landmark selection, which so far limited broad clinical applicability (Fig. 3.5).

Fluoroscopy simulation methods that generate virtual fluoroscopy for pose guidance rely on a preoperative CT scan or an anatomical atlas. In addition, they require an external tracking system to estimate the alignment between C-arm and CT each time the patient position is varied (Bott et al., 2011; Gong et al., 2014). Integrating tracking systems in the clinical workflow is challenging, mainly due to the related line-of-sight requirements and the additional hardware setup, which severely interferes with the surgical workflow. Consequently, the application of these approaches is currently focused on surgical training purposes. Otake et al. (2011); Bodart et al. (2018) proposed the use of X-ray visible fiducial markers placed intraoperatively to register different images. The placement of fiducial markers is a tedious and time-consuming procedure. Recently, a 3D-2D registration between a preoperative CT scan and initially acquired radiographs was shown to allow the generation of virtual fluoroscopy images in real-time (De Silva et al., 2017; Fallavollita et al., 2014). Since the 3D-2D registration is purely image-based, the registration performance is influenced by anatomical changes during the orthopedic procedure or metallic objects in the FoV. Further, a preoperative scan is still required. All previously mentioned approaches are designed to support the surgeon but not directly deliver the optimal pose.

### *3.2. C-arm positioning for fluoroscopy-guided spinal surgery*

---

Without the use of markers, Haiderbhai et al. (2019) presented a user interface for automatic C-arm positioning. The surgeon can define an optimal view based on simulated X-rays from a preoperative CT. This was converted to a C-arm position using inverse kinematics (Matthäus et al., 2007; Wang et al., 2012). To align the patient position during the surgery with the CT scan, a point cloud of the patient's body taken with a Kinect was registered to the point cloud generated from the surface body of the CT scan. It requires a registration of the preoperative CT and the patient each time either is moved. Since the registration is based on surface point cloud registration, the alignment of bones can be inaccurate, especially in case of fractures or unstable anatomical conditions. Maier et al. showed the potential of using inertial measurement units for object motion correction in C-arm imaging (Maier et al., 2018).

Shao et al. (2014) proposed a virtual reality aided positioning method. A 3D surface appearance model was reconstructed from multiple RGB images acquired by a camera augmented mobile C-arm. Subsequently, the preoperative CT was aligned with the appearance model so that a reference pose is determined in regard to the appearance. All mentioned approaches require a preoperative CT, which availability cannot be assumed for every procedure. Registration is required each time the patient or C-arm is moved. Pre- to intraoperative registration is additionally complicated by altered patient anatomy in the intervention or intraoperative presence of tools.

Given the model of the object of interest and at least one projection, Soltow and Rosenhahn (2015) presented a method for X-ray pose estimation based on contour matching. Given the segmented contour of the object in the images, the registration was initialized by matching computed Fourier descriptors to a database of pre-defined, simulated projections. The pose estimation was further refined by employing global and local optimization strategies to minimize the distance between the given contour and the model's projection given the current pose estimate. The approach was evaluated for pose estimation of radiopaque implants with clearly visible boundaries and not for bones where lower contrast differences with the surroundings hinder contour segmentation.

Without the use of marker or additional prior assumption like preoperative CT, Schaller et al. (2009) used a time-of-flight sensor to achieve automatic inverse C-arm positioning. Employing body part localization on the computed 3D surface point cloud, the required body part can be centered automatically within the isocenter of the C-arm. The method automatizes the initial patient to C-arm alignment; however, C-arm orientation is not considered in this work.

Other approaches do not estimate an optimal pose but address the problem of re-aligning the scanner with intraoperatively recorded C-arm views employing augmented reality (Fotouhi et al., 2019; Unberath et al., 2018a). The former requires a preoperative CT scan for initialization via 2D-3D registration, while the latter requires a technician with a head-mounted augmented reality display in the loop. Both methods do not automatically deliver the optimal pose but were designed to reproduce already defined C-arm positions.

In addition to the correct anatomical representation of X-ray projection generated from optimal poses, other factors can also be taken into account. Rodas et al. (2017) presented a MC approach to optimize the pose of a C-arm in a defined neighborhood around the known standard direction while reducing the radiation exposure. However, this so-called standard direction depends on the positioning of the patient relative to the C-arm and on the patient-specific anatomy, which is unknown in general. Zaech et al. (2019) presented an approach for task-aware trajectory recommendation to improve image quality for reconstruction in order to reduce metal artifacts in cone-beam computed tomography (CBCT). A VGG regression network (Simonyan and Zisserman, 2014) was trained to regress the optimal out-of-plane rotation offset in a defined neighborhood around the current pose. View optimality was assessed in terms of detectability measure, which is estimated only based on the current X-ray projection. The model was trained and evaluated on synthetic X-rays generated from five chest CTs. Thies et al. (2020) extended the evaluation to different noise levels and initialization angles and performed a retrospective phantom study with real X-rays. Reconstruction results indicate that task-aware trajectories reduce metal artifacts. Since each prediction is only based on one preceding X-ray, trajectory increments tend to jump for the out-of-plane rotation angle. Sequential modeling needs to be addressed to assure real-time trajectory scanning.

An alternative approach for image-guided positioning of C-arms based on landmark registration was presented by Binder et al. (2006). Therefore, the surgeon needs to manually define selected vertebra landmarks in two X-rays from different views. The a.-p. direction was automatically derived using triangulation methods. It comes with the limitation of requiring two distinct projection images with known pose offset and corresponding visible landmarks, manually depicted by the surgeon during the intervention.

One of the earliest deep learning-based pose regression approaches was presented by Miao et al. (2016) for 2D-3D registration. They divided the pose space into zones and trained zone-specific CNNs to iteratively refine the current pose estimate given the difference between the X-ray and the rendered DRR from the current pose estimate. Later, Miao et al. (2018) formulated the task as a Markov decision process and proposed a

multi-agent system to handle various artifacts in 2D X-ray images.

Pose regression without a preoperative scan has been proposed for slice transformation prediction with respect to a canonical atlas coordinate frame by Hou et al. (2017). On DRR computed from preprocessed CTs (identical intensity ranges, spacing, translation), they achieved sufficiently accurate performance (average translation error of 106 *mm* and 5.6° plane rotation for healthy patients) for the task of robust initialization of X-ray to CT registration. The method was not evaluated on real X-rays. Bui et al. (2017) presented a machine learning-based pose estimation for mobile X-ray imaging. They focused on industrial applications where the existence of a precise 3D CAD model can be assumed as prior knowledge, serving as an additional reprojection constraint. The experiments indicate that neural networks outperform other regression approaches in the prediction of object poses from simulated X-ray projections, with high pose prediction accuracy for object-specific models and decreasing accuracy for more generic CNN models trained on sets of different objects. The object-specific network was trained from one single CAD object and tested on unseen poses of the same object. In contrast, medical C-arm devices considered in this thesis need to handle data scarcity, inter-patient anatomical variation, and intraoperative requirements.

#### 3.2.4 Conclusion

This work aims at developing an approach for automatic C-arm positioning that considers the aforementioned challenges and limitations:

(1) To tackle the problem of automatic C-arm positioning in a deep learning-based approach requires annotated training data. Due to the lack of a constant reference frame in interventional fluoroscopy, X-rays with annotated pose labels do not exist. State of the art DRR simulation strategies can address the interventional data scarcity problem and facilitate robust deep learning.

(2) Fluoroscopic images involve specific challenges for image understanding due to overlapping anatomical structures, the presence of surgical implants, and changing viewing directions. Extracting semantic information from X-ray images, including segmentations and landmarks, can benefit fluoroscopic image understanding.

(3) Many state of the art C-arm positioning approaches rely on a preoperative CT scan to estimate the pose of an X-ray via 2D-3D registration. However, the availability of a CT scan cannot always be assumed and comes at the cost of additional radiation. In

addition to this requirement, the registration performance is limited by the fact that the patient's anatomy is altered during the orthopedic procedure, and surgical tools may be present in the intraoperative X-ray. Moreover, the application of assistance systems in the clinic is often limited due to the necessary hardware requirements that impede the clinical workflow. Thus, C-arm positioning is often performed manually, where the dose and time consumption depends on surgical expertise.

This work aims at tackling the C-arm positioning task without requiring a preoperative CT or other technical equipment to facilitate seamless integration into the clinical workflow. Further, the proposed approach should decrease the positioning radiation caused by the current manual procedure and reduce the overall positioning time. Deep learning-based pose estimation approaches have demonstrated superior performance on industrial objects. Regarding a medical C-arm device, the method has to learn variable anatomical representations to handle inter-subject anatomical variation that does not occur for industrial objects and to deal with data scarcity and confounding factors resulting from the intraoperative setting.

## Materials and Methods

This chapter presents the proposed methods for computer-assisted spinal surgery. Spinal instrumentation can be guided by different image-assisted procedures, e.g., CT-navigation or fluoroscopy-guidance. The following chapter is divided into two parts: Section 4.1 presents assistance methods for CT-navigated spinal surgery, focusing on automating the pedicle screw planning, and in Section 4.2 assistance methods for C-arm positioning during fluoroscopy-guidance are proposed.

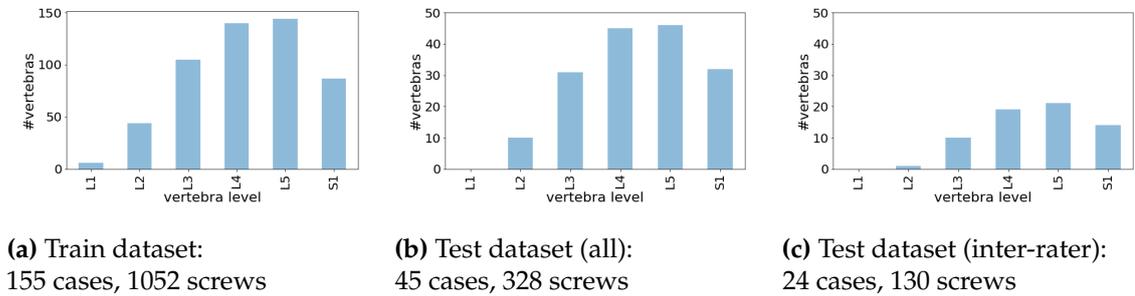
### 4.1 Pedicle screw planning for CT-navigated spinal surgery

#### Disclosures to this work:

Part of this work has been published in the SPIE Medical Imaging Conference on Image-Guided Procedures, Robotic Interventions, and Modeling and is submitted for publication to the Journal of Neurosurgery - Spine (Kausch et al., 2021a; Scherer et al., in submission).

Pedicle screw planning under CT navigation has been shown to significantly increase accuracy and safety by providing explicit 3D information (Mason et al., 2014; Kosmopoulos and Schizas, 2007). A necessary task in this process is the screw trajectory planning before surgery, which is often performed manually by the surgeon. Manual planning requires a high level of expertise and in-depth knowledge of the vertebra anatomy and pedicle orientation and comes at the cost of extended procedural time.

A computer-assisted method for automatic pedicle screw planning that tackles the challenges mentioned above is proposed in this thesis. For automatic image-based pedicle screw planning, a sizeable intraoperative planning dataset is leveraged. In contrast to the state of the art, the proposed approach does not rely on manually defined geometric and structural properties or atlas-based modeling, which is by design biased towards the mean anatomical shape, limiting its application to altered, deformed, degenerated, or fractured bones. Instead, the screw placement is implicitly learned from expert annotations that



**Figure 4.1: Vertebra level-specific screw planning dataset distribution.**

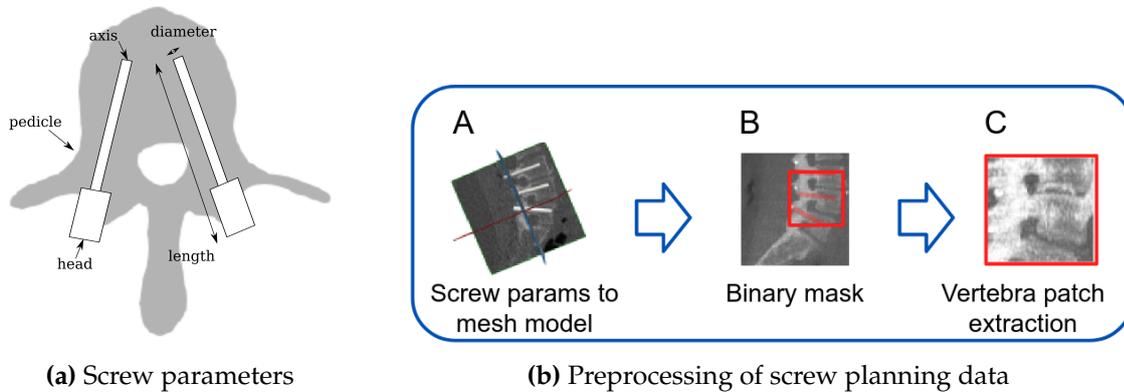
take patient-specific anatomical conditions and the surgeon’s preference into account. In addition, the patch-based design considers individual vertebrae at a time that benefits generalization towards spinal deformities.

The analysis focuses on the lumbar spine (first lumbar vertebra (L1)-first sacral vertebra (S1)) because fusion is most commonly performed in the lumbar region (Rajae et al., 2012). However, with available training data, the method can be extended to the cervical and thoracic level.

This chapter presents the retrospective screw planning data (Sec. 4.1.1), the baseline approach based on direct convolutional neural network (CNN) landmark regression (Sec. 4.1.2) and the proposed approach where the screw planning task is interpreted as a segmentation task (Sec. 4.1.3). The chapter is concluded in Sec. 4.1.4 with an overview of the developed software prototype for automatic screw planning that was employed for a clinical applicability evaluation.

### 4.1.1 Retrospective planning data

This section presents the anonymized retrospective planning datasets used for training and evaluation. All recordings were acquired at the university hospital Heidelberg with a Siemens SOMATOM<sup>®</sup> Emotion CT Scanner. The data was derived retrospectively from a registry of navigated spinal instrumentations (n=1660) performed from 01/2010 - 12/2018. The Institutional Review Board of Heidelberg University Medical School approved scientific evaluation of this retrospective data and the requirement for patient informed consent was waived (No. S-723/2017). The intraoperative computed tomography (CT) scans for screw planning were acquired in patients’ prone position, covering the desired spinal levels. Pedicle screws were manually planned by the operating surgeon in the acquired CT on a navigation workstation (*Stryker Spinemap 3D*). The planning process

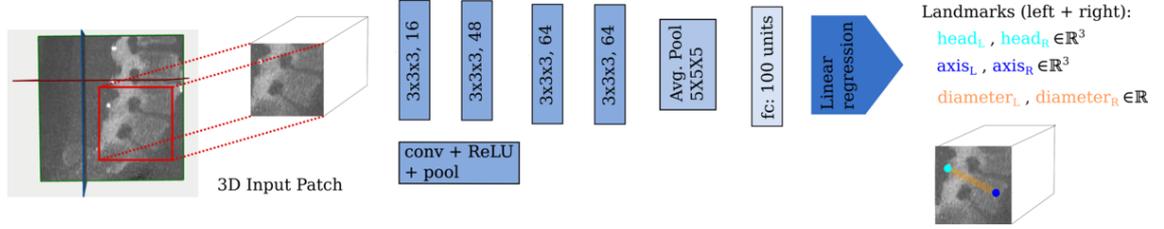


**Figure 4.2: Screw parameters and data preprocessing:** (a) Definition of screw parameters taken from Kausch et al. (2021a). (b) Preprocessing steps of screw planning data adapted from Kausch et al. (2021a).

involves manual interaction to position the 3D screw model at desired screw locations, correctly orientate the screw trajectory, and adjust screw length and dimension. The dataset considered in this thesis consists of 200 cases with 1380 screw plans created on the day of surgery that were randomly selected from the registry. The retrospective planning cohort consisted of spinal construct length of 1-5 vertebra levels. Screw parameters were extracted from the planning results and served as ground-truth reference (rater A). Screw parameters included: the screw *head* point, the screw *axis* point, the screw *length*, and the screw *diameter* as depicted in Fig. 4.2a.

The training dataset consisted of 155 cases (1052 screws) and the external test dataset of 45 cases (328 screws). A randomly selected subset of the test dataset (24 cases, 130 screws) was additionally annotated by a second specialized spinal neurosurgeon (rater B) who performed the screw planning twice at an interval of 8 weeks without considering previous planning results. This allowed for intra- and inter-rater variance assessment (rater B vs. rater B, rater A vs. rater B) of manual pedicle screw planning. Fig. 4.1 shows an overview of the training datasets, the test dataset (all), and the test dataset (inter-rater) along with the vertebra level-specific cohort distribution.

It has to be noted that the reference data was derived from surgical planning data and does not represent classic ground truth. The surgical planning data takes individual anatomical conditions into account, as well as the surgeon’s preference for screw placement.



**Figure 4.3: Instance-based screw planning based on direct CNN landmark regression** serves as baseline approach taken from Kausch et al. (2021a). A 3D vertebra patch serves as input for the CNN regression model. The screw *diameter*, *head* and *axis* point are predicted for the left and right screw.

**Preprocessing:** Since screws are only positioned in a selected subset of vertebrae, the screw planning task is tackled in a patch-based approach. Besides, screw planning was interpreted as a segmentation task. Therefore, the planned screw trajectories were converted to segmentation masks by performing a preprocessing consisting of the following steps (Fig. 4.2b): Given the screw parameters, a cylindrical mesh model was derived (A). The shape representation was modeled as cylindrical because it resembles the screw shape the best. The volume was resampled isotropically to balance the physical receptive field in terms of mm. Subsequently, binary masks were generated from the screw mesh models (B), and vertebra patches were extracted, cropped around the vertebra centroids (C). Thereby, the centroid  $c$  was computed from the screw parameters (Fig. 4.2a) by  $c = 1/2(\text{axis}_R + \text{head}_R + \text{axis}_L + \text{head}_L)$ . The lower indices  $R, L$  indicate the left and right screw parameters. For the training set, this resulted in 526 distinct vertebra patches covering vertebra levels L1-S1.

The following sections present the baseline approach for automatic pedicle screw planning based on direct CNN landmark regression (Sec. 4.1.2) and the proposed approach via an auxiliary segmentation task (Sec. 4.1.3).

### 4.1.2 Baseline approach

The baseline approach used for comparison is based on direct vertebra instance-based screw landmark regression with a CNN. The 3D CNN, shown in Fig. 4.3, is loosely adapted from the 2D network for facial landmark coordinate regression (Zhang et al., 2014). For training, vertebra-centered patches of size  $80 \times 80 \times 80$  voxels were created. The patch size was chosen large enough to contain a single vertebra of interest, and the random crops of  $72 \times 72 \times 72$  within these patches help to gain robustness towards varying vertebra patch initializations. The screw *head* and *axis* point were considered as coordinate

landmarks for the left and right screw, respectively.

**Training procedure:** The baseline approach was trained on the preprocessed vertebra patch training set resulting from surgical planning data (526 patches). The network was trained to predict the coordinate landmark locations relative to the patch center and the screw *diameters*, resulting in an 8-channel output. The other screw parameters, screw *length* and screw *direction* can be derived subsequently. One network configuration was trained across all lumbar vertebra levels (L1-S1).

The model was implemented using PyTorch v1.6.0 and trained with an 11 GB GeForce RTX 2080 Ti. The weight parameters were optimized using the adaptive moment estimation algorithm (Adam) optimizer (Kingma and Ba, 2014) with a base learning rate of  $\eta = 10^{-4}$  and batchsize of 64 until convergence. The baseline CNN regression model was trained in a 5-fold cross-validation on the training set.

**Objective function:** Let the screw parameters *head*, *axis*, *diameter* be abbreviated as  $h, a, d$ . Then, the model weights were optimized employing the L2 Loss function on the predicted and ground truth screw landmark coordinates  $\hat{h}, \hat{a}, h, a \in \mathbb{R}^3$  and screw diameters  $\hat{d}, d \in \mathbb{R}$  according to

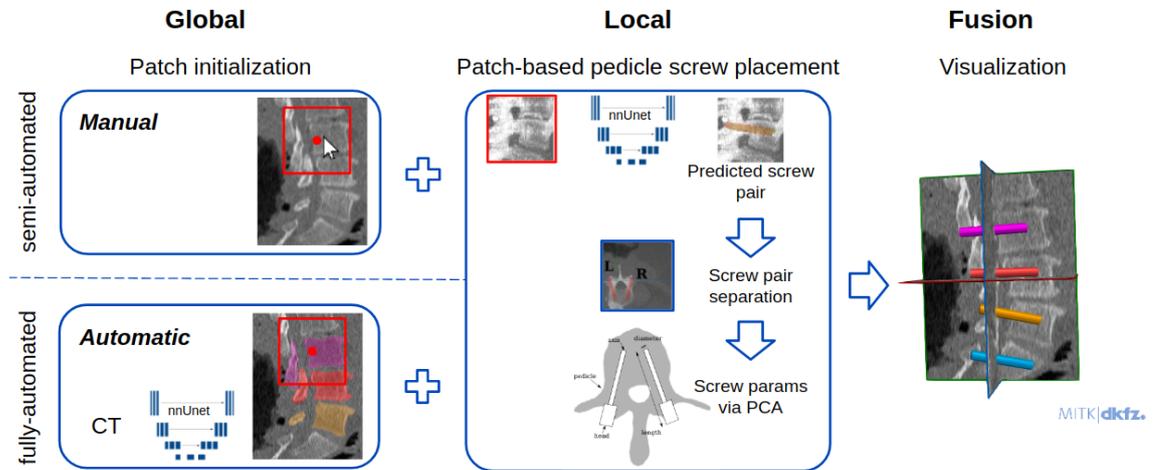
$$\mathcal{L}_{CNN}(\hat{h}, \hat{a}, \hat{d}, h, a, d) = \min_W \sum_{L,R} \left( \|\hat{h} - h\|^2 + \|\hat{a} - a\|^2 + \|\hat{d} - d\|^2 \right),$$

where  $R, L$  indicate the left and right screw parameters of one vertebra.

### 4.1.3 Proposed approach

In the proposed approach the pedicle screw planning was interpreted as a segmentation task. The complete workflow is shown in Fig. 4.4 and consists of the following three main steps:

**1) Patch initialization:** Vertebra-instance patch initialization can be performed *manual* or *automatic* resulting in a *semi-automated* or *fully-automated* screw planning pipeline. The *manual* patch initialization requires a user input in an interactive graphical user interface (*Medical Imaging Interaction Toolkit (MITK)*) as illustrated in Fig. 4.4, top left. The *automatic* patch initialization was based on a 3D U-Net (Isensee et al., 2021) trained on the large scale vertebrae segmentation challenge (VerSe) dataset for vertebra instance segmentation (Löffler et al., 2020) as illustrated in Fig. 4.4, bottom left. The vertebra centroids can be derived from the vertebra instance segmentation masks through connected component



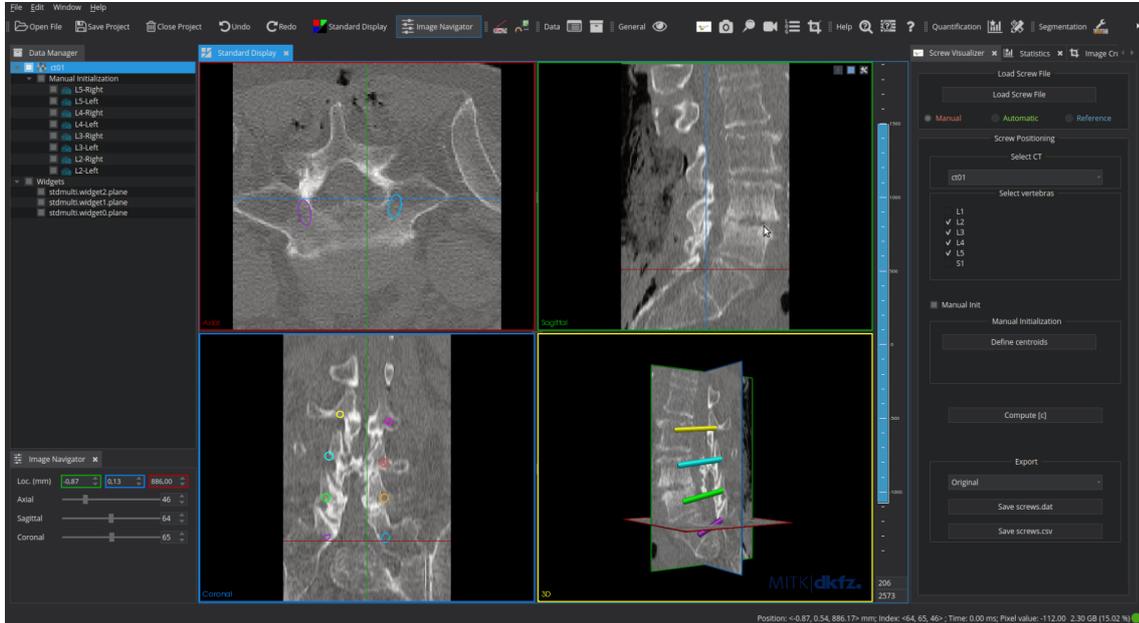
**Figure 4.4: Workflow for automatic pedicle screw placement** can be divided into vertebra instance patch initialization and instance-based pedicle screw planning. The pedicle screw planning consists of two steps: the instance-based screw pair segmentation and the derivation of screw parameters from the predicted screw masks.

bounding box computation. Given the automatically computed centroids and a set of selected vertebra levels, vertebra patches can be initialized without manual interaction. However, one limitation is inherited from the VerSe training dataset, where only vertebrae that are fully contained in the field-of-view (FoV) were labeled. In consequence, vertebrae close to the image boundary cannot be detected automatically.

**2) Patch-based pedicle screw placement:** Given a set of vertebra patches, it was hypothesized that a 3D U-Net could be exploited for pedicle screw placement though the 3D CT patches do not contain the actual screws. The network learns to predict a binary mask representing the position of pedicle screws based on the spatial context information of the input vertebra patch. The nnU-Net framework (Isensee et al., 2021) that automatically configures itself was employed for training. The instance-based screw segmentation was trained in a 5-fold cross-validation on the training set, similar to the baseline approach.

**3) Derivation of screw parameters from screw masks:** The derivation of screw parameters from the predicted screw masks is depicted in Fig. 4.4, bottom middle. At first, the screw pairs were separated into left and right single screw instances employing a connected component filter on the segmentation mask. The two largest connected components can be assigned to the left and right screw trajectory considering their coordinates. Subsequently, principal component analysis (PCA) was performed separately on the spa-

## 4.1. Pedicle screw planning for CT-navigated spinal surgery



**Figure 4.5: MITK Screw Planning Plugin for semi- and fully-automatic pedicle screw planning.** It offers a manual and automatic patch initialization. The computed screw plans are displayed in a medical viewer. Functionalities for screw parameter import and export are provided.

tial coordinates of the left and right screw instance. While the first eigenvalue represents the *length*  $l$  of the corresponding screw, the second eigenvalue determines the maximal *diameter*  $d$  of the predicted cylindrical mask. Finally, a least-square circle fit along the second principal component was performed to derive the screw diameter. The first principal component determines the screw direction. Consequently, with  $\mu$  indicating the mean of all coordinates, the screw *head*, and *axis* point can be computed by  $a = \mu + 1/2 \cdot l \cdot d$  and  $h = \mu - 1/2 \cdot l \cdot d$ .

### 4.1.4 Prototype implementation

The complete *semi-* and *fully-automated* pipeline for pedicle screw planning was developed and integrated as a plugin into the open-source framework MITK, which enables direct visualization of planning results in a medical viewer (Nolden et al., 2013) (Fig. 4.5). The developed *MITK Screw Planning Plugin* requires a selection of vertebra levels for screw planning and offers either *manual* or *automatic* patch initialization. Whereas the *manual* initialization requires that the user indicates the vertebra centroids in the interactive graphical user interface, the *automatic* initialization does not require any additional user

interaction. After initialization, the screw parameters are computed fully automatic by calling a python script in the background that initiates the deep learning prediction. After computation, the patch-based screw planning results are fused in the CT scan (Fig. 4.4, Fusion), and the predicted screws are displayed in the medical viewer. Additionally, the plugin offers functionalities for screw parameter loading and export. This allows for direct comparison of manually generated screw plans and automatically generated screw plans.

## 4.2 C-arm positioning for fluoroscopy-guided spinal surgery

### Disclosures to this work:

Part of this work has been published in the International Journal for Computer Assisted Radiology and Surgery (IJCARS) and in the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (Kausch et al., 2020, 2021b).

Guidance and quality control in orthopedic surgery increasingly relies on intraoperative fluoroscopy using a mobile C-arm. In this process, a key task of the surgeon involves the accurate acquisition of standardized and anatomy-specific projections. The corresponding manual positioning of the C-arm under repeated acquisitions or continuous fluoroscopy is error-prone, time-consuming, and strongly expertise-dependent. To reduce positioning time and radiation exposure for patients and clinical staff, standardize results, and avoid errors in fracture reduction or implant placement, in this thesis, computer-assisted methods for guiding – and in the long run automating – this procedure are developed.

The approaches developed in this thesis for automatic C-arm positioning are based on deep learning techniques. In contrast to the state of the art, the proposed methods do not require patient-individual prior knowledge like preoperative computed tomography (CT) scans or 2D-3D registration. Without requiring additional technical equipment, the necessary pose update for C-arm repositioning is directly predicted from the projection images themselves.

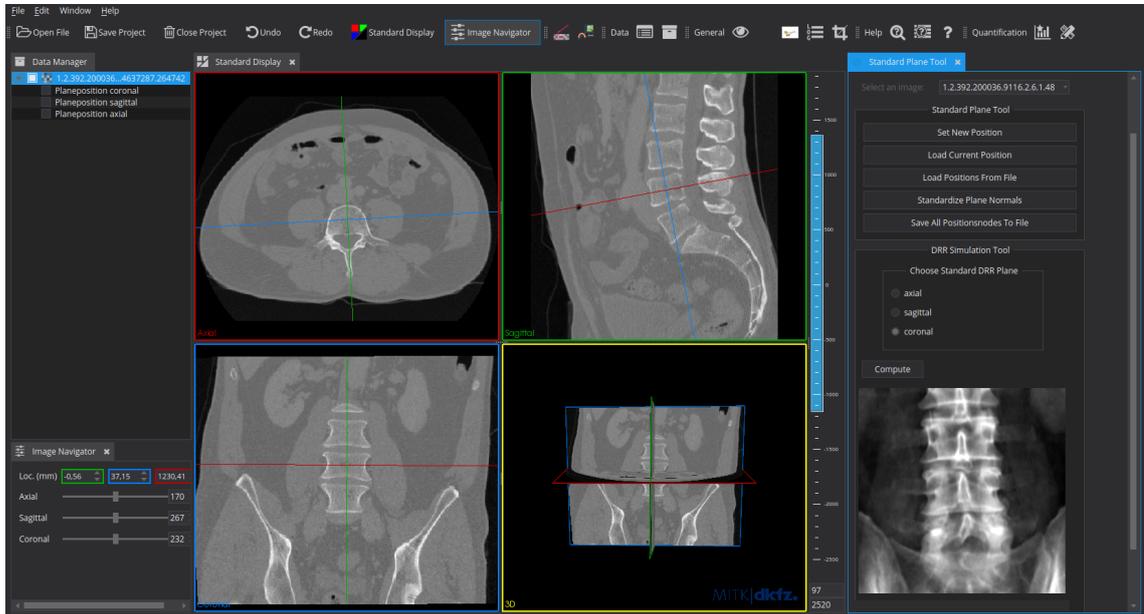
The analysis focuses on the spine's a.-p. and lateral standard projection because they are the most frequently acquired. The proposed methods are exemplarily evaluated for the fourth lumbar vertebra (L4). However, the methods are designed such that they are directly transferable to other standard projections, vertebra levels, and anatomical regions.

To tackle the lack of annotated intraoperative X-rays, a training data generation framework is introduced in Sec. 4.2.1. In Sec. 4.2.2 the specimens study for validation data acquisition is introduced. Sec. 4.2.3 derives different measures for manual C-arm positioning accuracy that can serve as a measure for acceptable automatic positioning deviations. In addition, the manual C-arm positioning efficiency was assessed. Subsequently, the baseline approach for automatic C-arm pose estimation and the proposed intensity-based and sequential C-arm pose regression approaches are presented (Sec. 4.2.4-4.2.6).

### 4.2.1 Training data generation

The mobile C-arm allows the acquisition of intraoperative images from arbitrary positions and orientations. The flexible positioning of the mobile C-arm relative to the patient implies the absence of a constant reference frame in interventional fluoroscopy. Consequently, X-rays with annotated pose labels do not exist in clinical routine. Thus, the necessary anatomical hints for efficient C-arm positioning were learned from *in silico* simulations leveraging retrospective CT scans. For training data generation, synthetic X-rays were derived from 47 full-body CTs acquired at different institutions and scanners. Prior to simulation, the CTs were cropped to an region-of-interest (ROI) around the L4 vertebra using the Medical Imaging Interaction Toolkit (MITK) *Image Cropper Plugin* which prevents other extremities from superimposing the simulated X-ray. Similarly, this is guaranteed in real intraoperative X-rays by the patient placement. For training and evaluating the following methods, the CTs were randomly divided into 60% training (28 CTs), 20% validation (10 CTs), 20% test (9 CTs), assuring that images of one patient are not shared between folds.

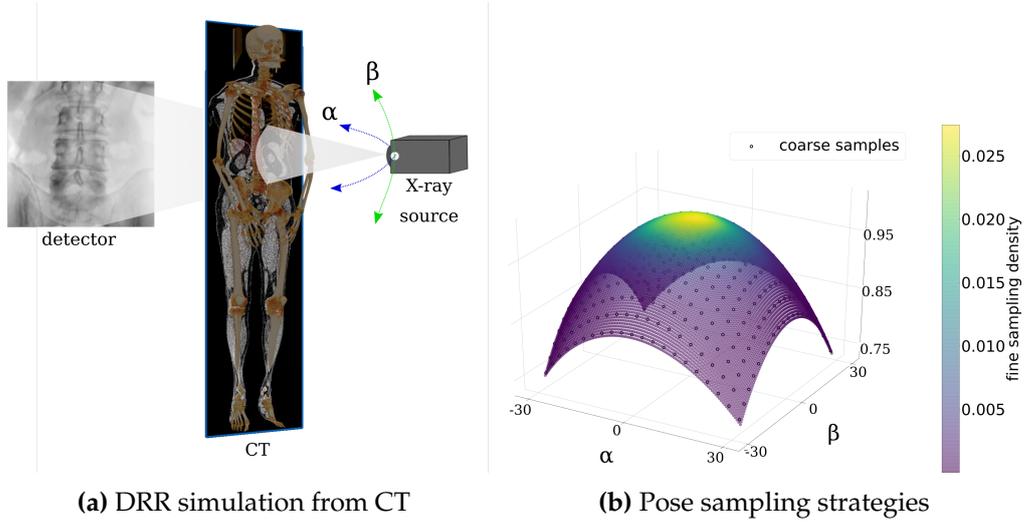
**Definition of ground truth reference:** To produce high-quality arbitrary view synthetic X-rays, a set of high-resolution CTs was provided by the university hospital Marburg. Independently, two clinical experts defined the two reference standard projection planes, corresponding to the a.-p. and lateral standard projection, for the L4 vertebra of the spine in each 3D volume. The 3D reference planes were defined using interactive plane translation and rotation of the interactive medical image viewer *MITK* (Nolden et al., 2013). The a.-p. standard was defined via the coronal plane and the lateral standard via the sagittal plane, respectively. The intersection point defines the center of the projection. It was placed in the centroid of the L4 vertebra body to assure that the vertebra of interest was centered in the synthetic X-ray. The reference plane orientations were defined, taking the requirements for a.-p. and lateral standard projection defined in Sec. 2.2.1 into account. To increase robustness of the defined reference standards, the planes were defined under the additional prior of pairwise plane orthogonality. The axial plane was aligned with



**Figure 4.6:** *MITK Standard Plane Plugin with DRR Preview.* The coronal plane defines the a.-p. standard and the sagittal plane defines the lateral standard. The plugin offers load and export functionalities for defined reference standard planes, as well as a corresponding digitally reconstructed radiograph (DRR) preview.

the vertebra endplates in the sagittal and coronal view. The sagittal plane was aligned with the symmetry axis of the vertebra in the axial view. For validation of the 3D-defined reference standard planes, a *DRR Preview Plugin* was developed and integrated into *MITK* (Fig. 4.6, left bottom). It allows to choose a reference standard plane and visualizes the corresponding DRR in real-time in a side preview. In addition, the plugin offers functionalities for loading and exporting standard plane parameters (Fig. 4.6, left top).

**Simulation framework:** To generate arbitrary view DRRs for training, the source and detector pose were varied around the reference C-arm pose (Fig. 2.2B), defined by the standard reference planes. The simulation framework *DeepDRR* was employed for a realistic forward projection of the 3D CT anatomy (Unberath et al., 2018b). In contrast to traditional ray-tracing algorithms, *DeepDRR* computes energy- and material-dependent attenuation images that are converted to synthetic X-rays. The choice of *DeepDRR* simulation was motivated by its generalization capabilities towards clinically acquired X-rays, as shown by Gao et al. (2019); Bier et al. (2018b). The C-arm system parameters were defined according to a Siemens Cios Spin<sup>®</sup> robotized C-arm. It has a  $300 \times 300 \text{ mm}^2$  detector at  $1952 \times 1952$  pixel resolution and a source-detector-distance of 1164 mm. DRRs are simulated for a

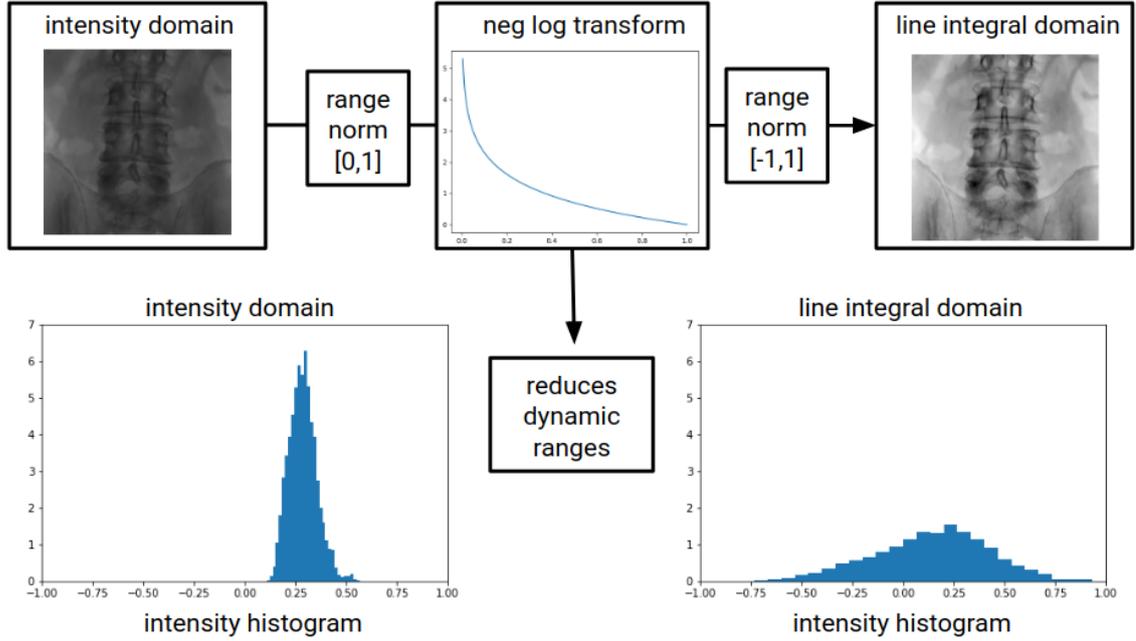


**Figure 4.7: DRR simulation framework.** Synthetic X-rays with ground truth reference poses are generated from CT scans 4.7a. Different sampling strategies (coarse, fine) around the reference beam direction are employed 4.7b. Figures taken from Kausch et al. (2020).

reduced resolution of  $512 \times 512$  pixels, which decreases the computation time. Ground truth pose labels were established by labeling DRRs according to their pose distance to the optimal view defined by the reference standard planes  $(d\alpha, d\beta, d\gamma, t)$ .

**Sampling strategies:** For training data generation, angular and orbital rotation were sampled from  $d\alpha, d\beta \in [-40^\circ, 40^\circ]$  in steps of  $3^\circ$  centered around the standard beam direction. The resultant uniform sampled training set, denoted by  $\mathcal{X}$ , has a size of 729 synthetic X-rays per CT volume. Validation was limited to a range of  $d\alpha, d\beta \in [-30^\circ, 30^\circ]$ . This assures that training examples equally cover the outer pose limits of the validation range as the inner poses. For training the proposed 2-step coarse to fine positioning approach, as introduced in Sec. 4.2.5, projections were sampled on two spherical segments: The coarse network was trained on the generated dataset with uniformly sampled pose offsets  $\mathcal{X}$ . The training set for subsequent fine positioning was simulated with pose parameters sampled from a Gaussian distribution so that the validation range was covered with 99% confidence. The resultant Gaussian sampled training set is denoted by  $\mathcal{X}'$ . The two different sampling strategies are visualized for comparison in Fig. 4.7b.

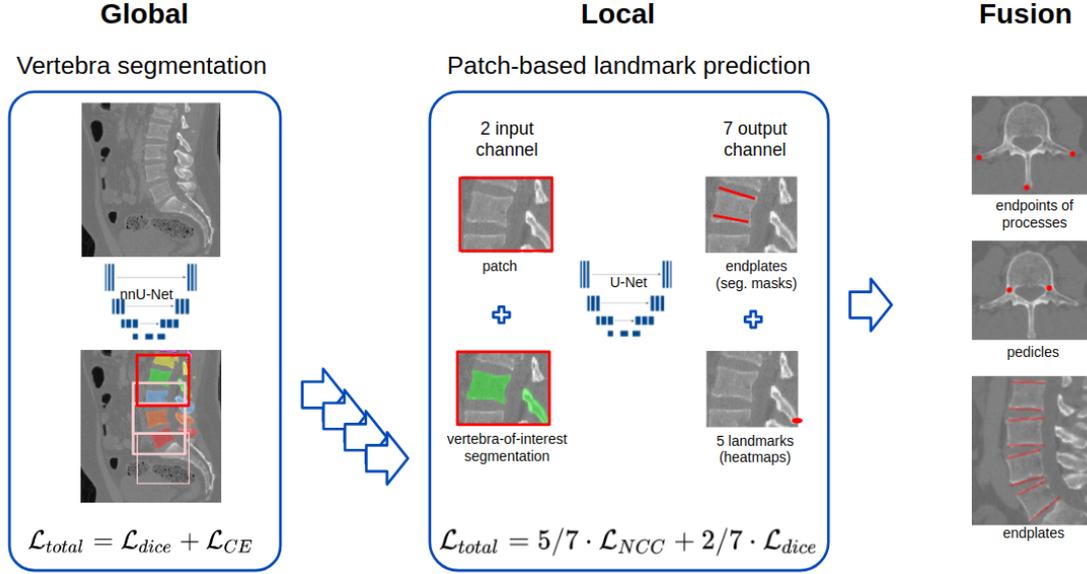
**Preprocessing:** The simulations' gray value intensity range was normalized to a range  $[0, 1]$  and converted from intensity to line integral domain by applying the negative Log transform (Fig. 4.8). The transformation decreases the dynamic range of the image



**Figure 4.8: Preprocessing of X-ray images.** The X-rays are converted from intensity to line integral domain to decrease the dynamic range.

intensities and increases the contrast for structures of interest, i.e. bone. Subsequently, the intensity range was normalized to  $[-1, 1]$ .

**Augmentation:** All proposed methods are trained with the same augmentation strategies introduced in the following to allow for cross-experiment comparison. The in-plane rotation  $\gamma$  was fixed during the simulation and was augmented online during training on the simulated dataset to limit the number of necessary simulations and hence computation time. At a reduced image resolution, detector rotation can be approximated by image rotation. For the uniform sampled training set  $\mathcal{X}$ , image rotation was augmented with rotation angles randomly sampled from a uniform distribution  $\gamma \in \mathcal{U}(-180, 180)$  and for the Gaussian sampled training set  $\mathcal{X}'$ , from a Gaussian distribution  $\gamma \in \mathcal{N}(\mu = 0, \sigma^2 = (30/2.576)^2)$  respectively. Modeling the in-plane rotation  $\gamma$  in addition to the out-of-plane rotations  $\alpha, \beta$  allows representing variable initial C-arm to patient orientations, that depends on the angle the C-arm is moved towards the patient table. Translation in the detector plane was simulated by augmenting random crop transforms with a translation of the image center  $t \in \mathcal{U}(-50 \text{ mm}, 50 \text{ mm})^2$  at a resolution of  $256 \times 256$  pixels. Simultaneously, the random crop eliminates order interpolation effects resulting from the image rotation. The labeled training set is given by  $\mathcal{S} = \{(\mathbf{x}_1, \mathbf{p}_1), \dots, (\mathbf{x}_N, \mathbf{p}_N)\}$



**Figure 4.9: Generation of ground truth 3D vertebra segmentations and landmarks** adapted from Kausch et al. (2021b). The selected landmarks relevant for standard projections are shown on the left (endpoints of processes, pedicles, and vertebra endplates). The global vertebra segmentation stage automatically locates the vertebrae for vertebra patch initialization. The local stage operates on the derived patches and outputs a 7 output channel feature map which contains 2 segmentation masks indicating the vertebra endplate locations and 5 heatmaps corresponding to the 5 point landmarks.

with  $\mathbf{p}_i = [\alpha_i, \beta_i, \gamma_i, t_x, t_y] \in \mathbb{R}^5$  and synthetic X-rays  $\mathbf{x}_i$ . Additional online augmentations were applied during training to increase the variety of training data: Image scaling with  $s \in \mathcal{U}(0.8, 1)$  was used, which approximates variation of the source-to-isocenter distance or bone sizes. Larger scale coefficients, which resemble moving the source closer to the anatomy, were excluded. This results in a smaller field-of-view (FoV) and a higher patient dose, which is not clinical practice. Moreover, contrast augmentation was performed by modifying each pixel  $p$  according to  $p = (p - \mu) * c + \mu$  for image-wide  $c \in \mathcal{U}(0.5, 1.5)$ , where  $\mu$  denotes the mean of the image. Random additive Gaussian noise, Gaussian blur with  $\sigma \in \mathcal{U}(1, 2)$ , and non-linear gamma transform  $p = p^\gamma$  with  $\gamma \in \mathcal{U}(0.5, 2)$  were applied (Isensee et al., 2020). All augmentations were applied with a probability 0.5.

**Generation of ground truth 3D segmentation and landmarks:** The simulated images were complemented with automatically generated segmentations and landmarks to mimic approaches used by spinal neurosurgeons in identifying correct standard projections. Annotating the landmarks in 3D and establishing 2D annotations by projection implies

2D consistency and reduces the labeling effort compared to manual 2D annotations. To automate the 3D annotation process, two networks were trained (Fig. 4.9): one for the task of vertebra segmentation (Fig. 4.9, Global), the other for the task of landmark detection (Fig. 4.9, Local).

**3D vertebra segmentation:** For 3D vertebra segmentation of CT volumes, a state of the art U-Net segmentation framework (Isensee et al., 2021) was employed that was trained on the large scale vertebrae segmentation challenge (VerSe) dataset for 3D vertebra instance segmentation (Löffler et al., 2020). The model was applied on all CT volumes, resulting in 3D multi-label segmentation masks.

**3D landmark detection:** Clinically meaningful landmarks to identify correct standard projections were selected in consultation with a spinal neurosurgeon: (1) endpoints of transverse and spinous processes, (2) pedicles, (3) vertebra endplates (Fig. 4.9, Fusion). The selected vertebra landmarks were manually annotated on a subset of 10 CTs chosen randomly from the training set. As shown in Fig. 4.9 on the right, the landmarks were defined for vertebra levels second lumbar vertebra (L2)-fifth lumbar vertebra (L5), taking also neighboring vertebrae of the considered vertebra level of interest L4 into account. For annotations, those vertebrae were selected that are usually contained in the FoV considering the L4 standard projection.

To automatically derive the landmark annotations for the remaining CTs (37), an automatic patch-based approach for landmark prediction was trained on the 40 annotated vertebra patches resulting from the ten annotated CTs. The complete pipeline is shown in Fig 4.9.

**Training procedure:** The 3D vertebra instance segmentation was employed to derive local vertebra patches from the 3D CTs. Local square patches of pixel size 128 were cropped around the vertebra centroid of the bounding box defined by the instance-based vertebra segmentation. The crop size was selected as small as possible to exclude neighboring vertebrae but large enough to contain the whole vertebra. The vertebra patch was complemented with its vertebra-of-interest segmentation and used as 2-channel input to train a 3D U-Net (Klein et al., 2019) for patch-based landmark prediction. The model predicts a 7-channel output where the vertebra endplates were represented as segmentation masks and the point landmarks were represented as heatmaps. The heatmaps were defined as symmetric 2D normal distribution with a mean equal to the spatial landmark location and standard deviation  $\sigma = 3 \text{ mm}$  in each direction. The value of  $\sigma$  was subjectively selected to represent the variation of manual landmark annotation. Subsequently, the heatmaps

---

## 4.2. C-arm positioning for fluoroscopy-guided spinal surgery

were normalized to the range  $[0, 1]$ .

The models were implemented using PyTorch v1.6.0 and trained with an 11 GB GeForce RTX 2080 Ti. For experiment organization the *trixi* package was used (Zimmerer et al., 2017). The networks were optimized using the adaptive moment estimation algorithm (Adam) optimizer (Kingma and Ba, 2014) with a base learning rate of  $\eta = 0.0002$  and batchsize of 16 until convergence.

The training procedure is described in detail in the following:

**Objective function:** Let

$$p : x \in \Omega \subset \mathbb{R}^3 \mapsto \mathbb{R}$$

be the CT volume patch and

$$\begin{aligned} \hat{y} : x \in \Omega \subset \mathbb{R}^3 &\mapsto [0, 1]^7 \\ y : x \in \Omega \subset \mathbb{R}^3 &\mapsto \mathbb{R}^7 \end{aligned}$$

be the ground truth annotation and the network output, respectively. The network output  $y$  can be derived from the model architecture with mapping function  $f$  by

$$f(\cdot, \cdot) : p \times W \mapsto y,$$

given the network weight configuration  $W$ .

Then weight parameters  $W$  of the model were optimized jointly for the segmentation and landmark detection task by solving the objective function

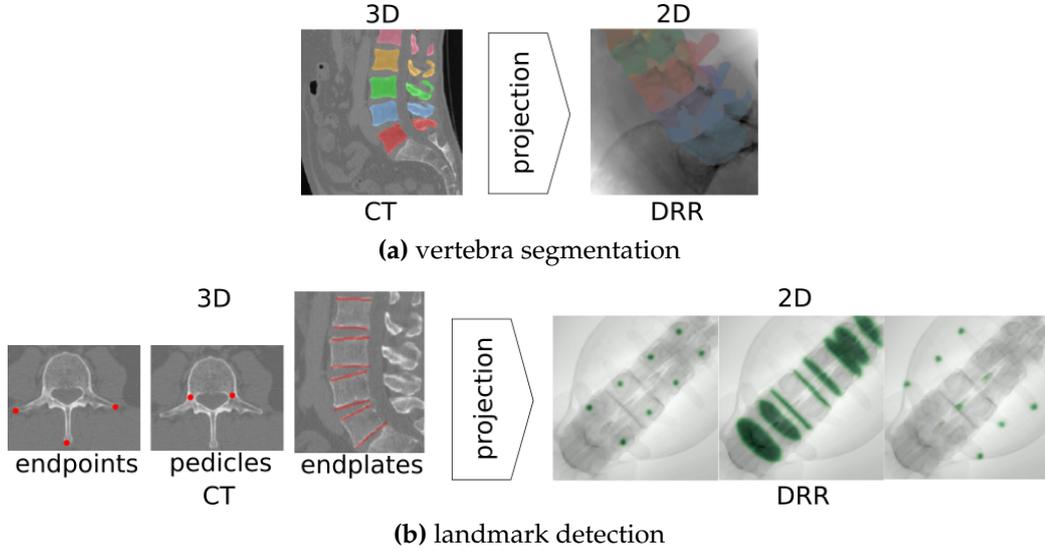
$$\mathcal{L}_{total}(\hat{y}, y) = \min_W \frac{5}{7} \mathcal{L}_{NCC}(\hat{y}, y) + \frac{2}{7} \mathcal{L}_{Dice}(\hat{y}, y).$$

For the segmentation task, the binary Dice Loss was employed for optimizing the network weights such that the criterion

$$\mathcal{L}_{Dice}(\hat{y}, y) = -2 \left( \sum_x \hat{y}(x) \sigma(y(x)) \right) / \left( \sum_x y(x) + \sum_x \sigma(y(x)) \right) \in [-1, 0]$$

was maximized with sigmoid activation function

$$\sigma(\cdot) : y \mapsto \exp(y) / (1 + \exp(y)) \in [0, 1].$$



**Figure 4.10: Generation of ground truth 2D annotations** for arbitrary poses. The projection of automatically generated 3D annotations from CTs assures consistent annotations and reduces the labelling effort.

For the heatmap task, the averaged normalized cross-correlation (NCC) was employed for optimizing the network weights such that the criterion

$$\mathcal{L}_{NCC}(\hat{y}, y) = -\frac{1}{N} \cdot \frac{\sum_x (y(x) - \text{mean}(y)) \cdot (\hat{y}(x) - \text{mean}(\hat{y}))}{\left(\sqrt{\text{var}(y) - \text{var}(\hat{y})}\right)} \in [-1, 1]$$

was maximized with  $N$  denoting the number of voxels per volume. For simultaneous optimization of segmentation masks and heatmaps, the NCC Loss was rescaled and shifted to  $[-1, 0]$  (Grupp et al., 2020).

**Augmentation:** To regularize training and decrease overfitting, the training data was augmented with different image transformations using the *batchgenerators* framework (Isensee et al., 2020). Augmentations included rotation  $\alpha_{rot} \in \mathcal{U}(-5^\circ, 5^\circ)$ , scaling  $s \in \mathcal{U}(0.8, 1.2)$ , translation  $t \in \mathcal{U}(0\text{mm}, 22\text{mm})$ , elastic transform  $(\alpha, \sigma) \in (\mathcal{U}(0, 10000), \mathcal{U}(40, 60))$ , contrast augmentation  $c \in \mathcal{U}(0.75, 1.5)$ , and subsequent range normalization to  $[-1, 1]$ . All augmentation were applied with probability 1.

The automatically generated 3D annotations were inspected visually for validation.

**Algorithm 1:** K-wire augmentation

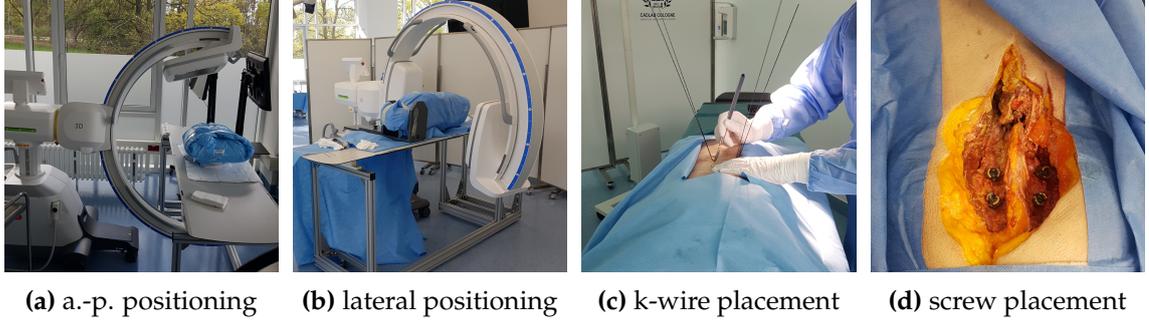
---

**Input** :  $x$  (input image, range  $[-1,1]$ ),  $y$  (spine or pedicle mask)  
Sample number of k-wire:  $k = \text{randint}(0,7)$ ;  
 $N := \text{batchsize}$ ;  
**for**  $i \leftarrow 0$  **to**  $k$  **do**  
    Sample contrast augmentation:  $k \leftarrow \text{random.uniform}(-0.25,0)$ ;  
    Sample length of k-wire:  $l \leftarrow \text{randint}(0.2 * n, 0.6n)$ ;  
    Sample direction of k-wire:  $\alpha \leftarrow \text{randint}(-180,180)$ ;  
    Compute direction:  $d \leftarrow [\cos \alpha, \sin \alpha]$ ;  
    Sample start control point:  $a$  uniformly sampled from inside  $y$ ;  
    Compute end control point:  $c \leftarrow a + l \cdot d$ ;  
    Compute and distort middle control point:  $b \leftarrow$   
     $a + \text{rand}(0.01,0.99) * l * d + [\text{randint}(-0.5 \cdot l, 0.5 \cdot l), \text{randint}(-0.5 \cdot l, 0.5 \cdot l)]$ ;  
    Render  $k - \text{th}$  bezier curve with control points  $a, b, c$ ; Replace non-zero  
    values with samples from  $\text{random.normal}(-0.85,0.05)$ ;  
    Image blending with contrast augmentation:  $x_{aug} \leftarrow x + r + k$   
**end**  
**Output:**  $x_{aug}$

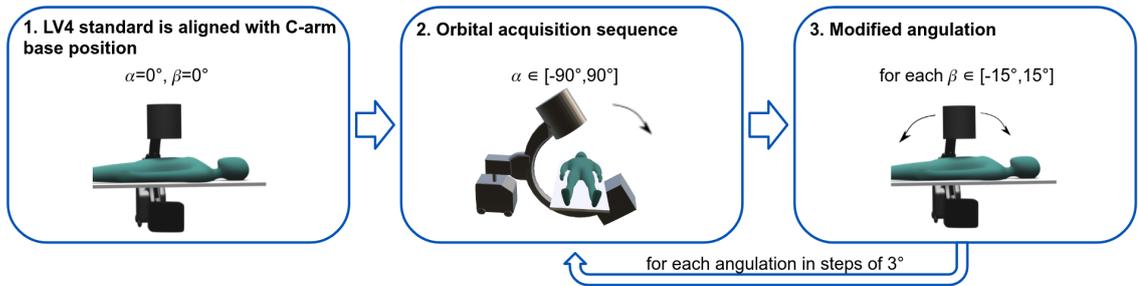
---

**Generation of ground truth 2D segmentations and landmarks:** The generated 3D annotations were projected for varying directions, thereby establishing 2D ground truth vertebra segmentations and landmark locations (Fig.4.17(C), (D)). In the projection domain, endpoints, pedicles and plates were represented in distinct image channels (Fig. 4.10). Landmarks were represented as heatmaps by convolution with a Gaussian kernel with  $\sigma = 3$ . Every pixel value represents the confidence proportion to its proximity to the exact landmark location. This reflects uncertainty, since 3D manual annotation can differ between several experts.

**K-wire and screw augmentation:** To address the domain gap between training DRRs and intraoperative X-rays with present confounding factors like metal implants, synthetic X-rays were additionally complemented with k-wire and screw simulation. K-wires were modeled with quadratic Bézier curves. The pseudocode of the k-wire simulation method is given in Alg. 1 and was integrated as an additional augmentation strategy into the training pipeline. For screw simulation, a realistic screw modeling pipeline was developed, consisting of the following steps: (1) Screw trajectories were automatically planned in CT volumes for vertebra levels L2-first sacral vertebra (S1) using the automatic pedicle screw planning prototype developed in Sec. 4.1.3 (Kausch et al., 2021a). (2) Pedicle screw computer-aided design (CAD) models (Marathakis, 2013) were aligned with the derived



**Figure 4.11: Preparation for specimens study.** The patient positioning for validation data acquisition differs for both standard projection to enable the automatic derivation of ground truth pose labels (a), (b). The placement of surgical implants is exemplary shown (c), (d).



**Figure 4.12: Workflow for validation data acquisition in specimens study** exemplary for the a.-p. standard.

screw parameters. (3) CAD models were converted to screw segmentation masks. Derived masks were projected in the detector plane using the *DeepDRR* framework (Unberath et al., 2019). The *DeepDRR* framework was extended to allow for an arbitrary number of tools. During simulation, the number of implants was selected randomly in the interval  $[1, 10]$ . Synthetic DRRs were complemented with k-wire and screw simulations at probability  $1/3$  for each case (no metal, with k-wire, with screws).

#### 4.2.2 Acquisition of validation data

To validate the proposed pose regression approaches and implant simulation strategies on real X-rays, a specimen study was performed. The ethics committee Rheinland-Pfalz approved the specimen study (No. 2020-15423). To simulate different stages of a real clinical procedure, k-wires and screws were inserted in selected pedicles of torso specimens (Fig. 4.11c, 4.11d). For acquisition, a Siemens Cios Spin<sup>®</sup> mobile C-arm was employed. The

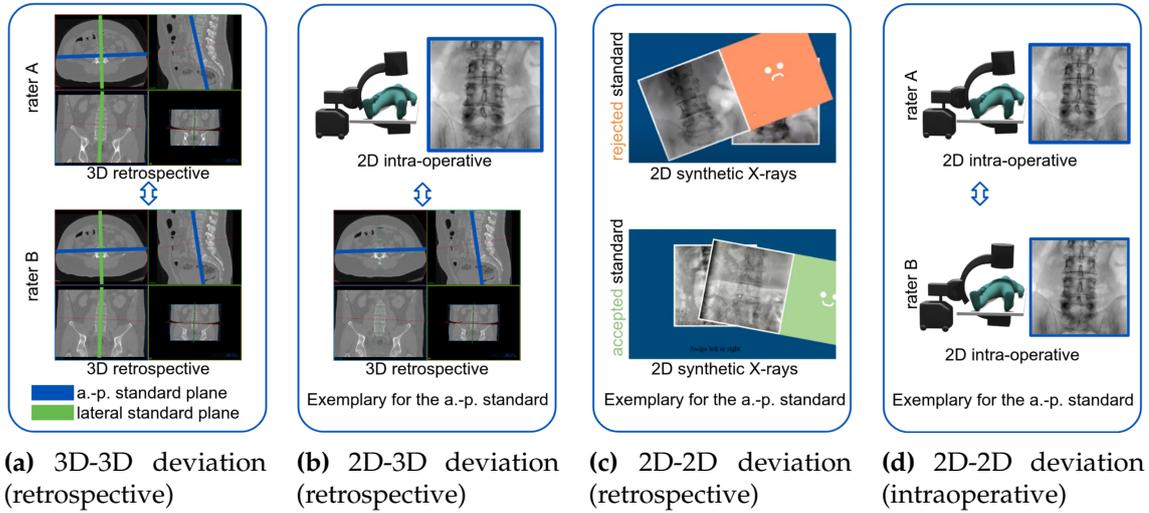
study workflow is depicted in Fig. 4.12. The torso specimens were initially aligned with the C-arm such that the C-arm's isocenter aligns with L4 and its base position at  $\alpha = 0^\circ$ ,  $\beta = 0^\circ$  corresponds to the standard pose (a.-p., lateral, respectively). For the anterior-posterior (a.-p.) standard the torso lies in prone position (Fig. 4.11a) and for the lateral standard a lateral positioning in which the torso was placed sideways on the operating room (OR) table was required (Fig. 4.11b). The described initial positioning enables the establishment of ground truth poses for the acquired datasets directly from the internal recorded C-arm poses. The Cios Spin<sup>®</sup> stores the  $\alpha/\beta$  angles in the metadata of every single acquisition which served as the ground truth labels after correction (Sec. 5.2.2). A dataset consists of orbital projection sequences acquired at different angulations ( $\beta \in [-15^\circ, 15^\circ]$ ) in steps of  $3^\circ$  and corresponding 3D reconstructions. Per specimen and standard projection, this implies 11 3D scans for the different angulations. Projections resulting from poses not covered by the scope of the algorithms were neglected during validation. Specifically, all projections with  $\alpha \in [-30^\circ, 30^\circ]$  were sampled from the orbital sequence and were considered for validation. The resultant validation set contains 1364 projection images per specimen covering a pose range of  $\alpha \in [-30^\circ, 30^\circ]$ ,  $\beta \in [-15^\circ, 15^\circ]$ . In total, 16 datasets were acquired: 7 specimens without metal, 5 specimens with k-wires, 4 specimens with screws.

### 4.2.3 Assessment of manual C-arm positioning accuracy and efficiency

So far, no study has been performed to assess the variability of standard projections after manual C-arm positioning. The inter- and intra-rater variability are difficult to assess due to two main reasons: (1) fluoroscopic imaging of patients requires clinical indication since it involves harmful X-ray radiation and thus a variance study in clinical routine is not feasible, (2) a constant reference frame does not exist in interventional fluoroscopy which hinders retrospective evaluation. To overcome this limitation, an *inter- and intra-rater specimens study* was performed in the scope of this thesis to assess the accuracy and efficiency of manual C-arm positioning for standard projections of the spine.

Four different accuracy measures for standard projections were assessed in this thesis. An overview is given in Fig. 4.13. They are introduced in the following regarding their definition, purpose, and interpretation:

**(a) 3D-3D deviation (retrospective):** The inter-rater accuracy of the reference standard planes, retrospectively defined in the CT volumes independently by two experts, was assessed. For evaluation, all CT volumes were considered. Specifically, the variance of the coronal plane orientation was assessed for the a.-p. standard, and the variance of the



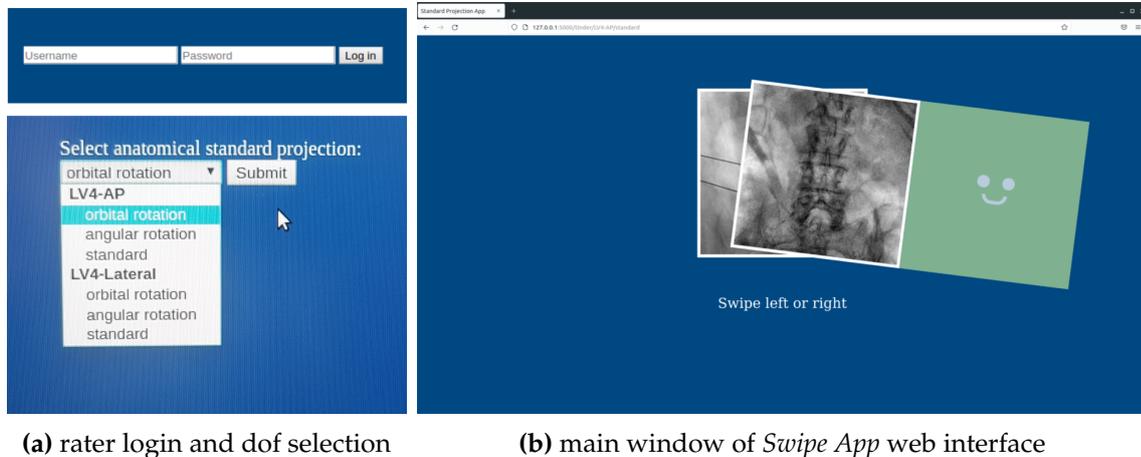
**Figure 4.13: Accuracy measures for standard projections.**

sagittal plane was assessed for the lateral standard. The measure reflects the accuracy of the training data and consequently forms an upper limit for the reachable accuracy of the trained algorithms. Standard planes were defined in the 3D volume under the additional prior of pairwise plane orthogonality. Thus, the inter-rater variances are by no means representative of the precision achieved in real clinical scenarios, where the variance is much higher due to the missing depth information. Results are reported in Sec. 5.2.1.

**(b) 2D-3D deviation (retrospective):** The variability of an intraoperatively acquired standard projection and a retrospectively defined standard plane in CT shows how much the pose varies in the respective standard projection. Assessment of this variance requires the availability of intraoperative standard projections and corresponding 3D reconstructions with a known spatial relationship. For evaluation, the acquired validation data (Sec. 4.2.2) was considered. Therefore, the standard planes (coronal for a.-p. standard, sagittal for lateral standard) were defined in all reconstructed 3D volumes, and the deviation to the intraoperative standard poses after manual C-arm positioning was computed. Results are summarized in Sec. 5.2.2.

**(c) 2D-2D deviation (retrospective):** To derive a measure for the variability of the standard projections, a standard plane clinical accuracy tool, called *Swipe App*, was developed, inspired by Bosco (2015) (Fig. 4.14). The *Swipe App* was implemented as a web application using Flask. It randomly presents X-ray projections that were labeled with the pose offset to the reference standard. The tool allows swiping left or right to accept or reject a

## 4.2. C-arm positioning for fluoroscopy-guided spinal surgery



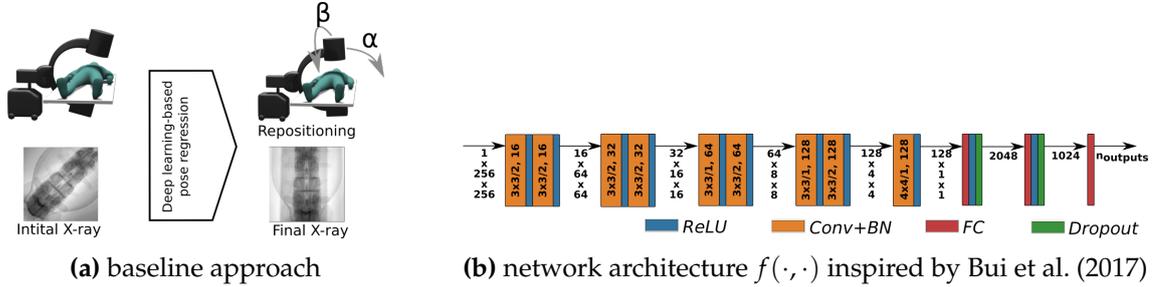
**Figure 4.14:** *Swipe App*: standard plane clinical accuracy web tool. After login and dof selection ( $\alpha$ ,  $\beta$ ), X-rays with known pose offset to a desired standard projection were presented randomly. The expert can accept or reject a presented X-ray as standard projection by swiping right or left. The orbital and angular dofs were assessed independently to simplify the decision space.

presented projection as standard projection. The ground truth angulation of all accepted projections serves as a baseline for the acceptable errors in clinical practice. Results are presented in Sec. 5.2.3.

**(d) 2D-2D deviation (intraoperative):** To assess the manual C-arm positioning accuracy, an inter- and intra-rater specimen study was performed. Therefore, two experts independently positioned the C-arm manually for the a.-p. and lateral standard, respectively. The standard projections were acquired for two specimens for the vertebra levels first thoracic vertebra (Th1)-L5. Additionally, a 3D scan was acquired corresponding to each acquired standard projection. Two standard projections of different experts can be assessed with respect to their pose agreement by registering the related 3D volumes. Results are reported in Sec. 5.2.3.

**Manual C-arm positioning efficiency:** In addition to the manual C-arm positioning accuracy, also the efficiency of manual C-arm positioning for standard projections was assessed. Therefore, for each expert, the time, the dose, and the number of X-rays required for iterative positioning were recorded separately for the two standard projections. Results are reported in Sec. 5.2.3.

In the following, the deep learning-based methods for automatic C-arm positioning



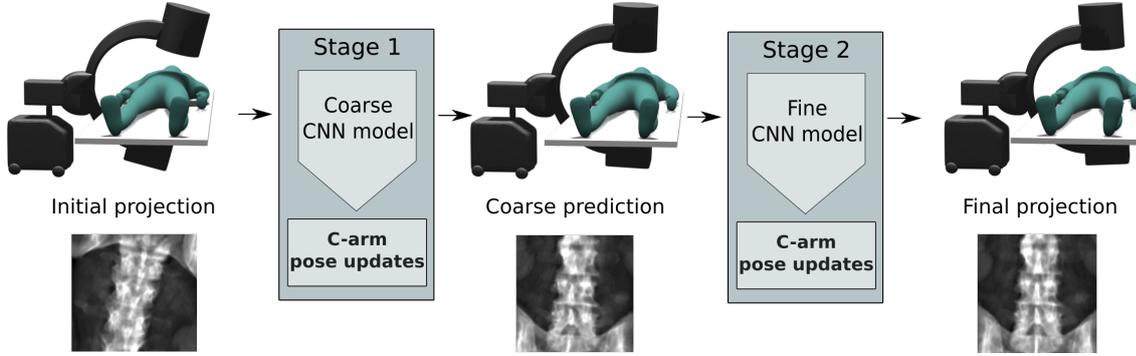
**Figure 4.15: Baseline approach for automatic C-arm positioning** for standard projections. Network architecture taken from Kausch et al. (2020).

developed in the scope of the thesis are introduced.

#### 4.2.4 Baseline approach

The baseline approach is based on a convolutional neural network (CNN) regression model that predicts five degrees-of-freedom (dof) pose updates directly from a first X-ray image. The setup and the model architecture are shown in Fig. 4.15. The network architecture is a modified version of PoseNet (Bui et al., 2017) with larger input size, batch normalization after each convolutional layer, and additional rectified linear unit (ReLU) activation after each 2nd convolutional layer, which showed to improve the convergence during training. The five dof represent the positioning dof of the C-arm, namely the orbital rotation  $\alpha$ , the angular rotation  $\beta$ , the image rotation in the detector plane  $\gamma$  and the translation parallel to the detector plane  $t_x, t_y$  (Fig. 2.1a). The translation along the beam direction was omitted because this translation only influences the image scale of the projection and thus does not alter the quality of standard projection. Since the scale also depends on the anatomy's size, this parameter cannot be optimized without a 3D reference model. The underlying modeling assumes that the structure of interest was initially located approximately midway between source and detector for the initial projection.

**Training procedure:** The baseline approach was trained on the synthetic training dataset  $\mathcal{X}$  generated from retrospective CT scans and uniform pose sampling. The DRRs were complemented with k-wire and screw simulations and augmentation strategies, as presented in Sec. 4.2.1. For a given initial X-ray, the model estimates the necessary pose update  $\hat{\mathbf{p}} = [\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{t}_x, \hat{t}_y]$  to acquire a desired standard projection. For each standard projection (a.-p., lateral), individual network weights were derived.



**Figure 4.16: 2-step intensity-based C-arm pose regression** taken from Kausch et al. (2020). In subsequent stages, the 5-dof relative pose update to reposition the C-arm for a desired standard projection was estimated, directly from an initial X-ray.

The model was implemented using PyTorch v1.6.0 and trained using an 11 GB GeForce RTX 2080 Ti. The weight parameters were optimized using the Adam optimizer (Kingma and Ba, 2014) with a base learning rate of  $\eta = 10^{-4}$  and batchsize of 32 until convergence.

**Objective function:** The model weights were optimized employing the L2 Loss function on the predicted and ground truth pose parameters  $\hat{\mathbf{p}} \in \mathbb{R}^6$ ,  $\mathbf{p} \in \mathbb{R}^5$ . Simultaneously, the loss was computed for the out-of-plane rotation  $\alpha$ ,  $\beta$ , the in-plane rotation  $\gamma$  and the detector translation  $t_x$ ,  $t_y$  by

$$\begin{aligned} \mathcal{L}_{pose}(\hat{\mathbf{p}}, \mathbf{p}) &= \min_W \mathcal{L}_{out-of-plane}(\hat{\mathbf{p}}, \mathbf{p}) + \mathcal{L}_{in-plane}(\hat{\mathbf{p}}, \mathbf{p}) + w \mathcal{L}_{detector-t}(\hat{\mathbf{p}}, \mathbf{p}) \\ &= \min_W \|\hat{\alpha} - \alpha\|^2 + \|\hat{\beta} - \beta\|^2 + \|\hat{c}_\gamma - \cos \gamma\|^2 + \|\hat{s}_\gamma - \sin \gamma\|^2 + \\ &\quad w^2 (\|\hat{d}_x - d_x\|^2 + \|\hat{d}_y - d_y\|^2), \end{aligned} \quad (4.1)$$

where  $w = \pi(180)^{-1}$  is a weighting factor to equally penalize orientation ( $^\circ$ ) and translation (mm) error. The in-plane rotation  $\gamma$  was converted to sin/cos-space and two values were regressed, which results in a continuous loss function, which can handle the circularity. This results in one additional output neuron. The parameters  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{c}_\gamma$ ,  $\hat{s}_\gamma$ ,  $\hat{d}_x$ ,  $\hat{d}_y$  were predicted by the pose net. The in-plane rotation can be derived by

$$\hat{\gamma} = \text{atan}_2 \left( \hat{s}_\gamma (\hat{s}_\gamma^2 + \hat{c}_\gamma^2)^{-\frac{1}{2}}, \hat{c}_\gamma (\hat{s}_\gamma^2 + \hat{c}_\gamma^2)^{-\frac{1}{2}} \right).$$

### 4.2.5 2-step intensity-based C-arm pose regression

Based on the baseline approach, a 2-step coarse and fine positioning approach (Fig. 4.16) was developed, which iteratively adjusts the C-arm pose to the desired standard projection. Both stages share the same network architecture shown in Fig. 4.15b. The first stage regressor covers a broader capture range, which makes the approach more robust toward higher initial deviations. It complies to the task of coarse alignment, while the second stage fine-tunes the prediction and thereby improves the accuracy. In clinical practice, this would reduce the iterative radiation to only two necessary X-rays.

**Training procedure:** The networks were trained separately on differently sampled synthetic training sets. The first stage resembled the baseline approach and was trained accordingly. The second stage for fine positioning was trained on the Gaussian sampled training set  $\mathcal{X}'$ . It was optimized accordingly, with the difference being that the  $\gamma$ -rotation was estimated directly and that a weighted L1 Loss was employed, which penalizes pose errors linearly. One individual network was trained for each of the anatomy-specific standard projections. The implementation and optimizer strategies were the same as for the baseline approach.

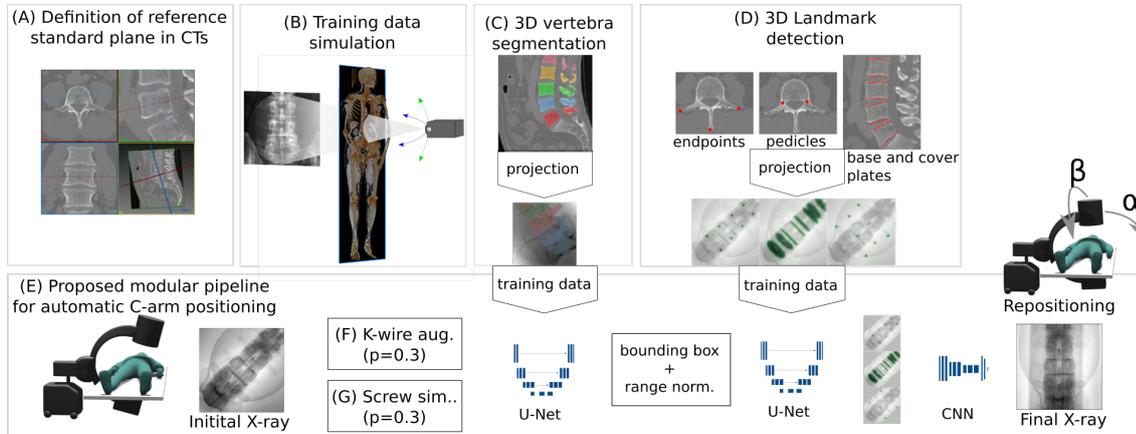
**Objective function:** Given the 2-stage predictions  $\hat{\mathbf{p}}^{(1)} \in \mathbb{R}^6$ ,  $\hat{\mathbf{p}}^{(2)} \in \mathbb{R}^5$  the model weights were optimized by minimizing the objective function

$$\begin{aligned}\mathcal{L}_{2\text{-step}}(\hat{\mathbf{p}}^{(1)}, \hat{\mathbf{p}}^{(2)}, \mathbf{p}) &= \min_W \mathcal{L}_{\text{coarse}}(\hat{\mathbf{p}}^{(1)}, \mathbf{p}) + \min_{W'} \mathcal{L}_{\text{fine}}(\hat{\mathbf{p}}^{(2)}, \mathbf{p}) \\ &= \min_W \mathcal{L}_{\text{pose}}(\hat{\mathbf{p}}^{(1)}, \mathbf{p}) + \min_{W'} \|\hat{\mathbf{p}}^{(2)} - \mathbf{p}\|_1.\end{aligned}$$

### 4.2.6 1-step sequential C-arm pose regression

Instead of direct intensity-based pose regression, a 1-step sequential automatic C-arm pose regression approach was proposed (Fig. 4.17). Building upon ideas of Bier et al. (2019); Grupp et al. (2020), domain knowledge was leveraged to guide robust decision-making. Specifically, the pipeline was designed to reflect and automate the decision approaches used by spinal neurosurgeons for identifying correct standard projections. It explicitly incorporates steps like ROI localization, detection of relevant view-independent landmarks, and subsequent pose regression. The design enables a successful transfer from simulated to real X-rays and also increases the interpretability of results. In developing the proposed pipeline, good surgical practice was followed:

## 4.2. C-arm positioning for fluoroscopy-guided spinal surgery



**Figure 4.17: 1-step pose regression pipeline** taken from Kausch et al. (2021b). The top row summarizes the steps for training data generation, which are connected to specific pipeline steps in the bottom row.

**ROI localization module:** Surgeons adjust the standard projection only based on the bone structures. This was reflected by constraining the input only to the bone region, employing a vertebra segmentation mask. Therefore, a U-Net model (Klein et al., 2019) for view-independent vertebra segmentation of spinal X-rays was trained on synthetic data. The 2D segmentations were derived from automatically generated 3D segmentations as described in Sec. 4.2.1. The model weights were optimized using a combination of binary cross-entropy and Dice Loss (Isensee et al., 2021). A rotating bounding box was computed from the predicted vertebra segmentation. The X-ray was masked by the estimated rotated bounding box, followed by range normalization to  $[-1, 1]$ . That reflects the use of collimators during surgery to increase image contrast. Besides, it restricts the focus to anatomical regions and prevents irrelevant regions, such as background, from distorting the pose estimation.

**Landmark detection module:** In the clinic, correct standard projections are typically evaluated based on 2D anatomy-specific landmarks. This provides implicit additional 3D information based on anatomical prior knowledge. The landmark detection module mimics human decision-making to identify correct standard projections. A U-Net model (Klein et al., 2019) was trained for view-independent landmark detection of spinal X-rays on simulated data after ROI localization. The 2D landmark annotations were derived from automatically generated 3D landmarks by forward projection as described in Sec. 4.2.1. The model weights were optimized using L2 Loss on the difference between target and predicted heatmaps.

**Pose regression module:** The predicted heatmaps were used as input for the pose regression network. Landmark prediction uncertainties were covered by training the pose regressor on the predicted landmarks instead of the ground truth.

**Training procedure:** The models were implemented using PyTorch v1.6.0 and trained with an 11 GB GeForce RTX 2080 Ti. Optimization was performed with the Adam optimizer (Kingma and Ba, 2014), a base learning rate of  $\eta = 10^{-4}$  and a batchsize of 32.

**Objective function:** Given the ROI localization model with weight parameters  $W_{ROI}$ , the landmark detection module with weight parameters  $W_{LM}$  and the pose regression model with weight parameters  $W_{pose}$ , the models were trained by optimizing the objective function

$$\mathcal{L}_{total} = \min_{W_{ROI}} \mathcal{L}_{ROI} + \min_{W_{LM}} \mathcal{L}_{LM} + \min_{W_{pose}} \mathcal{L}_{pose}$$

with

$$\begin{aligned} L_{ROI} &= L_{Dice} + L_{CE} \\ L_{LM} &= L2 \\ L_{pose} & \quad (4.1). \end{aligned}$$

During inference, test time ensembling was employed for all proposed C-arm positioning methods: The initial X-ray was augmented with different in-plane rotations, and the median of all derived pose predictions was computed to increase robustness.

The different proposed approaches for C-arm pose estimation for standard projections were solely trained on simulated X-rays since they uniquely provide ground truth labels for arbitrary training examples. The transfer from simulated training to real X-rays is additionally challenged by intraoperative confounding factors, e.g., surgical implants. In Sec. 5.2.4, 5.2.5 the proposed methods are compared with respect to their performance on synthetic and real X-rays. The performance was assessed in terms of accuracy, robustness, and generalization capabilities.

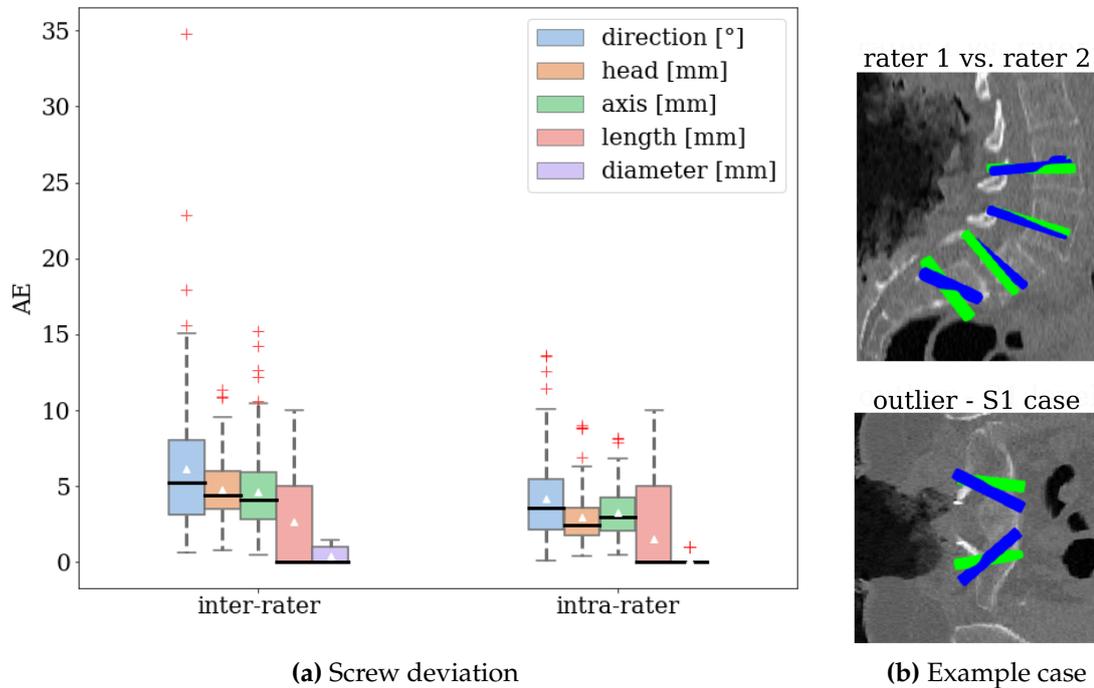
## Experiments and Results

The proposed computer-assisted methods for CT-navigated and fluoroscopy-guided spinal instrumentation were carefully evaluated. The experiments and results presented in this chapter are structurally aligned with the previous chapter: The first part presents the results for pedicle screw planning for CT-navigated spinal surgery (Section 5.1) and the second part summarizes the results for C-arm positioning for fluoroscopy-guided spinal surgery (Section 5.2).

### 5.1 Pedicle screw planning for CT-navigated spinal surgery

This section presents results on manual and automatic screw planning accuracy and efficiency (Sec. 5.1.1, 5.1.2). Therefore, an intra- and inter-rater study on manual screw planning was performed. Results were compared to automatic plans obtained with the proposed screw planning methods via the auxiliary task of segmentation. Automatic positioning results were clinically evaluated by a spinal neurosurgeon using the Gertzbein-Robbins (GR) classification (Sec. 5.1.3). Additionally, the proposed approach was compared to a baseline approach based on direct convolutional neural network (CNN) landmark regression. In Sec. 5.1.4 a study is presented assessing how accurately the pre-planned screw trajectories were transferred to the patient using CT navigation. The structure of the section is guided by research questions (RQ).

**Evaluation metrics:** For accuracy evaluation of screw planning, the Euclidean distance of *head*, *axis* point, the absolute error of *length*, *diameter*, and the angular deviation of screw trajectories were reported. The deviation of screw trajectories  $AE_{direction}$ , given by the angle formed by the intersection of the vectors representing the predicted and reference screw *direction* ( $d_p, d_r \in \mathbb{R}^3$ ), was computed by  $AE_{direction}(d_p, d_r) := \arccos(\frac{\langle d_p, d_r \rangle}{\|d_p\| \cdot \|d_r\|})$ . Besides the screw parameter deviations, also the geometric *overlap* of corresponding screw plans ( $S_i, S_j$ ) was measured by the Sørensen-Dice coefficient, a commonly used evaluation metric for image segmentation. It is given by  $Dice(S_i, S_j) := 2 \cdot |S_i \cap S_j| / (|S_i| + |S_j|) \in [0, 1]$ , where  $|\cdot|$  indicates the volumes of the sets (Dice, 1945).



**Figure 5.1: Manual pedicle screw planning accuracy.** Accuracy was assessed in an intra- and inter-rater study on a subset of 24 cases with 130 screws (a). Outliers can be related to the S1 vertebra level where the screw trajectory has more degrees-of-freedom (dof) as visualized for one exemplary case (b).

### 5.1.1 Manual pedicle screw planning accuracy

**RQ 1.1:** How much variation results from manual planning?

The manual pedicle screw planning accuracy was assessed in an intra- / inter-rater study. Screw plans from the day of surgery, extracted from the registry, served as ground truth reference (rater A). A randomly selected subset (24 cases, 130 screws) covering vertebra levels second lumbar vertebra (L2) (2 screws), third lumbar vertebra (L3) (20 screws), fourth lumbar vertebra (L4) (38 screws), fifth lumbar vertebra (L5) (42 screws), first sacral vertebra (S1) (28 screws) (Fig. 4.1c). The subset was additionally complemented with retrospective screw annotations by a second rater B, who assessed the computed tomography (CT) twice, as described in Sec. 4.1.1. Fig. 5.1b, top shows the planning results of rater A (blue) and rater B (green) exemplary for one CT scan. Results of the manual pedicle screw planning deviations assessed in an intra-rater (rater B vs. rater B) and

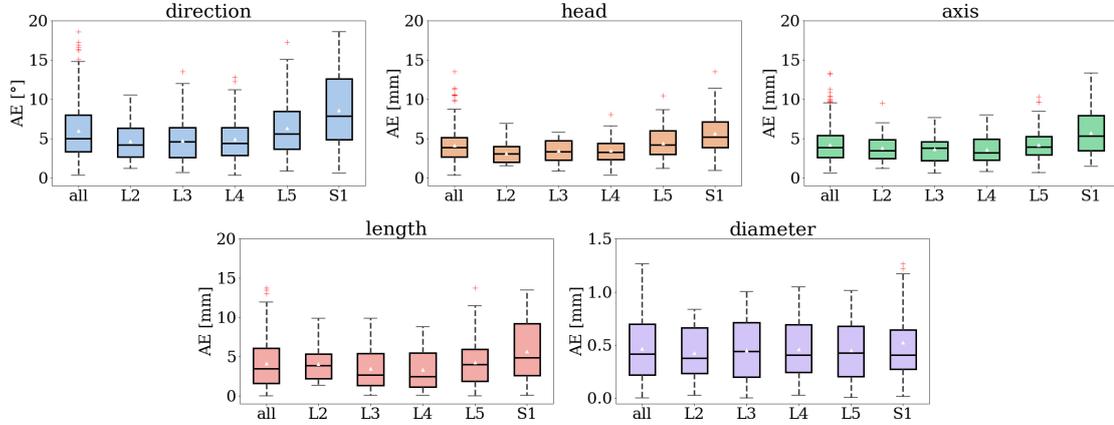
inter-rater (rater A vs. rater B) evaluation on the whole subset are shown in Fig. 5.1a across all vertebra levels. For the different screw parameters the inter-rater analysis (rater A vs. rater B) resulted in deviations ( $\mu \pm \sigma$ ) of *direction* =  $6.1^\circ \pm 4.6^\circ$ , *head* =  $4.8\text{mm} \pm 1.9\text{mm}$ , *axis* =  $4.6\text{mm} \pm 2.6\text{mm}$ , *length* =  $2.7\text{mm} \pm 2.9\text{mm}$ , *diameter* =  $0.4\text{mm} \pm 0.5\text{mm}$  and the intra-rater analysis (rater B vs. rater B) resulted in *direction* =  $4.1^\circ \pm 2.7^\circ$ , *head* =  $2.9\text{mm} \pm 1.7\text{mm}$ , *axis* =  $3.3\text{mm} \pm 1.6\text{mm}$ , *length* =  $1.5\text{mm} \pm 2.5\text{mm}$ , *diameter* =  $0.2\text{mm} \pm 0.4\text{mm}$ . The intra-rater variations between repeated manual planning of a single expert was significantly lower than inter-rater variations across all screw parameters ( $p \ll 0.01$ ). The outliers in screw *direction* result from the S1 vertebra level as shown exemplary for the inter-rater screw planning results with the highest deviation in Fig. 5.1b, bottom.

### 5.1.2 Automatic pedicle screw planning accuracy

**RQ 1.2:** Do the automatic screw plans agree with the manual plans?

Automatic screw planning results were compared to manually planned screws serving as the ground truth. The computed deviations were additionally set into context to inter- and intra-rater differences of manual screw planning, and the planning time was compared. Further, the proposed approach via the auxiliary segmentation task was compared to the baseline approach based on:) direct CNN landmark regression.

**Quantitative results:** The proposed methods for automatic pedicle screw planning, trained in a 5-fold cross validation, was compared on an external test set (55 cases, 328 screws) not seen during training, thereby testing its ability to handle inter-anatomical variation. The complete test set covered the vertebra levels L2 (20 screws), L3 (62 screws), L4 (90 screws), L5 (92 screws), S1 (64 screws) (Fig. 4.1b). Quantitative results of automatic pedicle screw planning with the proposed approach via segmentation (Sec. 4.1.3) are shown in Fig. 5.2, where the colors indicate the screw deviations for the different dof separated into vertebra levels. The label *all* indicates the results for the whole test set independent of the corresponding vertebra level. Across all vertebra levels mean screw deviations and corresponding standard deviations of *direction* =  $5.9^\circ \pm 3.8^\circ$ , *head* =  $4.1\text{mm} \pm 2.0\text{mm}$ , *axis* =  $4.2\text{mm} \pm 2.2\text{mm}$ , *length* =  $4.1\text{mm} \pm 3.1\text{mm}$ , *diameter* =  $0.5\text{mm} \pm 0.3\text{mm}$  were computed. To evaluate the agreement of screw plans with regard to the spinal level, a vertebra level-specific evaluation was performed. It illustrates that the deviations increase the more caudal spinal instrumentation was performed. The highest individual errors were observed for the S1 vertebra level.

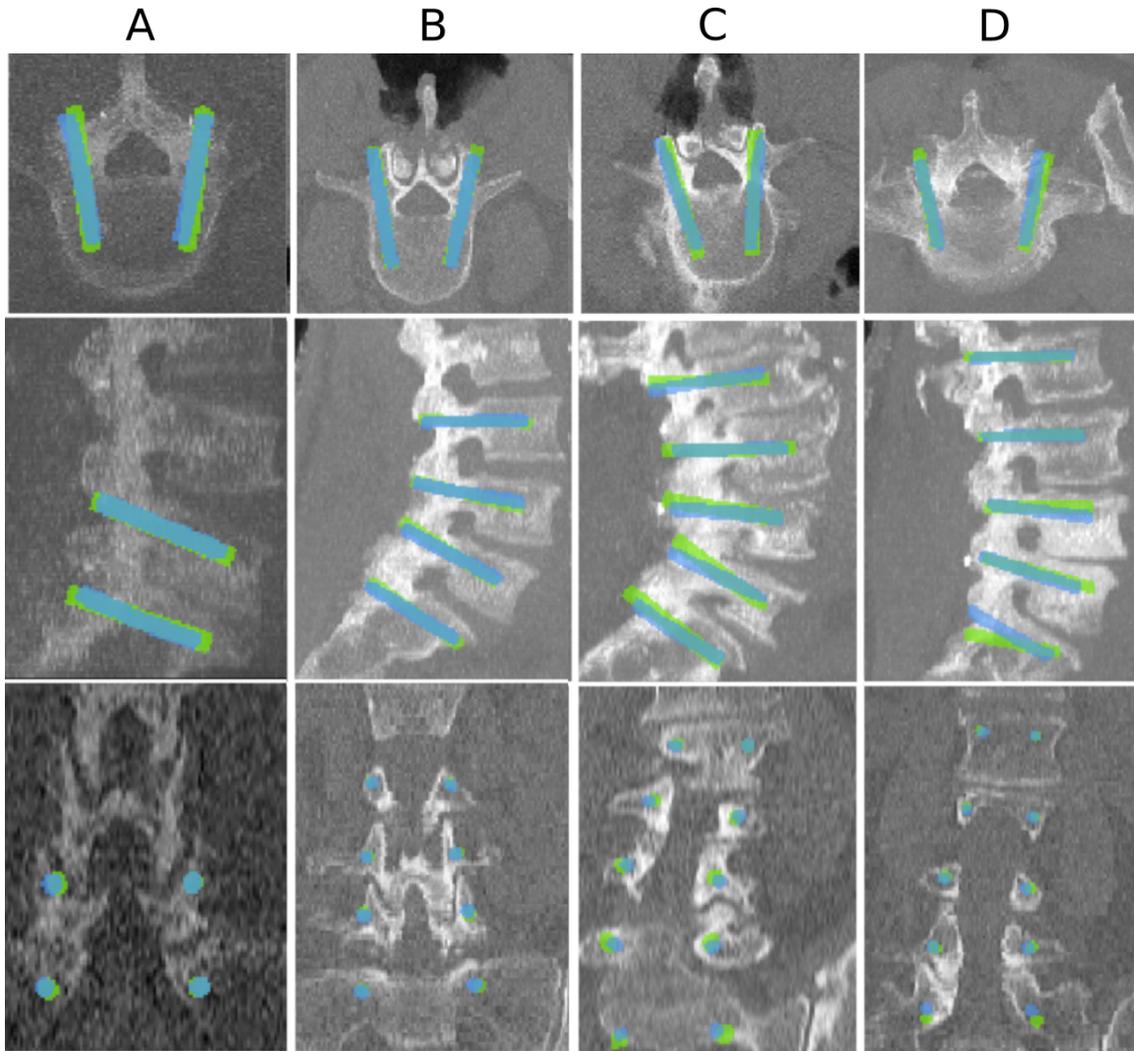


**Figure 5.2: Quantitative results of automatic pedicle screw planning** extended for larger test dataset from Kausch et al. (2021a). The predictions were obtained after ensembling of the models trained in a 5-fold cross-validation. It is shown the absolute error for the different screw parameters: *direction*, *head*, *axis*, *length*, *diameter*.

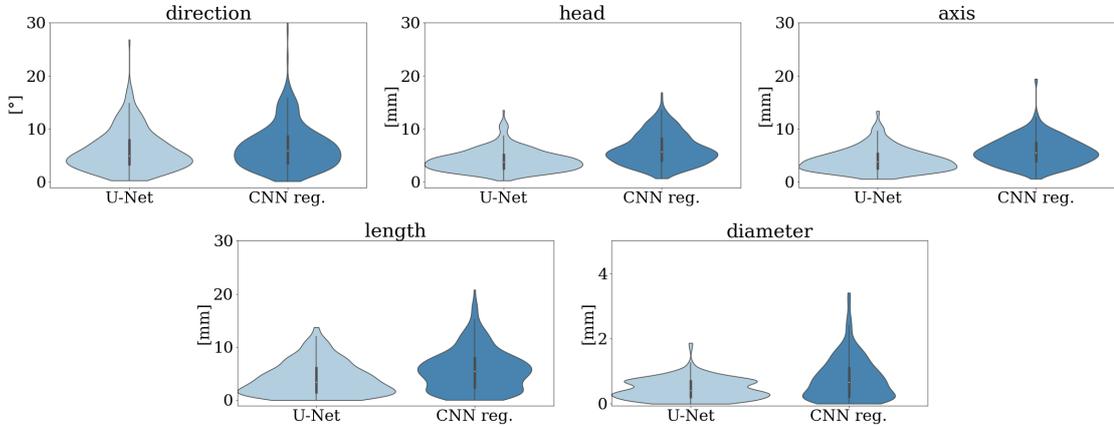
**Qualitative results:** Qualitative results of the proposed approach are shown in Fig. 5.3 for four exemplary cases (column-wise) and selected axial, sagittal and coronal slices (row-wise) for visualization. A high agreement between automatically generated screw plans (highlighted in blue) and manually created screw plans (highlighted in green) can be observed. The cases were selected to visualize a range of clinical scenarios: A and B show standard anatomical cases with short and long-range instrumentation, C shows a scoliosis patient, and D shows different planning results for the S1 level in the sagittal view.

**RQ 1.3:** Does the proposed approach via an auxiliary segmentation task benefit performance compared to direct CNN landmark regression?

**Comparison to baseline approach:** The performance of the proposed automatic pedicle screw planning via an auxiliary segmentation task was compared to the baseline approach based on direct CNN landmark regression (Sec. 4.1.2). While the *head*, *axis* and *diameter* parameter were directly regressed, the *direction* and *length* was subsequently computed. The performance comparison on the whole test set (55 cases, 328 screws) is visualized in Fig. 5.4. The proposed approach significantly outperforms the baseline approach across all screw parameters ( $p \ll 0.01$ ) except for the *direction* where the difference is not statistically significant.



**Figure 5.3: Visual results of automatic pedicle screw planning** taken from Kausch et al. (2021a). Visual comparison of automatically obtained (blue) and manually defined (green) pedicle screws for three different illustrative cases, one with scoliosis (middle column). Each column visualizes one case in selected axial, sagittal and coronal planes.



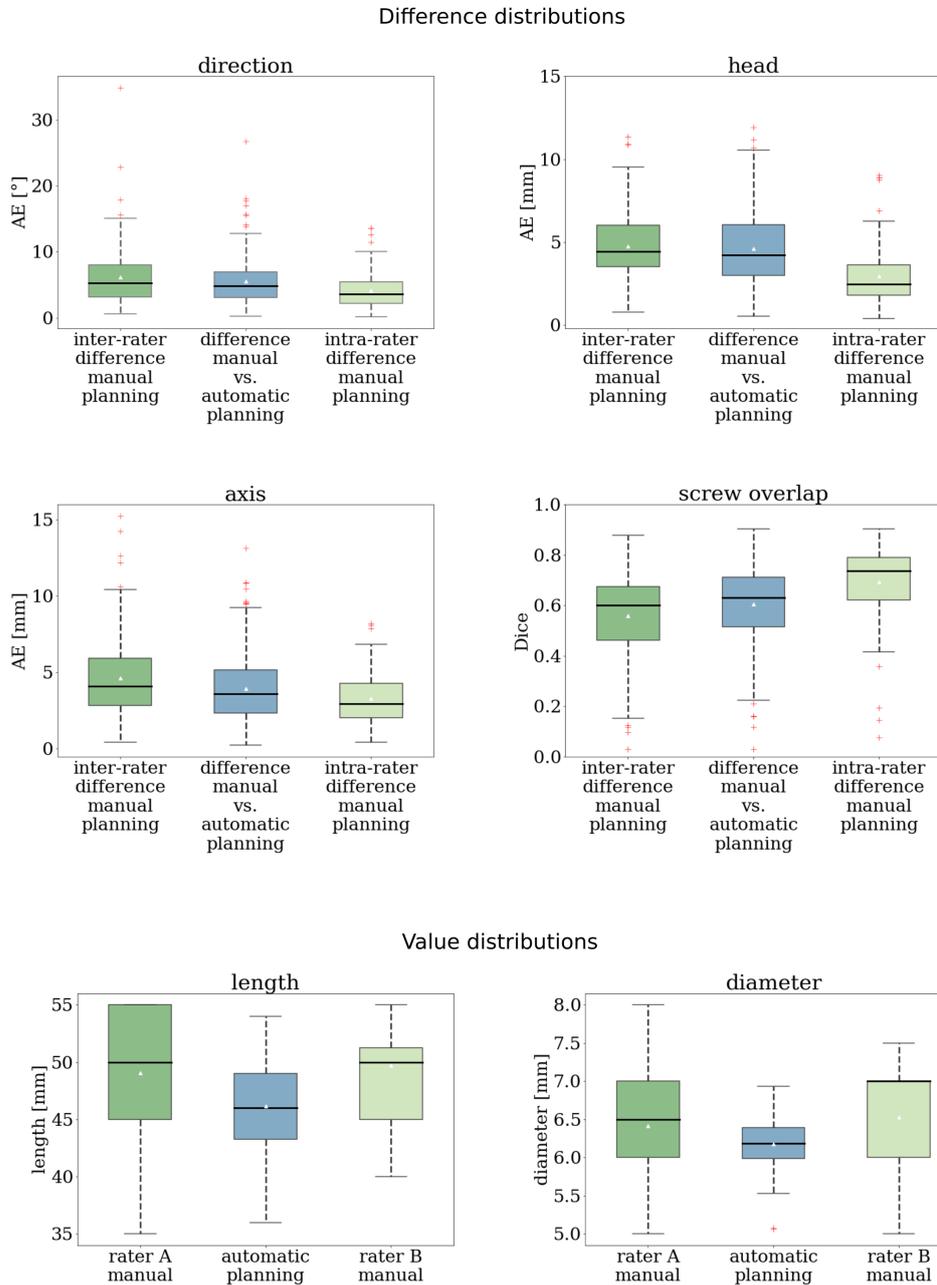
**Figure 5.4: Performance comparison: direct regression (dark blue) vs. segmentation (light blue) extended for larger test dataset from Kausch et al. (2021a).**

**RQ 1.4:** Do the manual vs. automatic planning deviations comply with the acceptable variation observed in manual planning?

**Inter- / Intra-rater study:** To further evaluate the performance of the proposed method, the automatic pedicle screw planning accuracy was compared to the inter- and intra-rater manual planning accuracy on a randomly selected subset of retrospective CTs (24 cases, 130 screws) (Sec. 5.1.1). Agreement between manual and automatic plans was assessed by computing the screw parameter deviations and the Sørensen-Dice coefficient. Furthermore, the value distributions of screw *length* and *diameter* parameters were compared between manual and automatic plans, respectively. An overview of the qualitative results for the different parameters is given in Fig. 5.5, along with a statistical significance evaluation in Tab. 5.1. For significance testing, the non-parametric Wilcoxon-signed-rank-test was employed after Shapiro-Wilk-test for normality distribution.

The screw *direction* varied between automatic and manual plans with  $5.51 \pm 3.64^\circ$ . This difference is comparable to the observed inter-rater difference of  $6.14 \pm 4.58^\circ$  ( $p = 0.04$ ) between two screw plans manually performed by rater A and B but significantly greater than the intra-rater difference of  $4.15 \pm 2.71^\circ$  ( $p \ll 0.01$ ) between two consecutive screw plans performed by rater B. Similarly, the deviations of the screw *head* points with  $4.61 \pm 2.27\text{mm}$  between automatic and manual plans were comparable to the inter-rater difference of  $4.77 \pm 1.95\text{mm}$  ( $p = 0.30$ ) but significantly greater than the intra-rater difference of

## 5.1. Pedicle screw planning for CT-navigated spinal surgery



**Figure 5.5: Comparison of automatic vs. manual screw planning accuracy** adapted from Scherer et al. (in submission). Manually planned screws serve as the ground truth and differences were additionally set into context by comparison to inter-rater and intra-rater differences of manual screw planning, respectively.

	<b>difference manual vs. automatic planning</b>	<b>inter-rater difference manual planning</b>		<b>intra-rater difference manual planning</b>	
	$\mu \pm \sigma$	$\mu \pm \sigma$	p-value	$\mu \pm \sigma$	p-value
<i>direction</i> [°]	5.51 ± 3.64	6.14 ± 4.58	0.04	4.15 ± 2.71	≪ 0.01
<i>head</i> [mm]	4.61 ± 2.27	4.77 ± 1.95	0.30	2.94 ± 1.70	≪ 0.01
<i>axis</i> [mm]	3.96 ± 2.19	4.64 ± 2.63	≪ 0.01	3.30 ± 1.65	≪ 0.01
<i>overlap</i> [Dice]	0.61 ± 0.16	0.56 ± 0.18	≪ 0.01	0.69 ± 0.14	≪ 0.01

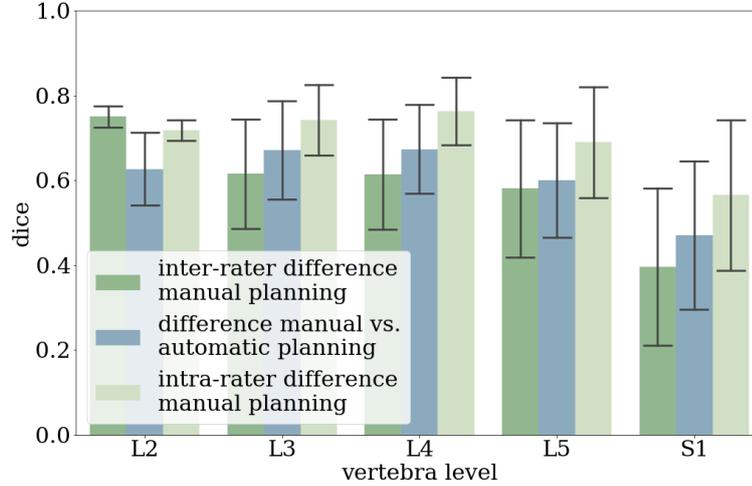
  

	<b>automatic planning</b>	<b>rater A manual</b>		<b>rater B manual</b>	
	$\mu \pm \sigma$	$\mu \pm \sigma$	p-value	$\mu \pm \sigma$	p-value
<i>length</i> [mm]	46.18 ± 3.63	49.08 ± 4.48	≪ 0.01	49.73 ± 3.87	≪ 0.01
<i>diameter</i> [mm]	6.18 ± 0.33	6.42 ± 0.63	≪ 0.01	6.53 ± 0.69	≪ 0.01

**Table 5.1: Statistical significance evaluation of automatic vs. manual screw planning** adapted from Scherer et al. (in submission). On the top, the difference of automatic vs. manual screw plans is compared against inter- and intra-rater variations for different screw parameters (*direction*, *head*, *axis*, *overlap*). On the bottom, the distribution of screw parameters *length*, *diameter* is compared across automatic and manual plans. For significance testing, the Wilcoxon-signed-rank-test was employed.

$2.95 \pm 1.70\text{mm}$  ( $p \ll 0.01$ ). For the screw *axis* points, a deviation of  $3.96 \pm 2.19\text{mm}$  was measured between automatic and manual plans, which was significantly lower than the observed inter-rater variation of  $4.64 \pm 2.63\text{mm}$  ( $p \ll 0.01$ ) and intra-rater differences with  $3.30 \pm 1.65\text{mm}$  ( $p \ll 0.01$ ). The Dice coefficient measuring the proportional overlap between corresponding screw plans was  $0.61 \pm 0.16$  between automatic and manual plans showing a significantly higher agreement than for the inter-rater study with  $0.56 \pm 0.18$  ( $p \ll 0.01$ ). The greatest overlap was achieved by the intra-rater study with  $0.69 \pm 0.14$  with significant difference to both groups ( $p \ll 0.01$ ).

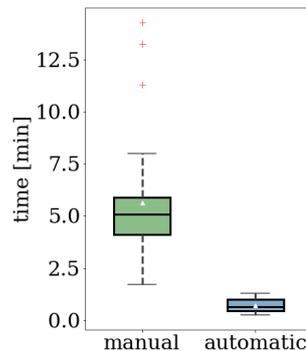
Comparing the distribution of screw *length* and *diameter* between automatic and manual planning, screw *length* was  $46.2 \pm 3.6\text{mm}$  in automatic and  $49.5 \pm 4.1\text{mm}$  in manual planning which indicates that automatically screws were planned significantly shorter ( $p \ll 0.01$ ). Similarly, the screw *diameter* resulting from automatic planning was significantly thinner compared to manually planning ( $6.2 \pm 0.3\text{mm}$  vs.  $6.5 \pm 0.7\text{mm}$ ,  $p \ll 0.01$ ). Between the manual planning performed by rater A and rater B no significant difference could be measured for screw *length* and *diameter*.



**Figure 5.6: Vertebra-level specific Dice evaluation for manual/automatic screw planning** adapted from Scherer et al. (in submission). The bars indicate the mean Dice and the error bars represent the standard deviation. The Dice is visualized separately for the different groups (inter-rater, automatic vs. manual, intra-rater).

To analyze the vertebra-level dependence of screw planning, the Dice coefficient is visualized separately for the vertebra levels L2-S1 (Fig. 5.6). According to ANOVA followed by Dunn’s multiple comparison test, Dice values were comparable across all levels, except the S1 vertebra level. Across all groups (inter-rater, automatic vs. manual, intra-rater), Dice coefficients were significantly higher for the lumbar levels than for the S1 level ( $p \ll 0.01$ ). This indicates higher planning variability, resulting from the S1 anatomy. However, within the S1 vertebra level, the observed automatic vs. manual planning deviations were comparable to the inter-rater variances ( $p = 0.07$ ).

The time consumption for manual screw planning of one CT scan was reported across the experiments and was compared to the *time* required by the proposed planning algorithm (Fig. 5.7). The automatic planning for one CT scan, including manual patch initialization, patch-based screw planning, and visualization, took on average  $41.8 \pm 20.2 \text{ sec}$  and was significantly faster than manual screw planning with  $338.7 \pm 186.3 \text{ s}$  (5 min 38 s,  $p \ll 0.01$ ) by 8-fold. Screw plans for one vertebra patch were automatically computed in  $4.9 \pm 0.1 \text{ s}$  on average. Automatic patch initialization via vertebra-instance segmentation took on average  $31.2 \pm 6.4 \text{ s}$ .



**Figure 5.7: Comparison of automatic vs. manual screw planning efficiency** adapted from Scherer et al. (in submission). The necessary *time* for planning one CT scan is compared.

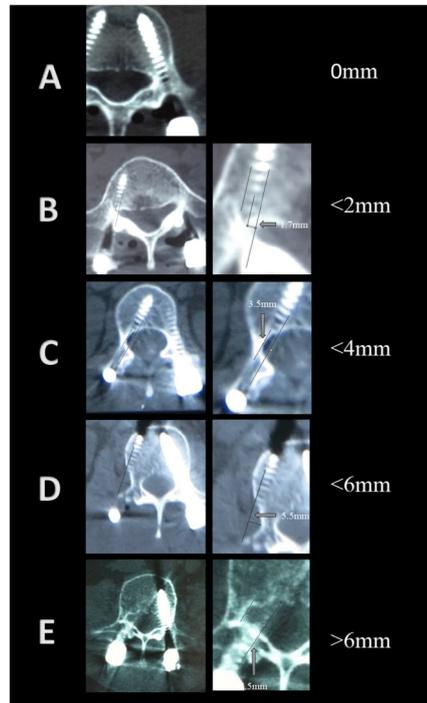
### 5.1.3 Clinical acceptability

**RQ 1.5:** Are the automatic screw plans clinically acceptable?

The clinical acceptability of the automatically planned screws (24 cases, 130 screws) was evaluated by a spinal neurosurgeon using the GR classification introduced in the following. Further, proximal facet violations and overall clinical acceptability on a binary scale were assessed.

**Gertzbein-Robbins classification:** To determine the screw placement accuracy, the GR classification is most commonly utilized (Aoude et al., 2015). It rates screws according to their positioning within cortical bone limits and scores the extent of possible cortical breach based on 2 mm increments (Fig 5.8). The grading system includes the following levels: The screw is entirely within the pedicle (GR Grade A), the screw breaches the pedicle’s cortex by < 2 mm (GR Grade B), pedicle cortical breach < 4 mm (GR Grade C), pedicle cortical breach < 6 mm (GR Grade D) and pedicle cortical breach > 6 mm (Grade E).

For the evaluation, it has to be considered that the reference data consisted of surgical planning data and therefore does not represent classic ground truth. 125/130 (96.2%) were classified as GR Grade A and 5/130 (3.8%) as GR Grade B, indicating that the automatically planned screws either showed no or only minor (< 2mm) cortical breach. Proximal facet violations were observed in 5/130 automatically planned screws resulting from 4 different patients affecting the L5 level. Despite the qualitative differences, 125/130 planned screws



**Figure 5.8:** Gertzbein-Robbins classification scores for pedicle screw accuracy are visualized on CT scans, taken from Fan et al. (2018). The deviation of the screw from the optimal trajectory is indicated, and screws are classified according to their extent of cortical breach into five levels A-E.

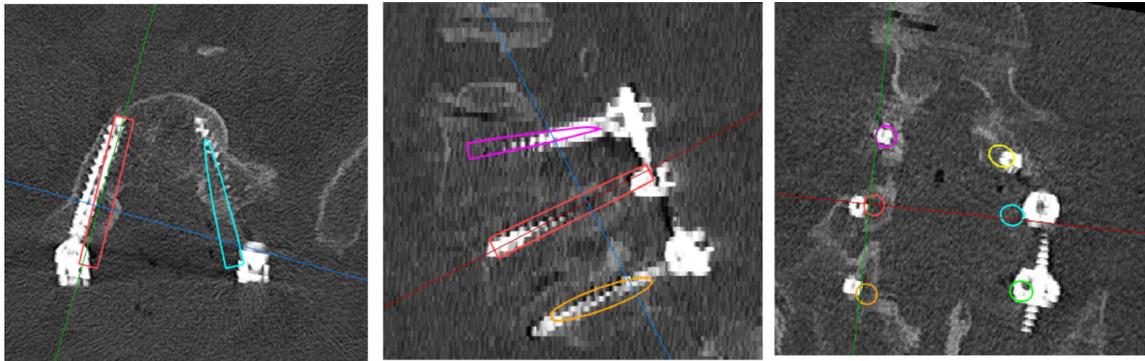
were classified as clinically acceptable for implantation (96.2%). All revisions were caused by proximal facet violations.

#### 5.1.4 Post-operative screw accuracy

**RQ 1.6:** How accurate are the screw plans transferred to the patient using CT navigation?

A study was performed to assess how accurately the preplanned screw trajectories were transferred to the patient using CT navigation.

**Retrospective planning data:** For validation, 24 cases with post-operative *control* CT for screw placement verification were collected from the registry of the university hospital Heidelberg. The CT-navigated spinal instrumentations were performed between 02/2014



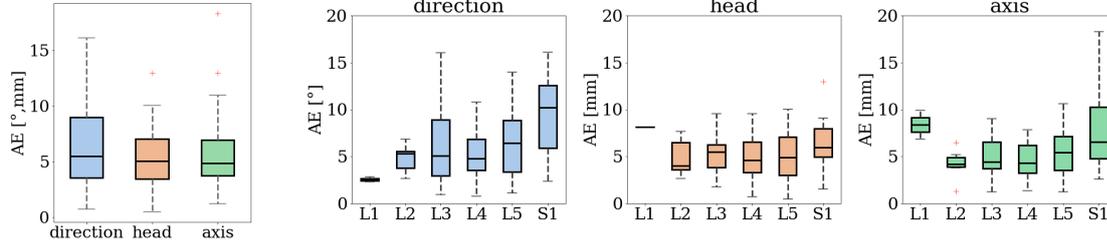
**Figure 5.9:** Example visualization of *navi* and *control* screws using MITK (Nolden et al., 2013). A variation between the *navi* screws (colored) and the *control* screws (white) is easily seen.

and 06/2016. Prior to surgery, 3D screw trajectories were planned in an intraoperative CT on a navigation workstation (*Stryker Spinemap 3D*). In total, the collected dataset contained 140 screws, the length of spinal instrumentation constructs ranged from 2-5 vertebra levels per patient. The complete test set covered the following vertebra levels: first lumbar vertebra (L1) (2 screws), L2 (8 screws), L3 (24 screws), L4 (40 screws), L5 (44 screws), S1 (22 screws).

**Experiment:** A spinal neurosurgeon retrospectively aligned screw trajectories with the screw locations in the post-operative *control* CTs using the commercial system *Stryker Spinemap 3D*. In the following, the screw trajectories obtained by navigation are called *control* screw trajectories, whereas the screw trajectories created prior to surgery on the intraoperative CT volume are denoted by *navi* screw trajectories. To assess the positioning difference between *navi* and *control* screws, the intraoperative and control CT were registered. Rigid registration was performed with the *Medical Imaging Interaction Toolkit (MITK) Registration Plugin*. Subsequently, the resultant rigid transformation was applied to the *navi* screw parameters, such that implanted screws can be compared to *navi* screw trajectories in the *control* CT using *MITK*.

**Results:** Fig. 5.9 shows the agreement between registered *navi* screws (colored) and *control* screws (white) in the *control* CT for one exemplary case. Variations in screw *direction*, *head*, *axis* point were derived and are shown in Fig. 5.10. Across all spinal level, a mean deviation of  $6.3 \pm 3.6^\circ$  was obtained for the screw *direction*,  $5.2 \pm 2.4\text{mm}$  for the screw *head* point and  $5.5 \pm 2.7\text{mm}$  for the screw *axis* point, respectively. To evaluate the results, the registration error of the navigation system, has to be considered which was on

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



(a) Overall screw placement deviation

(b) Screw placement deviation grouped by spinal level

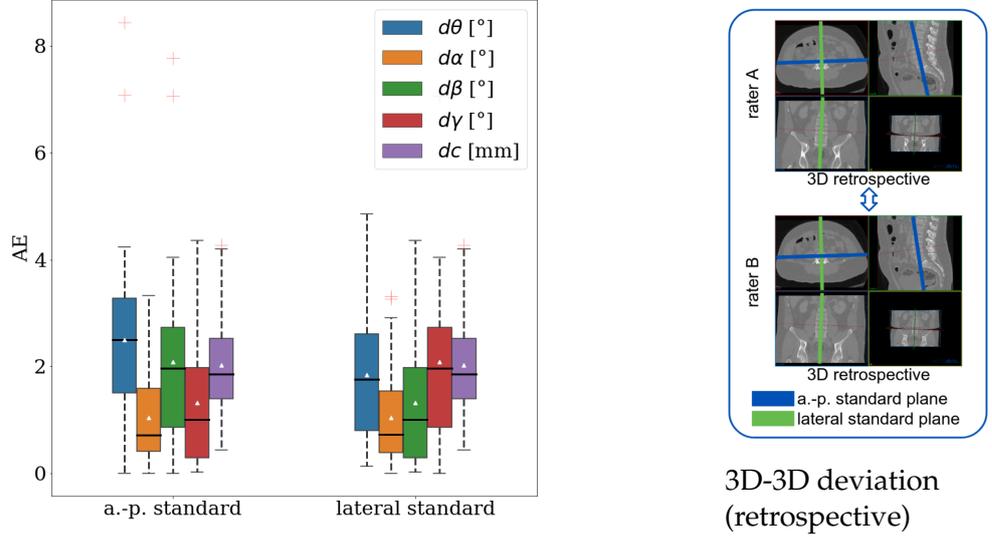
**Figure 5.10: Comparison of post-operative screw placement to intraoperative planning.**

average 1.9 mm (range 1-7mm).

## 5.2 C-arm positioning for fluoroscopy-guided spinal surgery

The proposed pipeline for automatic C-arm positioning solely trained on simulated data was carefully evaluated on synthetic digitally reconstructed radiograph (DRR) generated from computed tomography (CT) not seen during training and on real X-rays acquired in a large specimens study, covering different stages of a realistic clinical procedure, e.g., inserted k-wires and screws. At first, in Sec. 5.2.1 the accuracy of the training data is evaluated, serving as an upper bound for the reachable accuracy limit. The accuracy of the acquired validation data, consisting of real X-rays with pose labels, is analyzed in Sec. 5.2.2 and the pose annotations serving as ground truth were corrected accordingly. In Sec. 5.2.3 the two studies for assessing the manual C-arm positioning accuracy are presented, namely the *Swipe App* and the *inter- / intra-rater specimens study* of manual C-arm positioning. The first three sections collect the results of different approaches to derive relevant standard projection accuracy bounds, which were introduced in Sec. 4.2.3. Results are presented for the different approaches by referring to Fig. 4.13. Finally, Sec. 5.2.4 and 5.2.5 compare the performance of the baseline approach for C-arm positioning, the proposed 2-step direct intensity-based pose regression, and the 1-step sequential pose regression on synthetic and real X-rays, respectively. Influencing factors of specific design choices of the sequential approach on the overall performance were evaluated in Sec. 5.2.6. The structure of the section is guided by research questions (RQ).

**Evaluation metrics:** The pose accuracy was evaluated based on the angle  $\theta := \arccos(\langle v_{pred}, v_{gt} \rangle)$  between the principal rays of the ground truth  $v_{gt}$  and predicted pose  $v_{pred}$ . In addition, the mean absolute error ( $\mu$ ) and standard deviation ( $\sigma$ ) of rotation along



**Figure 5.11: Training data accuracy for standard projections**, i.e. inter-rater variance between 3D volume vs. 3D volume (assessed on 47 CTs).

individual axes denoted by  $d\alpha$ ,  $d\beta$ ,  $d\gamma$  are reported. Further,  $dc := \|c_{pred} - c_{gt}\|$  denotes the isocenter distance between the predicted and ground truth isocenter  $c_{pred}, c_{gt} \in \mathbb{R}^3$ . The translation of the projection in the detector plane is denoted by  $tx, ty \in \mathbb{R}$ .

### 5.2.1 Training data accuracy

**RQ 2.1:** How accurate is the reference data used as ground truth?

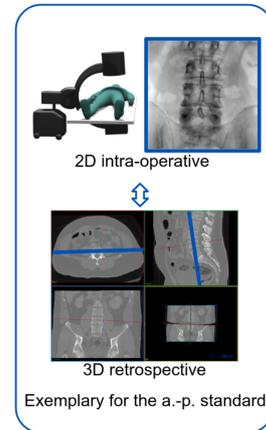
As described in Sec. 4.2.3 (a), the inter-rater variability of the 3D reference standard planes independently defined by rater A and B is computed. The a.-p. and lateral standards are in the vast majority of anatomical regions orthogonal to each other, including the lumbar spine. Thus, the a.-p. and lateral standard planes were defined for the fourth lumbar vertebra (L4) level under the additional prior of pairwise plane orthogonality. The two raters assessed the 47 retrospective CT volumes considered for DRR simulation. Let  $\mathbf{x}_i, \mathbf{n}_i^{standard} \in \mathbb{R}^3$ ,  $i \in \{A, B\}$ ,  $standard \in \{a.-p., lateral\}$  denote the intersection point and the unit normal of the defined reference planes for the a.-p. and lateral standard by rater A and B. Then, the inter-rater variance of the defined reference standard planes was assessed in terms of orientation differences  $d\theta = \arccos(\langle \mathbf{n}_A^{standard}, \mathbf{n}_B^{standard} \rangle)$  and 3D translational differences  $dc = \|\mathbf{x}_A - \mathbf{x}_B\|$ . Additionally, the orientation variation was

assessed specifically for the different rotational axis by taking rater *A* as reference. The rotational offset along the axis defined by the intersection of the coronal and sagittal plane is denoted by  $d\alpha$  (corresponds to the orbital C-arm rotation), the offset along the axis defined by the coronal (resp. sagittal) and axial plane is denoted by  $d\beta$  (corresponds to the angular C-arm rotation), and the in-plane rotation for the a.-p. (resp. lateral) standard. Fig. 5.11 visualizes the results separately for the a.-p. and lateral standard projection. For the a.-p. standard,  $\mu \pm \sigma$  inter-rater deviations of  $d\theta = 2.50 \pm 1.57^\circ$ ,  $d\alpha = 1.05 \pm 0.91^\circ$ ,  $d\beta = 2.09 \pm 1.55^\circ$ ,  $d\gamma = 1.32 \pm 1.17^\circ$ ,  $dc = 2.02 \pm 0.91\text{mm}$  were computed and for the lateral standard  $d\theta = 1.85 \pm 1.24^\circ$ ,  $d\alpha = 1.05 \pm 0.91^\circ$ ,  $d\beta = 1.32 \pm 1.17^\circ$ ,  $d\gamma = 2.09 \pm 1.55^\circ$ ,  $dc = 2.02 \pm 0.91\text{mm}$ , respectively.

### 5.2.2 Validation data accuracy

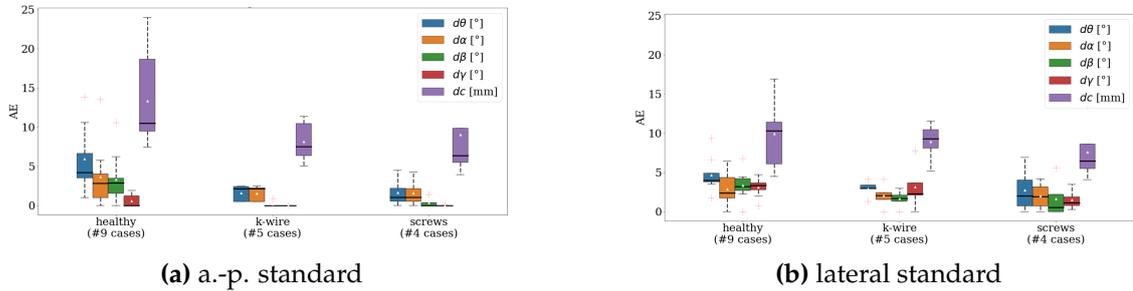
**RQ 2.2:** How accurate are standard projections positioned during the validation data acquisition?

To evaluate the proposed methods for automatic C-arm positioning, validation data was acquired in a specimens study. The experimental setup described in Sec. 4.2.2 enables direct pose labeling of the acquired dataset based on the metadata information of the C-arm. The pose labels denote the pose offset to a desired standard projection. However, if the manual positioned C-arm was not ideally aligned with the standard projections, the derived pose labels are distorted, leading to faulty ground truth. Therefore, pose labels were corrected by retrospectively defining the standard projections in the 3D reconstruction. The derivation between retrospectively defined 3D standard planes and 2D manual positioned standard projection is a more realistic measure for the clinical observed and acceptable variance of standard projections, as described in Sec. 4.2.3 (b). It was assessed on the 18 validation specimens. Fig. 5.12 illustrates the results for both standard projections, separated for the specimens without metal and those with metal. The specimens with metal were opened for instrumentation, such that the bone landmarks were visible and simplified the manual C-arm positioning. For the specimens without metal, a variation  $\mu \pm \sigma$  of  $d\theta = 5.96 \pm 3.98^\circ$ ,  $d\alpha = 3.67 \pm 4.12^\circ$ ,  $d\beta = 3.34 \pm 3.32^\circ$ ,  $d\gamma = 0.59 \pm 0.82^\circ$ ,  $dc = 13.33 \pm 5.83\text{mm}$  was computed for the a.-p. standard and  $d\theta = 4.70 \pm 2.18^\circ$ ,  $d\alpha = 2.87 \pm 2.06^\circ$ ,  $d\beta = 3.36 \pm 1.83^\circ$ ,  $d\gamma = 3.13 \pm 1.17^\circ$ ,  $dc = 9.97 \pm 4.29\text{mm}$  for the lateral



2D-3D deviation (retrospective)

projections, separated for the specimens without metal and those with metal. The specimens with metal were opened for instrumentation, such that the bone landmarks were visible and simplified the manual C-arm positioning. For the specimens without metal, a variation  $\mu \pm \sigma$  of  $d\theta = 5.96 \pm 3.98^\circ$ ,  $d\alpha = 3.67 \pm 4.12^\circ$ ,  $d\beta = 3.34 \pm 3.32^\circ$ ,  $d\gamma = 0.59 \pm 0.82^\circ$ ,  $dc = 13.33 \pm 5.83\text{mm}$  was computed for the a.-p. standard and  $d\theta = 4.70 \pm 2.18^\circ$ ,  $d\alpha = 2.87 \pm 2.06^\circ$ ,  $d\beta = 3.36 \pm 1.83^\circ$ ,  $d\gamma = 3.13 \pm 1.17^\circ$ ,  $dc = 9.97 \pm 4.29\text{mm}$  for the lateral



**Figure 5.12: Validation data accuracy after manual C-arm positioning, i.e. 3D volume vs. intraoperative projection.** Variance of standard beam direction between the standard projection obtained with manual C-arm positioning and the standard plane retrospectively defined in the reconstructed volume. Evaluation is performed separately for specimens without metal and with metal (open reduction). The variance serves as a measure for clinical acceptable deviations.

standard, respectively.

The corrected a.-p. and lateral standard projections are visualized in Fig. 5.13 for all 18 validation specimens.

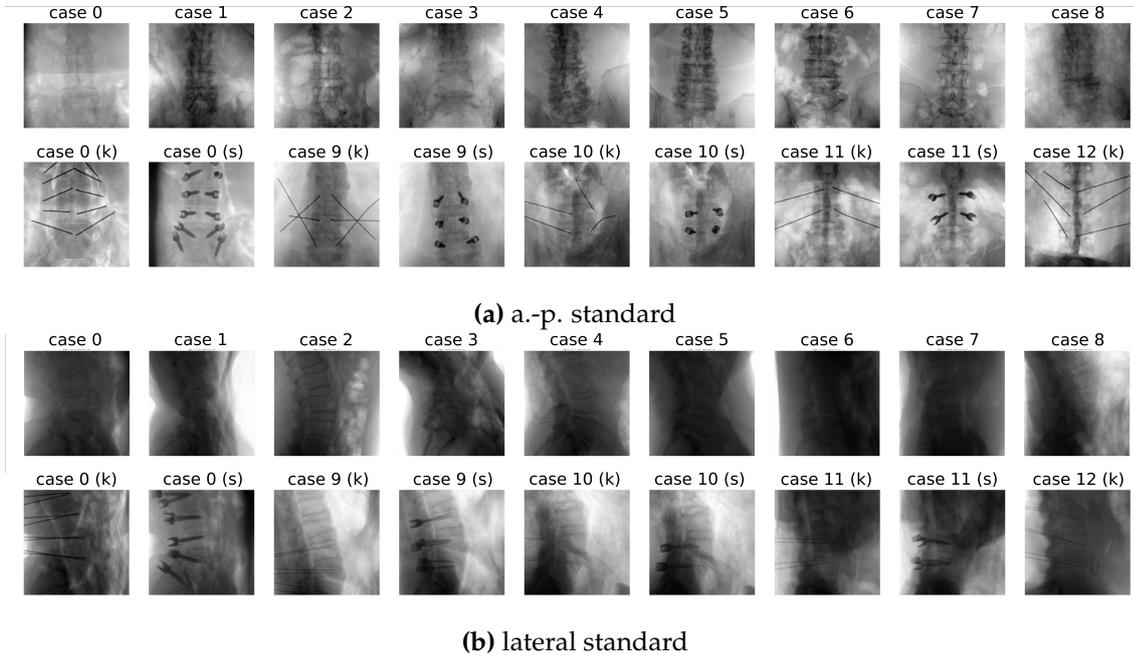
### 5.2.3 Manual C-arm positioning accuracy and efficiency

Two studies performed to assess the clinical accuracy of standard projections retrospectively and intraoperatively, as described in Sec. 4.2.3 (c), (d).

**RQ 2.3:** Which pose deviations are observed among clinically accepted standard projections?

**Retrospective assessment of standard projection accuracy (*Swipe App*):** To measure the angular variation of accepted standard projections (a.-p., lateral), a standard plane clinical accuracy tool (*Swipe App*) was developed that allows swiping to left or right to accept or reject a presented projection as standard projection (Sec. 4.2.3 (c)). Utilizing this, the generated DRR dataset with pose labels (Sec. 4.2.1) was presented to experts separately for varying angular ( $\beta$ ) and orbital rotation ( $\alpha$ ) for the a.-p. and lateral standard projection. This significantly reduces the amount of projection images and simplifies the decision for the expert. Three independent clinical experts carried out this study, who evaluated 2271 projection images in total for the a.-p. standard and 2191 images for the lateral standard,

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



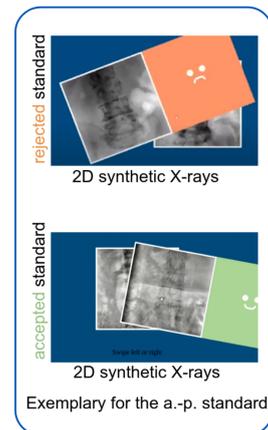
**Figure 5.13: Reference standard projections of 18 acquired validation specimens case 0 - case 12.** Thereby, k and s indicate the presence of k-wires or screws.

respectively.

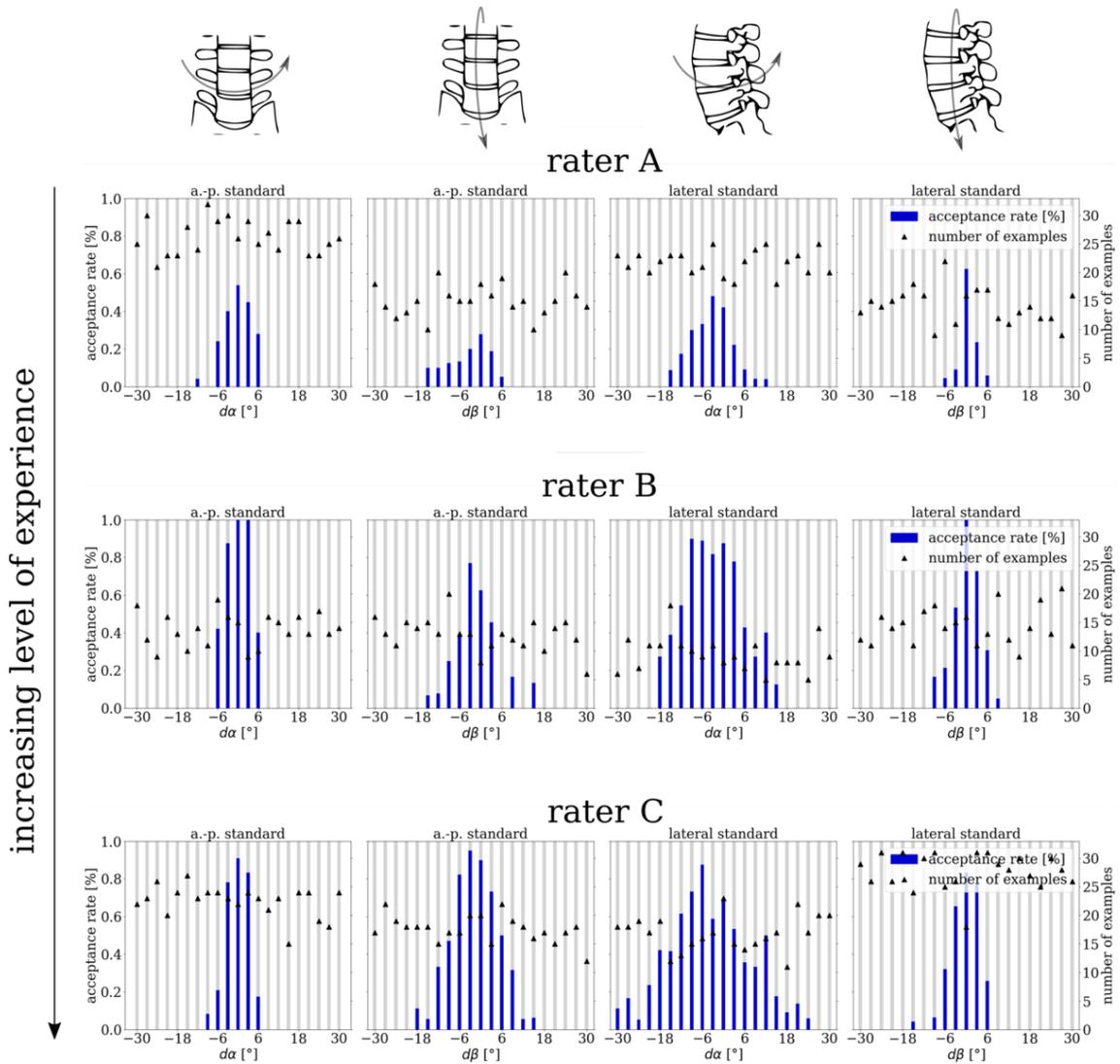
Tab. 5.2 details how many examples each expert assessed separated into angular degrees-of-freedom (dof), standard projection.

Fig. 5.14 visualizes the results of the rating for each expert and standard projection dof separately. The blue bars indicate the acceptance rate of presented projections as standard projection for different ground truth offsets from the reference standard projection ( $x$ -axis). The black triangles indicate how many examples were considered for each ground truth offset pose. The number varies since examples were presented randomly by the *Swipe app*.

To evaluate the angular variation of the a.-p. and lateral standards quantitatively, the ground truth angulations of all accepted projections (indicated by the blue bars in Fig. 4.14) were averaged, resulting in a deviation  $\mu \pm \sigma_\mu$  of  $d\alpha = 2.90 \pm 0.20^\circ$ ,  $d\beta = 5.17 \pm 0.22^\circ$  for the a.-p. standard and  $d\alpha = 7.59 \pm 1.77^\circ$ ,  $d\beta = 2.63 \pm 0.75^\circ$  for the



2D-2D deviation  
(retrospective)



**Figure 5.14: Swipe App: Measuring clinically acceptable deviations of standard projections.** Three raters assessed the acceptability of synthetic X-rays with varying pose offsets to the reference standard plane (a.-p. and lateral standard). Pose offsets were varied independently for the orbital rotation  $\alpha$  and the angular rotation  $\beta$ . Synthetic X-rays were simulated from 47 CTs, and experts performed the study for a fixed time frame of 15 mins per standard and dof. The raters have an increasing level of experience from top-bottom

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery

# examples	a.-p. standard			lateral standard		
	rotational dof			rotational dof		
	$\alpha$	$\beta$	$\Sigma$	$\alpha$	$\beta$	$\Sigma$
rater A	551	318	869	457	298	755
rater B	289	270	559	198	302	500
rater C	473	370	843	354	582	936
$\Sigma$			2271			2191

**Table 5.2: *Swipe App*: Number of examples assessed for each standard projection.** The number of examples  $N$  assessed by each expert for each separate dof ( $\alpha$ ,  $\beta$ ) is additionally indicated.

$\mu \pm \sigma$	a.-p. standard		lateral standard	
	rotational dof		rotational dof	
	$d\alpha$ [°]	$d\beta$ [°]	$d\alpha$ [°]	$d\beta$ [°]
rater A	$3.17 \pm 2.50$	$4.89 \pm 4.61$	$5.43 \pm 4.23$	$1.59 \pm 2.15$
rater B	$2.82 \pm 2.22$	$5.42 \pm 4.29$	$7.57 \pm 5.17$	$2.93 \pm 2.89$
rater C	$2.70 \pm 2.19$	$5.20 \pm 4.14$	$9.76 \pm 7.54$	$3.36 \pm 2.56$
global mean	$2.90 \pm 0.20$	$5.17 \pm 0.22$	$7.59 \pm 1.77$	$2.63 \pm 0.75$

**Table 5.3: *Swipe App*: Angular deviations ( $d\alpha$ ,  $d\beta$ ) for each standard projection.** In addition, the angular deviation is reported separately for each expert.

lateral standard, averaged across all raters. Rater-individual evaluations are summarized in Tab. 5.3.

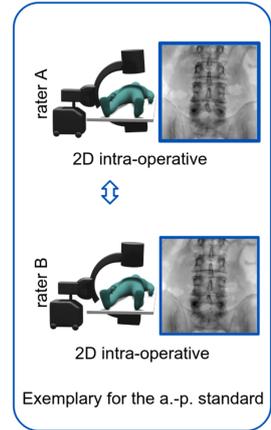
The derived retrospective measure serves as a baseline for the acceptable errors in clinical practice.

**RQ 2.4:** How accurate and efficient is the manual C-arm positioning?

**Intraoperative assessment of standard projection accuracy and efficiency (*Inter-rater study*):** The standard projection accuracy and efficiency under manual C-arm positioning were assessed in an inter- / intra-rater specimens study. Therefore, two experts independently positioned the C-arm to acquire a.-p. and lateral standard projections of vertebral levels first thoracic vertebra (Th1)-fifth lumbar vertebra (L5) on two specimens. This results in 17 standard projections, defined by each expert for each specimen. To allow for intra-rater variance assessment, one expert positioned the C-arm for one specimen

a second time. For manual C-arm positioning efficiency assessment, the number of X-rays, the total dose, and the time for manual C-arm positioning were reported additionally.

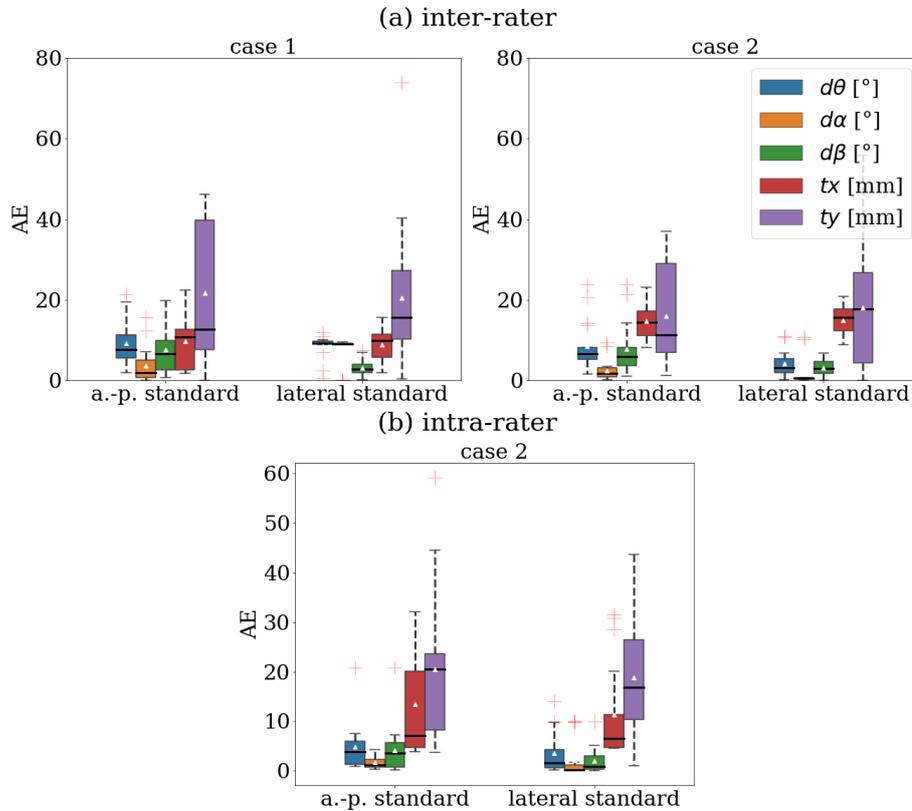
**Manual C-arm positioning accuracy:** To assess the intra- and inter-rater pose difference, the 3D reconstruction corresponding to the respective vertebra-level specific standard projection were registered. Results are visualized in Fig. 5.15. In the inter-rater study, mean positioning differences of  $d\theta = 8.89 \pm 5.77^\circ$ ,  $d\alpha = 3.06 \pm 3.65^\circ$ ,  $d\beta = 7.71 \pm 6.15^\circ$ ,  $tx = 12.24 \pm 5.80mm$ ,  $ty = 18.85 \pm 15.11mm$  were obtained for the a.-p. standard and  $d\theta = 6.36 \pm 3.71^\circ$ ,  $d\alpha = 4.58 \pm 4.46^\circ$ ,  $d\beta = 3.08 \pm 2.09^\circ$ ,  $tx = 11.83 \pm 4.88mm$ ,  $ty = 19.24 \pm 16.98mm$  for the lateral standard, respectively. Specimen-specific inter-rater pose differences are given in the Chapter A, Tab. A.1. In the intra-rater study, mean positioning differences of  $d\theta = 8.89 \pm 5.77^\circ$ ,  $d\alpha = 3.06 \pm 3.65^\circ$ ,  $d\beta = 7.71 \pm 6.15^\circ$ ,  $tx = 12.24 \pm 5.80mm$ ,  $ty = 18.85 \pm 15.11mm$  were obtained for the a.-p. standard and  $d\theta = 6.36 \pm 3.71^\circ$ ,  $d\alpha = 4.58 \pm 4.46^\circ$ ,  $d\beta = 3.08 \pm 2.09^\circ$ ,  $tx = 11.83 \pm 4.88mm$ ,  $ty = 19.24 \pm 16.98mm$  for the lateral standard, respectively.



2D-2D deviation (intraoperative)

**Manual C-arm positioning efficiency:** The number of X-rays, time, and dose requirements of manual C-arm positioning for the two standard projections are reported in Fig. 5.16, separated for each specimen and the two experts. Expert 1 assessed 51 vertebrae (1 round for *case 1*, 2 rounds for *case 2*), and expert 2 assessed 34 vertebrae (1 round for *case 1* and *case 2*). Averaged across the specimens and rounds, expert 1/expert 2 required a median number of 5/8 (range 2-15/range 3-29) X-rays for a.-p. standard positioning and 4/6 (range 1-30/range 1-28) for lateral standard positioning. Expert 2 required significantly more X-rays for manual C-arm positioning ( $p \ll 0.01$ , t-test). The number of required positioning X-rays correlates with the experience level of the surgeon. Regarding the time for manual positioning, expert 1/expert 2 required on average 74/80.5 s (range 22-193 s/range 26-292 s) for the a.-p. standard and 44/76.5 s (range 5-474 s/range 5-306 s) for the lateral standard. For the first acquired vertebra (Th1), positioning acquisitions were significantly increased, which can be attributed to the localization task. After initial positioning to Th1, the remaining vertebrae are easier to locate along the spinal curvature. Median doses of  $25/42.5\mu Gy \cdot cm^2$  (range 6-205 $\mu Gy \cdot cm^2$ /range 9-144 $\mu Gy \cdot cm^2$ ) for the a.-p. standard and  $53/85.5\mu Gy \cdot cm^2$  (range 3-836 $\mu Gy \cdot cm^2$ /range 12-606 $\mu Gy \cdot cm^2$ ) for the lateral standard were computed. Expert 2 requires significantly more dose than expert 1 for both stan-

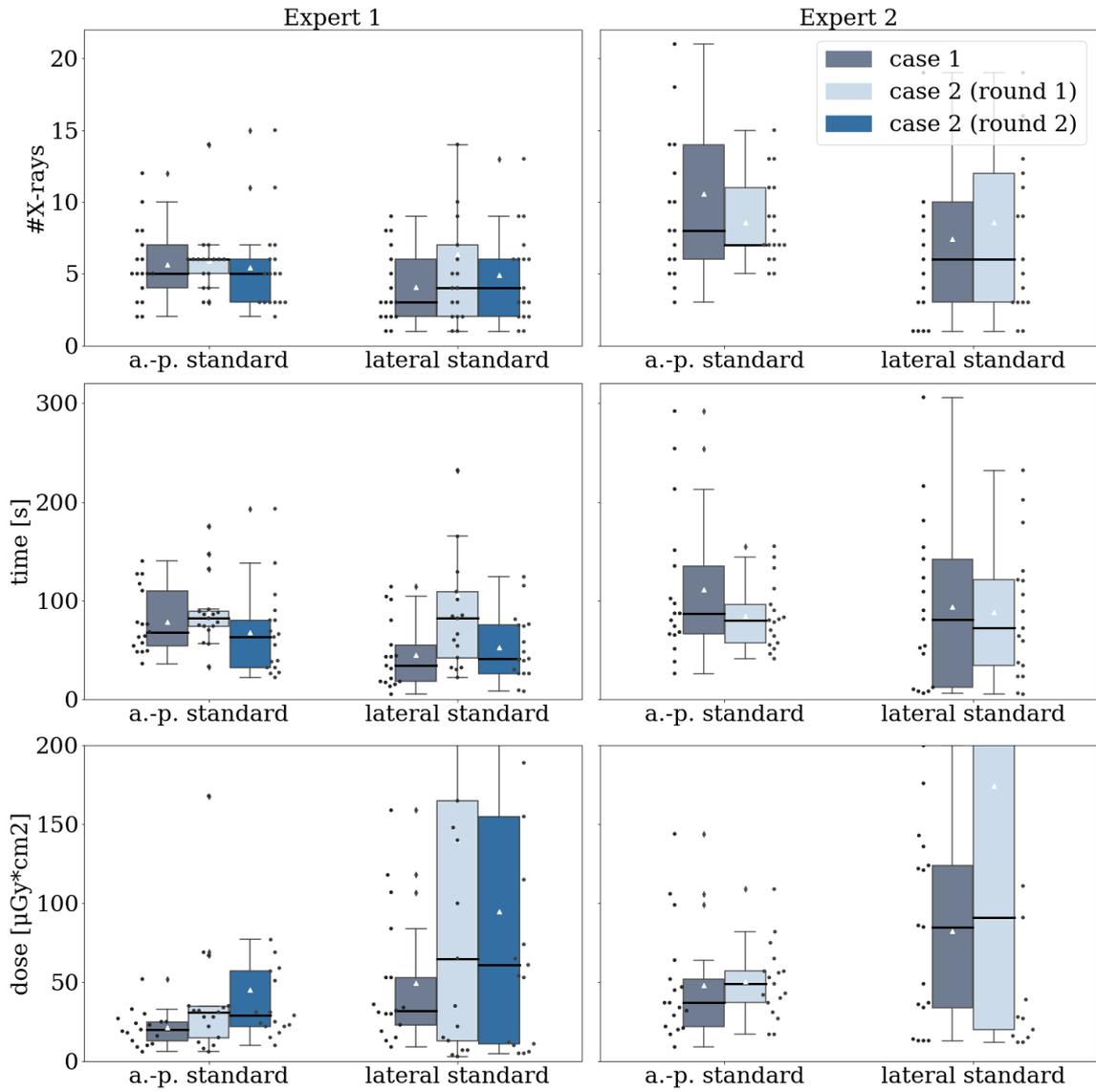
## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



**Figure 5.15: Inter-/Intra-rater study to assess manual C-arm positioning accuracy.** Two experts defined the a.-p. and lateral standard for vertebra levels Th1-L5 for two specimens. The pose differences were assessed in an inter- and intra-rater evaluation.

dards. The dose required for the lateral projection is significantly higher than for the a.-p. projection ( $p < 0.015$ ), which reflects the different body transmission depths. *Case 2* was anatomical more challenging than *case 1* due to severe lumbar scoliosis, which is reflected by a higher number of X-rays and time for lateral standard positioning of vertebra levels Th12-L5. A vertebra-level specific evaluation is shown in Fig. A.1. The number of X-rays and the time are correlated, while the dose also depends on the tissue penetration depth.

Tab. 5.4 summarizes the results for the different accuracy measures for standard projections and compares them to the obtained results in automatic C-arm positioning presented in the following sections.



**Figure 5.16: Efficiency of manual C-arm positioning** involving the number of X-rays, time, and dose requirement. Two experts manually positioned the C-arm for two specimens across vertebra level Th1-L5 in the a.-p. and lateral standard.

**Table 5.4: Comparison of mean positioning accuracies across different studies.** The measures are evaluated on different datasets. The inter-rater 3D deviation was assessed on the 47 retrospective CT volumes (L4 level). The 2D-3D deviation was assessed on the 18 validation specimens (L4 level). The retrospective 2D-2D positioning variance was assessed on a randomly selected subset from the simulated DRR set (20727 DRRs from 47 CTs). The intraoperative 2D-2D manual C-arm positioning differences were assessed on 2 specimens in vertebra levels Th1-L5.

(a) a.-p. standard				
	$d\theta[^\circ]$	$d\alpha[^\circ]$	$d\beta[^\circ]$	$d\gamma[^\circ]$
<b>3D deviation</b>	$2.50 \pm 1.57$	$1.05 \pm 0.91$	$2.09 \pm 1.55$	$1.32 \pm 1.17$
<b>2D-3D deviation</b>	$5.96 \pm 3.98$	$3.67 \pm 4.12$	$3.34 \pm 3.32$	$0.59 \pm 0.82$
<i>Swipe App</i>	n/a	$2.90 \pm 2.30$	$5.17 \pm 4.35$	n/a
<b>Inter-rater study</b>	$8.89 \pm 5.77$	$3.06 \pm 3.65$	$7.71 \pm 6.15$	n/a
<b>Intra-rater study</b>	$4.79 \pm 4.74$	$1.68 \pm 1.15$	$4.12 \pm 4.99$	n/a
<b>Simulated experiments</b>	$3.48 \pm 2.48$	$2.21 \pm 1.97$	$2.33 \pm 2.24$	$1.65 \pm 1.31$

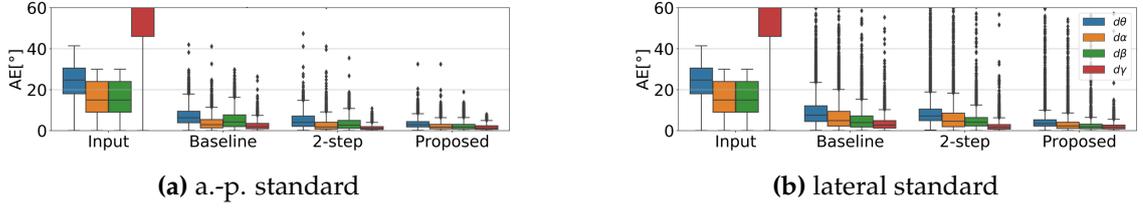
  

(b) lateral standard				
	$d\theta[^\circ]$	$d\alpha[^\circ]$	$d\beta[^\circ]$	$d\gamma[^\circ]$
<b>3D deviation</b>	$1.85 \pm 1.24$	$1.05 \pm 0.91$	$1.32 \pm 1.17$	$2.09 \pm 1.55$
<b>2D-3D deviation</b>	$4.70 \pm 2.18$	$2.87 \pm 2.06$	$3.36 \pm 1.83$	$3.13 \pm 1.17$
<i>Swipe App</i>	n/a	$7.56 \pm 5.65$	$2.63 \pm 2.53$	n/a
<b>Inter-rater study</b>	$6.36 \pm 3.71$	$4.58 \pm 4.46$	$3.08 \pm 2.09$	n/a
<b>Intra-rater study</b>	$3.51 \pm 4.20$	$2.17 \pm 3.88$	$1.97 \pm 2.62$	n/a
<b>Simulated experiments</b>	$4.83 \pm 6.00$	$3.53 \pm 5.16$	$2.69 \pm 3.78$	$2.28 \pm 7.31$

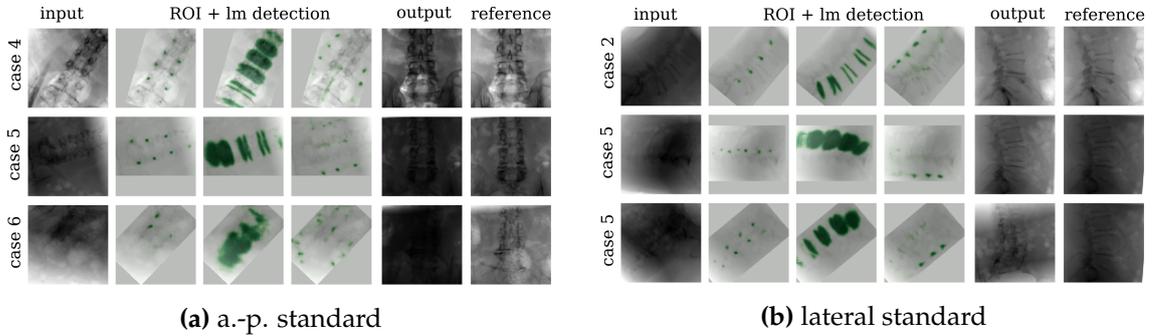
### 5.2.4 Simulated experiments

**RQ 2.5:** Can the 2-stage intensity-based pose regression (Kausch et al., 2020) be replaced by a sequential 1-stage approach without decreasing robustness and accuracy?

The accuracy and robustness of the baseline, 2-step and 1-step approach (Sec. 4.2.4, 4.2.5, 4.2.6) for automatic C-arm positioning from one initial X-ray were compared on synthetic X-rays. In total, 3969 DRRs simulated from the 9 CTs previously assigned to the test set were employed for evaluation. The performances of the different approaches are compared in Fig. 5.17 separately for the a.-p. and lateral standard. The test examples covered an initial pose offset distribution of  $d\theta = 23.8^\circ \pm 8.8^\circ$  around the a.-p. standard and lateral standard, respectively. The baseline approach corresponds to the first step of the 2-step



**Figure 5.17: Quantitative evaluation of proposed pose regression approach on simulated data.** Comparison of the proposed method to the baseline approach and the 2-step direct intensity-based pose regression evaluated on 3969 DRRs of 9 CTs. The proposed method outperforms Kausch et al. (2020) already in a single prediction step. Fig. 5.17a adapted and extended from Kausch et al. (2021b).

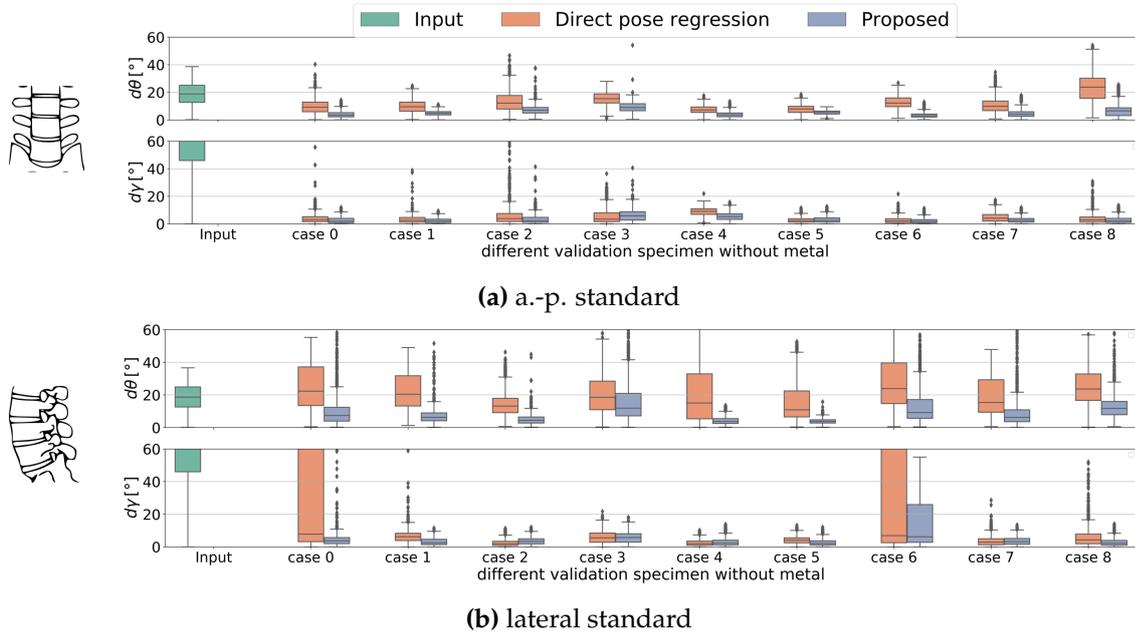


**Figure 5.18: Qualitative evaluation of proposed pose regression approach on simulated data.** The initial X-ray is shown along with the predicted ROI and landmark detection for examples with best, average, and worst performance (top-to-bottom). The predicted pose update is verified by forward projection and shown side-by-side with the reference standard.

approach. Starting from one initial projection, the mean absolute pose error to the desired standard pose is iteratively reduced across both anatomy-specific standard projections ( $d\theta_{a.-p.}^{(1)} = 7.2^\circ \pm 4.6^\circ$ ,  $d\theta_{a.-p.}^{(2)} = 5.0^\circ \pm 4.0^\circ$ ,  $d\theta_{lateral}^{(1)} = 9.9^\circ \pm 8.7^\circ$ ,  $d\theta_{lateral}^{(2)} = 9.6^\circ \pm 8.7^\circ$ ). The proposed 1-step approach performed even more accurate and robust, while only requiring one prediction step ( $d\theta_{a.-p.} = 3.5^\circ \pm 2.5^\circ$ ,  $d\theta_{lateral} = 4.8^\circ \pm 6.0^\circ$ ). Significant accuracy improvement in beam direction of the proposed approach (1-step) compared to the 2-step approach was confirmed by a t-test on  $d\theta$ ,  $d\alpha$ ,  $d\beta$ ,  $d\gamma$  with  $p \ll 0.01$ . Visual examples for best, average and worst accuracy in terms of  $d\theta$  are shown in Fig. 5.18.

In addition to the rotational dof the translation in the detector plane was estimated, resulting in a MAE of  $dtx_{a.-p.} = 2.56 \pm 1.85\text{mm}$ ,  $dtx_{a.-p.} = 3.59 \pm 7.95\text{mm}$

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



**Figure 5.19: Quantitative evaluation of proposed pose regression approach on real X-rays without metal.** The generalization capabilities from synthetic training to validation on real X-rays without metal is assessed. The proposed approach is compared to direct intensity-based pose regression with  $AE_{d\theta}$  and  $AE_{d\gamma}$  distribution given separately for each specimen *case 0 - case 8*. The proposed method shows improved generalization capabilities compared to direct pose regression. Fig. 5.19a adapted and extended from Kausch et al. (2021b).

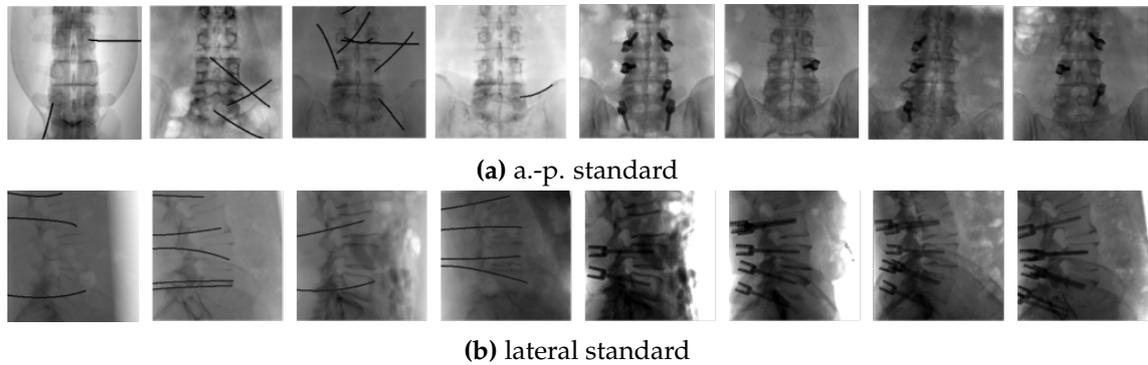
and  $dtx_{lateral} = 3.56 \pm 2.75mm$ ,  $dt\gamma_{a.-p.} = 3.32 \pm 7.28mm$ .

The runtime per X-ray was 0.12 s on average.

### 5.2.5 Specimens experiments

**RQ 2.6:** How does the proposed method solely trained on simulated DRRs generalize to real X-rays without metal acquired in a specimen study?

To assess the transfer capabilities from simulation-based training to real X-rays, validation was performed on 9 specimens without metal for two distinct standard projections (a.-p., lateral). For each specimens and standard, 1364 X-rays were sampled from different poses around the corresponding standard reference pose. An initial pose offset distribution



**Figure 5.20: Exemplary standard DRRs complemented with proposed random k-wire augmentation and realistic screw simulation.**

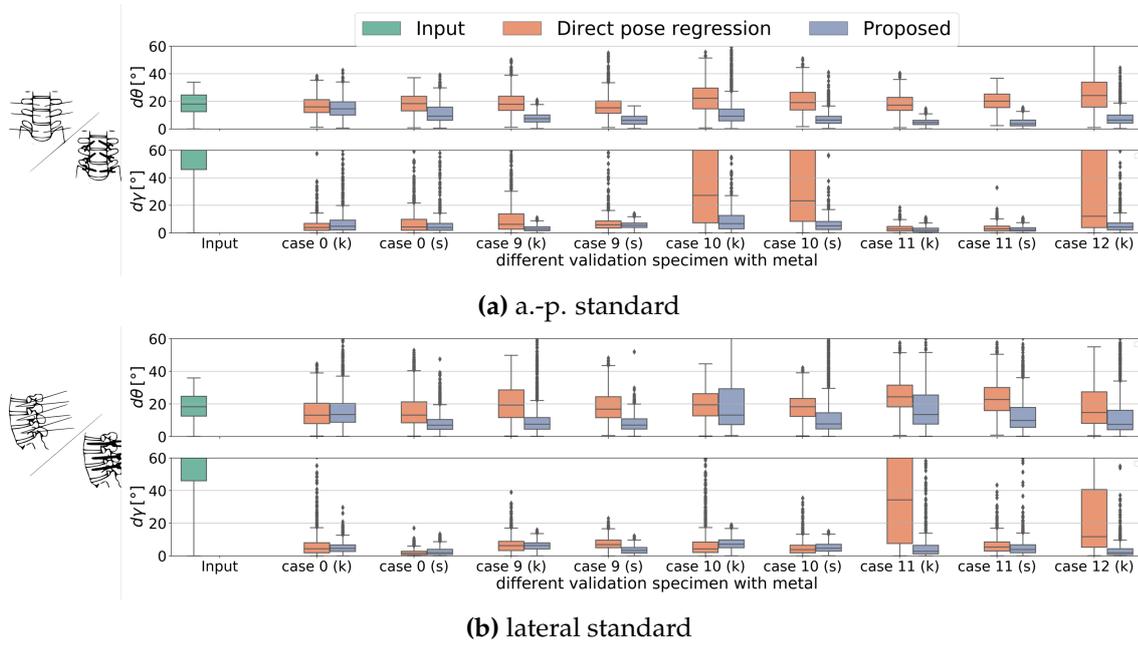
of  $d\theta = 18.3 \pm 7.6^\circ$  was computed. Fig. 5.19 shows that the proposed method outperforms direct intensity-based pose regression (Kausch et al., 2020) across all validation specimens and both standard projections. For the a.-p. standard, 6/9 (*case 0, case 1, case 4, case 5, case 6, case 7*) showed comparable mean performance as the test DRRs, indicating that there is no intrinsic domain gap between synthetic and real X-rays. 2 Specimens (*case 2, case 3*) showed lower performance due to superimposing air pockets resulting in blurred bone boundaries (cf. Fig. 5.13 (a)). *Case 8* showed severe lumbar scoliosis. For the lateral standard without metal, *case 3* and *case 6* show decreased performance, resulting from lateral arm overlays and superposition of air pockets (Fig. 5.13). Fig. 5.22 (top row) gives visual results of initial X-rays, along with their predicted region-of-interest (ROI) and landmarks, and the output X-ray after employing the predicted pose offset for C-arm repositioning. For comparison, the corresponding reference standard is shown on the right. From top to bottom, the selected cases represent good, mean and bad performance examples.

A more detailed performance visualization including all rotational dof is provided in Fig. A.2.

**RQ 2.7:** Does the proposed k-wire and screw simulation techniques improve generalization towards surgical instrumentation?

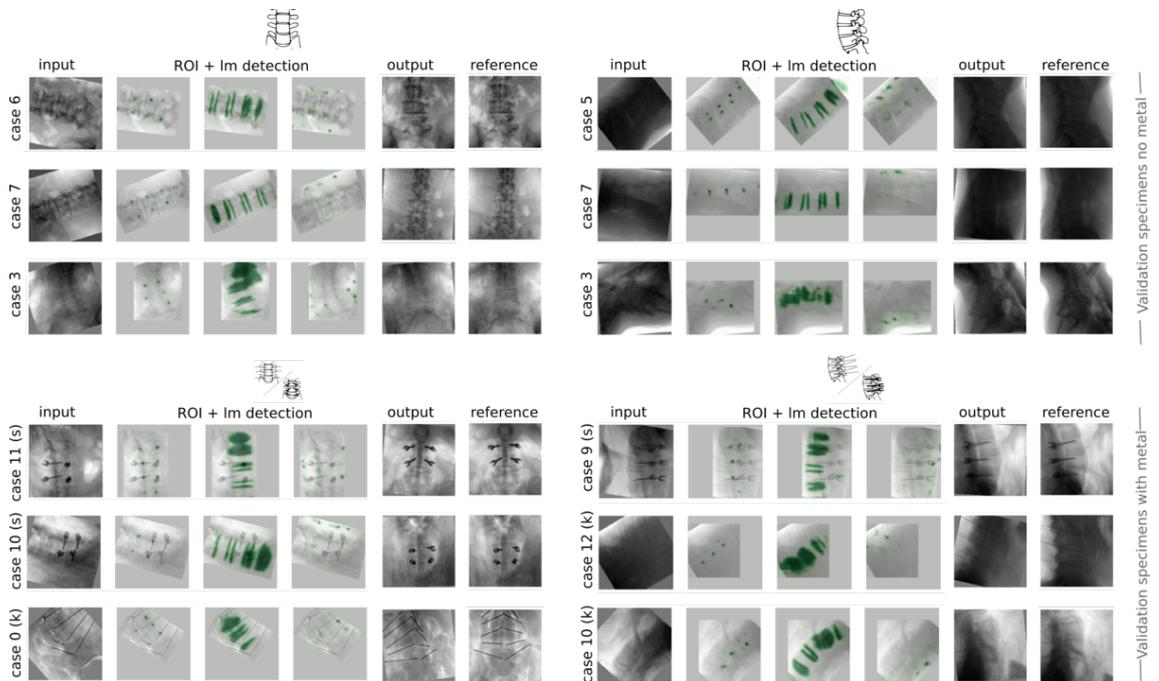
The proposed implant simulation strategies were used to complement the synthetic training data with k-wire and screw. Visual results for the a.-p. and lateral standard projection are given in Fig. 5.20.

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



**Figure 5.21: Quantitative evaluation of proposed pose regression approach on real X-rays with metal.** The domain adaptation capabilities from synthetic training to validation on real X-rays with spinal implants is assessed. The proposed approach is compared to direct intensity-based pose regression with  $AE_{d\theta}$  and  $AE_{d\gamma}$  distribution given separately for each specimen *case 0*, *case 9* - *case 12*, k and s indicating the presence of k-wires or screws. The proposed implant simulation strategies combined with the sequential pose regression shows to improve generalization towards spinal instrumentation.

To investigate if the proposed k-wire augmentation and realistic screw simulation address the domain gap, another specimens study was conducted, including 5 specimens with screws and 4 specimens with k-wires (Sec. 4.2.2). For each specimen, 1364 X-rays were acquired for each standard projection. The validation data was annotated with ground truth pose offsets to the respective standard projection. Applying the proposed approach trained without the suggested implant simulation strategies did not improve upon the initial error distributions. However, the proposed approach with simulated metal augmentation outperforms direct intensity-based pose regression in all specimens for the a.-p. standard and in all except one specimen for the lateral standard (Fig. 5.21). Evaluation on specimens *case 9*, *case 11*, *case 12* with k-wires and screws indicate that the proposed 1-step approach combined with the implant simulation strategies successfully address the domain gap regarding the a.-p. standard. Specimens *case 0*, *case 10* show reduced accuracy, which can be explained by projection artifacts resulting from the body

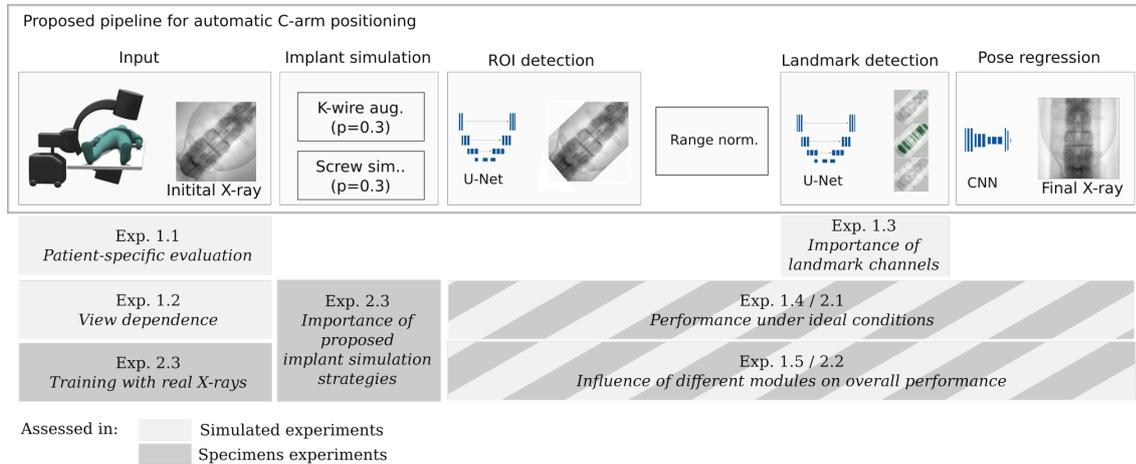


**Figure 5.22: Qualitative evaluation of proposed pose regression approach on real X-rays.** The initial X-ray is shown along with the predicted ROI and landmark detection for cases with good, mean, and bad performance (top-to-bottom). The predicted output is sampled from the acquisition sequence and shown side-by-side with the reference standard.

bag (*case 0*) and edge overlays resulting from corpse decay air pockets (*case 10*) (Fig. 5.13). For the lateral standard, the strongest performance is shown for the *case 9* with k-wire and screws. The other cases show limited performance that can be related to specific scenarios not represented during training, including cadaver bag artifacts (*case 0*), corpse decay air pockets (*case 10*), lateral arm superpositions (*case 11*). In most cases, the in-plane rotation  $\gamma$  was accurately estimated. For specific cases, outliers corresponding to a  $180^\circ$  shift were observed, which can be explained by the anatomical ambiguity within the spine height axis if the hip is not inside the field-of-view (FoV).

For both spinal standard projections and both validations without metal vs. with spinal implants, projections resulting from the predicted pose updates are visualized in Fig. 5.22. Exemplary cases were selected for good, mean, and bad performance. The method showed robust performance even for challenging anatomical deformations (*case 6*). Bad performance was observed for low contrast X-rays due to shadowing by intra-abdominal air (*case 3*) or strike artifacts from the corpse bag (*case 0*), leading to inaccurate

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery



**Figure 5.23: Overview of analyzed influencing factors on simulated and specimens experiments.**

ROI localization and landmark detection.

### 5.2.6 Systematic evaluation of influencing factors

Several factors can influence the final performance of pose estimation. To determine these factors, a variety of different experiments was conducted. For the simulated experiments, this included

**Exp 1.1 Patient-specific evaluation:** The proposed method for automatic C-arm positioning was evaluated separately for all DRRs simulated from one CT to determine whether the outlier cases can be attributed to specific patients.

**Exp 1.2 View dependence:** To analyze the behavior of the algorithm in different offset regimes, i.e., far from and close to the desired standard view.

**Exp 1.3 Importance of landmark channels:** The importance of the three different landmark channels (endplates, pedicles, endplates) was assessed by removing one channel successively during inference and comparing the overall performance in terms of pose accuracy.

**Exp 1.4 Performance under ideal conditions:** To determine the influence of individual sequential processing steps, the ROI detection module and the view-independent landmark detection module were subsequently replaced with ground truth annotations.

**Exp 1.5 Influence of different modules on overall performance:** Different variants of the proposed pipeline were compared with respect to their pose regression performance. Thereby, the effect of each design step on the overall performance was assessed. Results on synthetic X-rays are presented along with real X-rays in **Exp. 2.2**.

For the specimens experiments, the following experiments were conducted

**Exp 2.1 Performance under ideal conditions:** To assess the performance in case of ideal landmark predictions, the selected landmarks were manually annotated in all specimens, and 2D annotations corresponding to the acquired X-rays were automatically derived by forward projection. The derived annotations were used for subsequent pose regression.

**Exp 2.2 Influence of different modules on overall performance:** The effect of modular design choices on the pose regression performance was assessed.

**Exp 2.3 Training with real X-rays:** Training was resumed on a subset of real X-rays covering all validation specimens to derive an upper upper bound accuracy bound. A leave-one-out study on the validation specimens was conducted to determine if a domain gap between simulated X-ray and real X-ray training influences the performance.

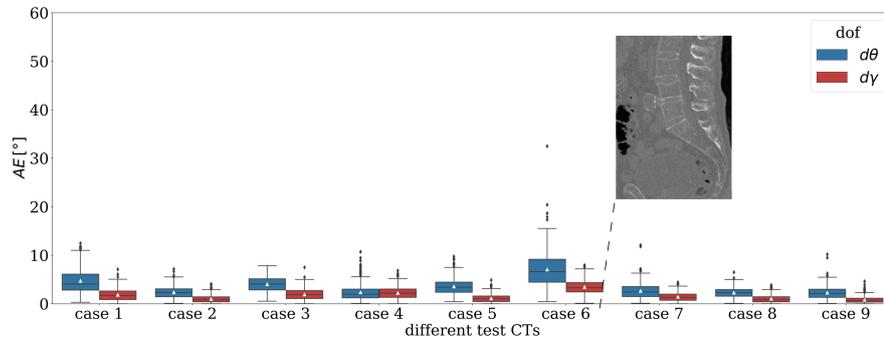
**Exp. 2.4 Importance of proposed implant simulation strategies:** The proposed 2D k-wire augmentation and 3D screw simulation strategies were omitted to measure their generalization effect toward real X-rays with spinal implants.

**Exp. 1.1 Patient-specific evaluation** The patient-specific evaluation in Fig. 5.24 illustrates that outliers can be attributed DRRs simulated from specific CT volumes. For the a.-p. standard this affects the *case 6* CT, where intra-abdominal air shadows the vertebrae in projection domain (cf. Fig. 5.25). For the lateral standard *case 5* and *case 6* show an increased amount of outliers. For the test CT volume *case 6* part of the arm and fingers are present at the lateral front, and for *case 5* lateral intra-abdominal air was observed.

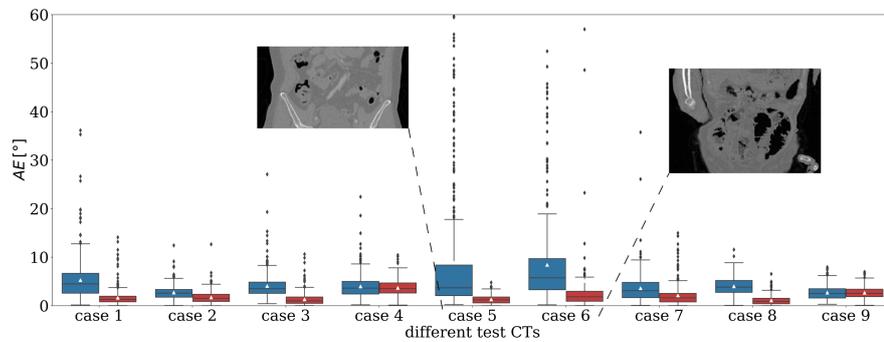
**Exp. 1.2 View dependence** Fig. 5.26 illustrates that the residual pose error is not proportional to the initial offset from the desired standard pose, neither for the angular rotational offset  $d\alpha$ , nor for the orbital rotational offset  $d\beta$ . This indicates that an iterative prediction would not improve upon the 1-step predictions, which was experimentally confirmed.

**Exp. 1.3 Importance of landmark channels** Fig. 5.27 visualizes the contribution of each landmark channel (pedicle, endplates, endpoints) on the overall performance. Removing

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery

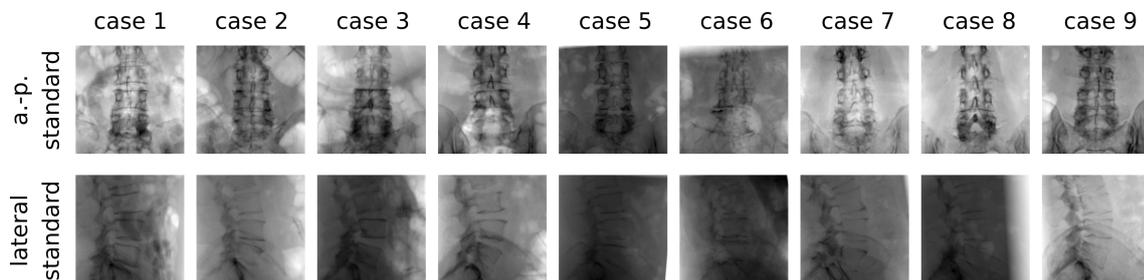


(a) a-p. standard

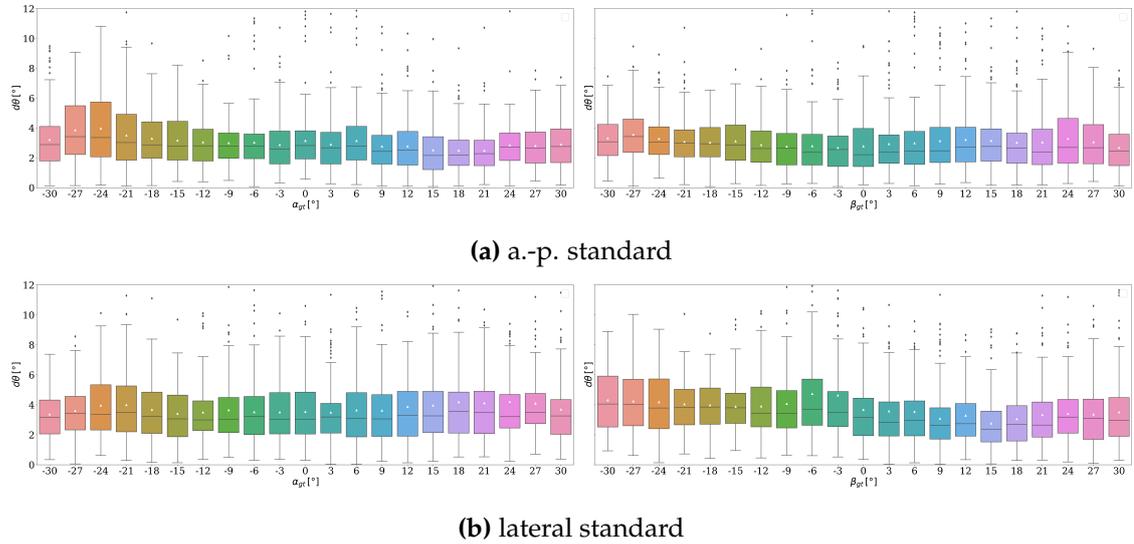


(b) lateral standard

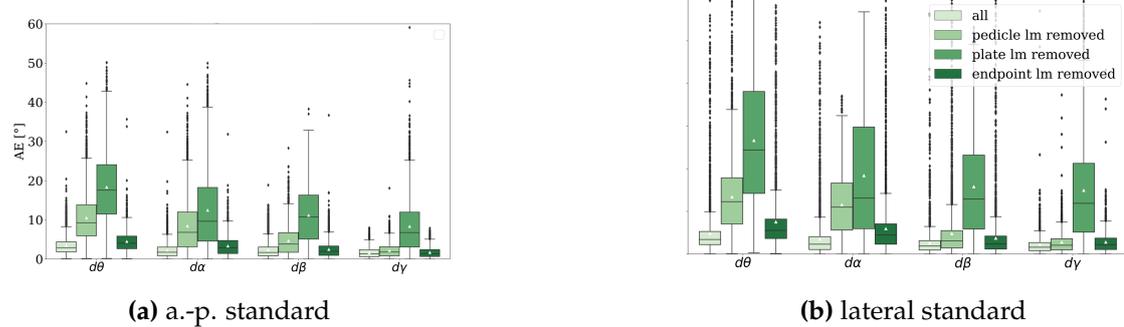
**Figure 5.24: Patient-specific evaluation of proposed approach on test DRRs.** For each patient, 1394 DRRs were generated with poses equidistantly sampled from  $\alpha, \beta \in [-30^\circ, 30^\circ]$ . Outliers can be attributed to specific CTs.



**Figure 5.25: Simulated reference standards for all test CTs.** A-p. and lateral standard are shown top-to-bottom.



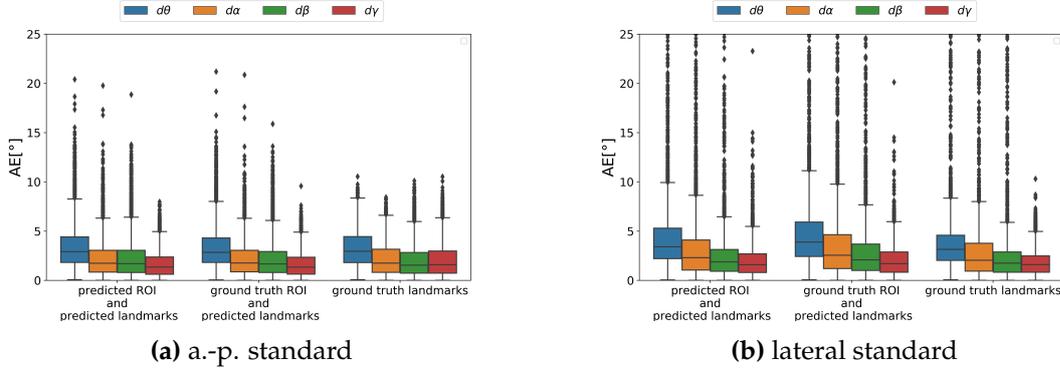
**Figure 5.26: View dependence of proposed 1-step approach on test DRRs.** From each CT, 441 DRRs were generated with pose offsets equidistantly sampled from  $\alpha, \beta \in [-30^\circ, 30^\circ]$  around the desired standard.



**Figure 5.27: Importance of landmark channels for pose regression performance on test DRRs.** One landmark channel is removed at a time during inference to analyze its contribution.

the endplates results in the highest performance drop, followed by the pedicles, whereas the endpoints of the vertebrae processes have the least influence. Best results are obtained by a combination of all three proposed landmark channels.

**Exp 1.4: Performance under ideal conditions** The effect of the performance of individual pipeline steps on the overall pose regression performance was analyzed. Individual

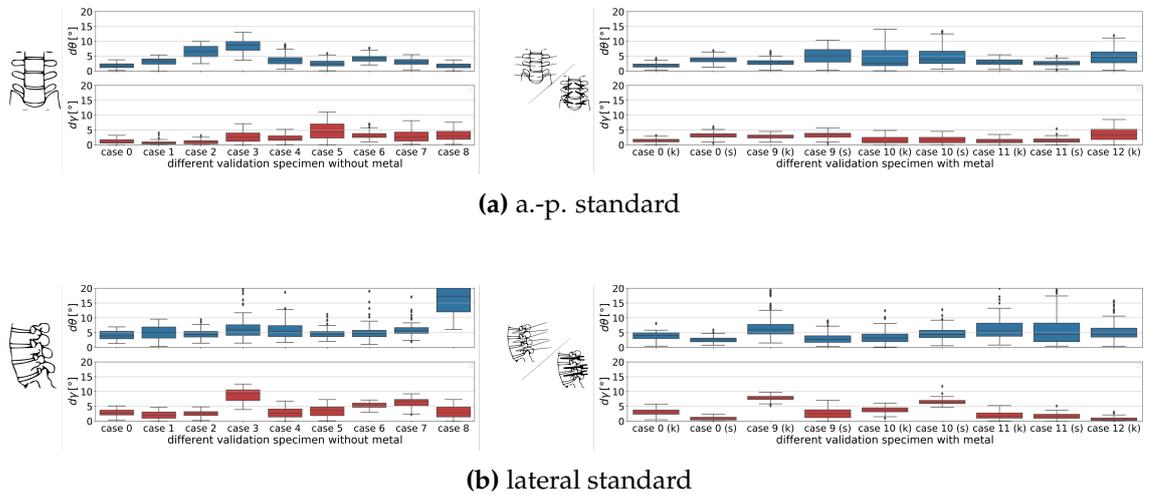


**Figure 5.28: Performance under ideal conditions evaluated on simulated data.** Comparison of fully automatic proposed approach to performance with ground truth landmarks.

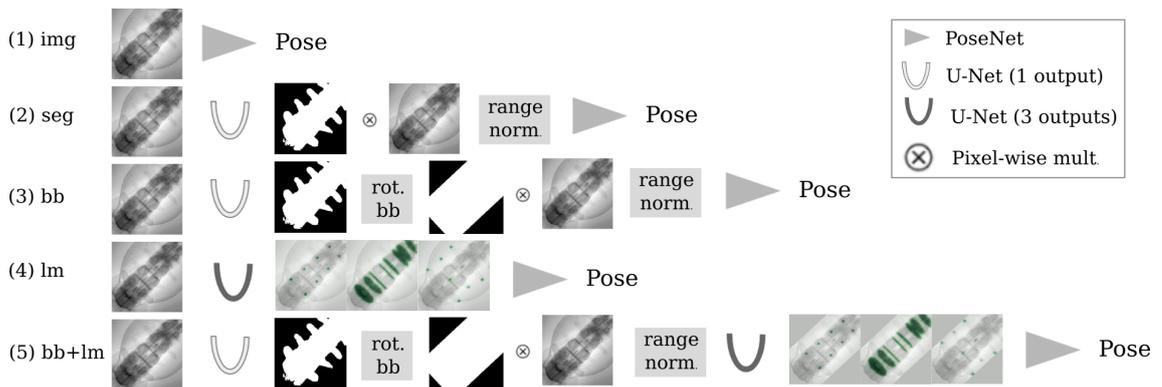
pipeline steps (ROI localization, landmark detection) were replaced with ground truth annotations. Fig. 5.28 shows a performance comparison on simulated data. The fully automatic pipeline with predicted landmarks ( $d\theta_{a.-p.} = 3.48 \pm 2.48^\circ$ ,  $d\theta_{lateral} = 4.83 \pm 6.00^\circ$ ) almost reaches the upper accuracy bound given by the pose regression from ground truth landmark annotations ( $d\theta_{a.-p.} = 3.26 \pm 1.83^\circ$ ,  $d\theta_{lateral} = 4.36 \pm 6.27^\circ$ ). The ROI detection module can replace ground truth annotations without performance degradation. After excluding the two outlier cases (Exp. 1.1) both standard projections show similar accuracy and robustness ( $d\theta_{a.-p.} = 2.96 \pm 1.89^\circ$ ,  $d\theta_{lateral} = 3.74 \pm 2.63^\circ$ ) with the automatic approach.

**Exp 2.1: Performance under ideal conditions** Fig. 5.29 illustrates the pose regression accuracy with ground truth landmark annotations which defines an upper limit for the reachable accuracy using the proposed approach. Under ideal conditions, the proposed method reaches acceptable accuracies in line with the derived clinical accuracy bound ( $d\theta = 8.89 \pm 5.77^\circ$ ) for all specimens with and without metal regarding the a.-p. standard. For the lateral standard, the derived clinical accuracy bound ( $d\theta = 6.36 \pm 3.71^\circ$ ) was reached by all specimens except *case 8*, which showed severe lumbar scoliosis.

**Exp. 2.2: Influence of different modules on overall performance** Five C-arm pose regression pipeline variants integrating different modules of task-specific prior knowledge derived from clinical experience into the pipeline were compared to justify the design of the proposed method. The modules were designed to reflect and automate approaches used by spinal neurosurgeons for identifying correct standard projections. An overview is given in Fig. 5.30. The variant *Image (img)* corresponds to the direct pose regression and the variant *Bounding box + Landmark (bb+lm)* to the proposed approach. The other variants are



**Figure 5.29: Performance under ideal conditions evaluated on real X-rays. 1-step pose regression.** Pose regression is computed based on manual reference landmark annotations. For each specimen,  $AE_{d\theta} \pm SD_{d\theta}$ ,  $AE_{d\gamma} \pm SD_{d\gamma}$  are indicated, k and s indicating the presences of k-wires or screws. Given ground truth landmarks, the pose regressor shows robust performance across different specimens and standard projections.



**Figure 5.30: Variants of the pose regression pipeline** integrating different stages of task-specific expert knowledge derived from clinical routine: (1) *img*, (2) *seg*, (3) *bb*, (4) *lm*, (5) *bb+lm*.

outlined in the following:

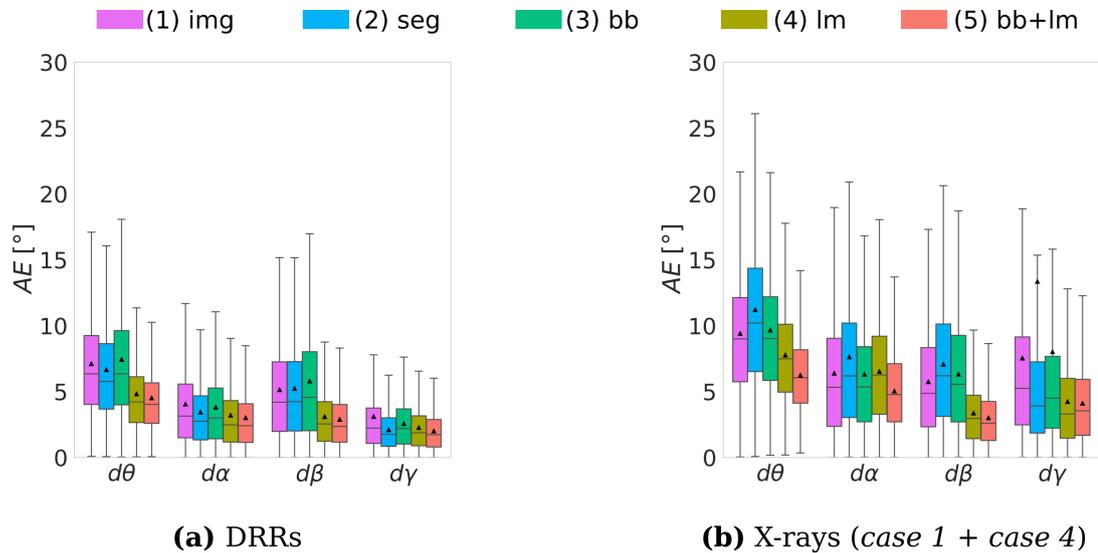
(2) *Segmentation (seg)*: The variant (*seg*) masks the initial X-rays with a vertebra segmentation mask to restrict the focus on bone structures.

(3) *Bounding box (bb)*: The previous approach was generalized in the variant (*bb*) by predicting a bounding box mask instead of a segmentation mask. This reflects the use of collimators employed in real clinical procedures to increase image contrast. The rotating bounding box can be computed from the predicted vertebra segmentation. The initial X-ray was multiplied pixel-wise by the estimated rotating bounding box and passed as input to the pose regression network.

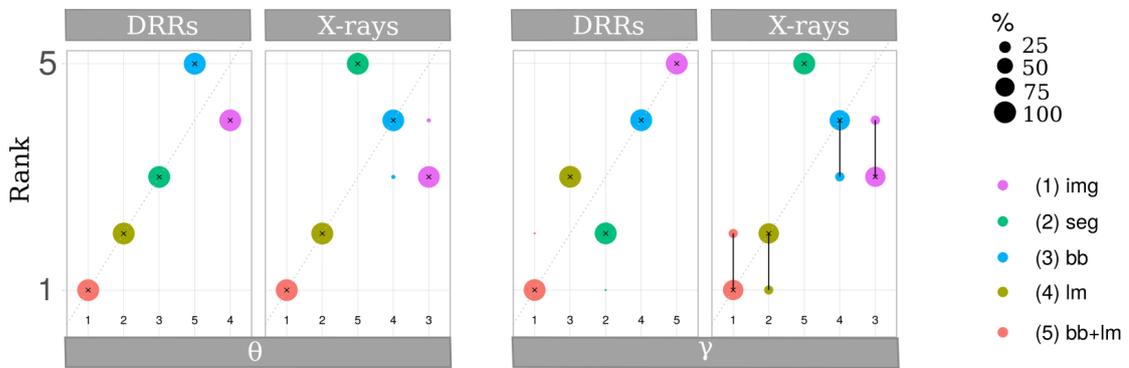
(4) *Landmark (lm)*: Anatomy-specific relevant landmarks were directly derived from the initial X-ray by a view-independent U-Net detection module (Klein et al., 2019) and predicted heatmaps were used for subsequent pose regression.

On the example of the a.-p. standard, the models were trained without implant simulation strategies and compared with respect to their performance on the simulated test set and real X-rays derived from two exemplary validation specimens. Quantitative results compared across the different variants are shown in Fig. 5.31. The MAE is reduced by the landmark-based variants compared to the other purely image-based variants in all pose parameters across all validation data. Best results were observed for the proposed method (*bb+lm*). A systematic robustness analysis using the *ChallengeR* tool developed by Wiesenfarth et al. (2021) confirms this finding; it can be seen in Fig. 5.32 that the proposed method (*bb+lm*) achieves the first rank across all datasets.

**Exp. 2.3: Training with real X-rays** It was analyzed if retraining of the derived model on real X-rays benefits the performance. Therefore, training was resumed on a subset of 137 X-rays equidistantly sampled from the acquired validation (1379 X-rays per specimen). The model was evaluated on the views not seen during training. Results are shown in Fig. 5.33 and can serve as an upper upper bound. Similar accuracies are achieved for both standard projections, indicating that the selected landmarks are equally suitable. However, the anatomy and close by views were present during training, consequently memorizing cannot be excluded. To verify if training on real X-rays improves performance compared to simulated training, a leave-one-out experiment was performed on the acquired validation data. One specimen used for testing was excluded during training, while the annotated X-rays of the remaining specimens were integrated into the training pipeline besides the synthetic X-rays. The observed results did not improve performance on individual test cases upon simulated training (Fig. 5.19).

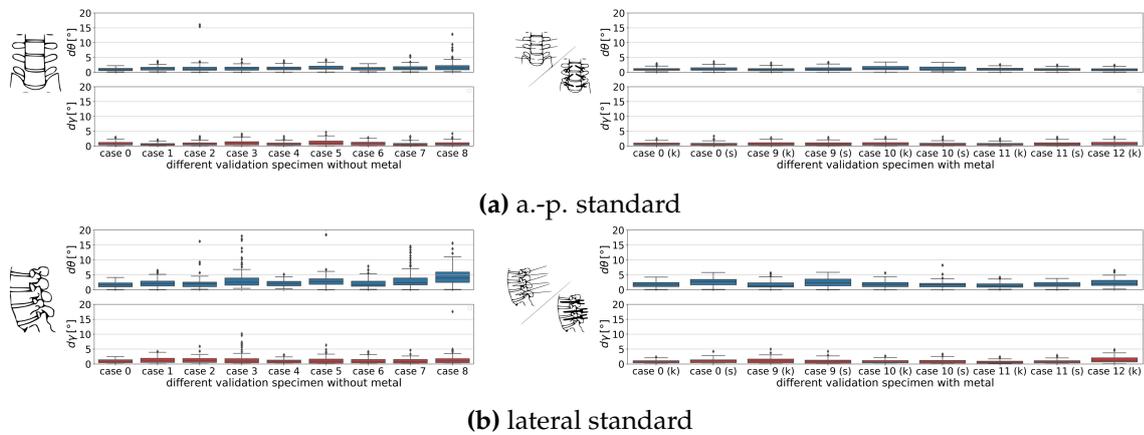


**Figure 5.31: Quantitative comparison of C-arm pose regression pipeline variants trained on simulated X-rays and evaluated on different test sets: (a) test DRRs, (b) real X-rays.**



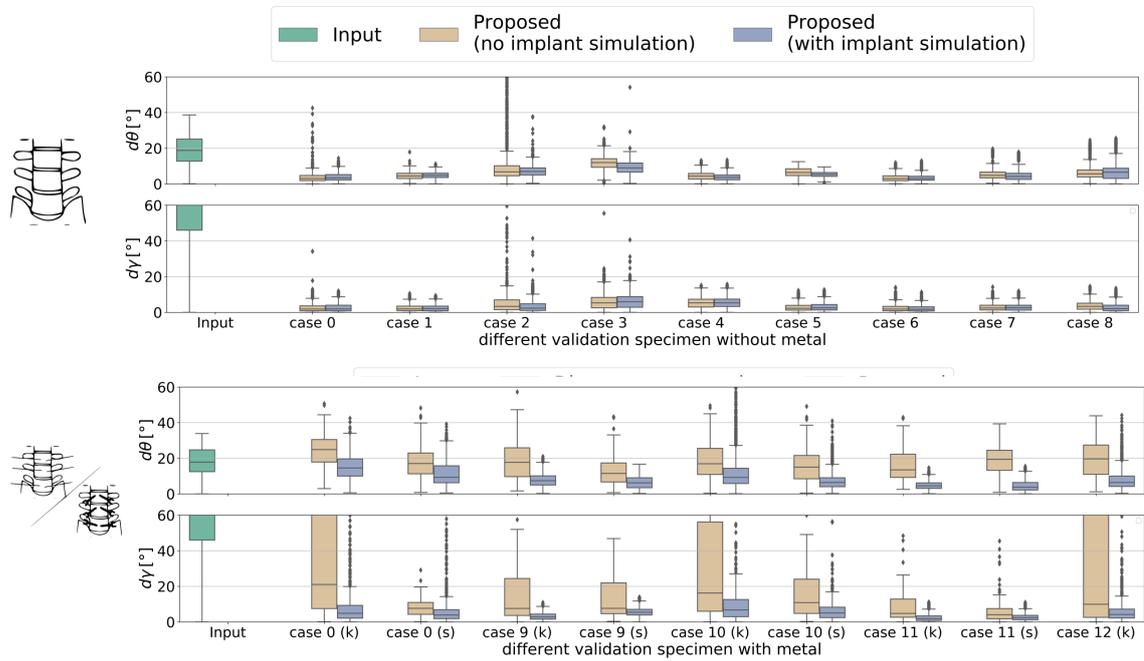
**Figure 5.32: Ranking stability of proposed variants** when applied on different validation datasets (test DRRs, real X-rays). Here the rank of each variant on each dataset (1: best, 5: worse) is measured with respect to the accuracy in viewing direction  $\theta$  and in-plane rotation  $\gamma$  generated by Wiesenfarth et al. (2021). Each variant is color-coded and the area of each blob ( $A_i; r_j$ ) is proportional to the relative frequency each variant  $A_i$  achieved rank  $j$  for 1000 bootstrap samples across the different validation datasets. The median rank of each variant is indicated by a black cross and 95% bootstrap confidence intervals across bootstrap samples are indicated by black lines.

## 5.2. C-arm positioning for fluoroscopy-guided spinal surgery

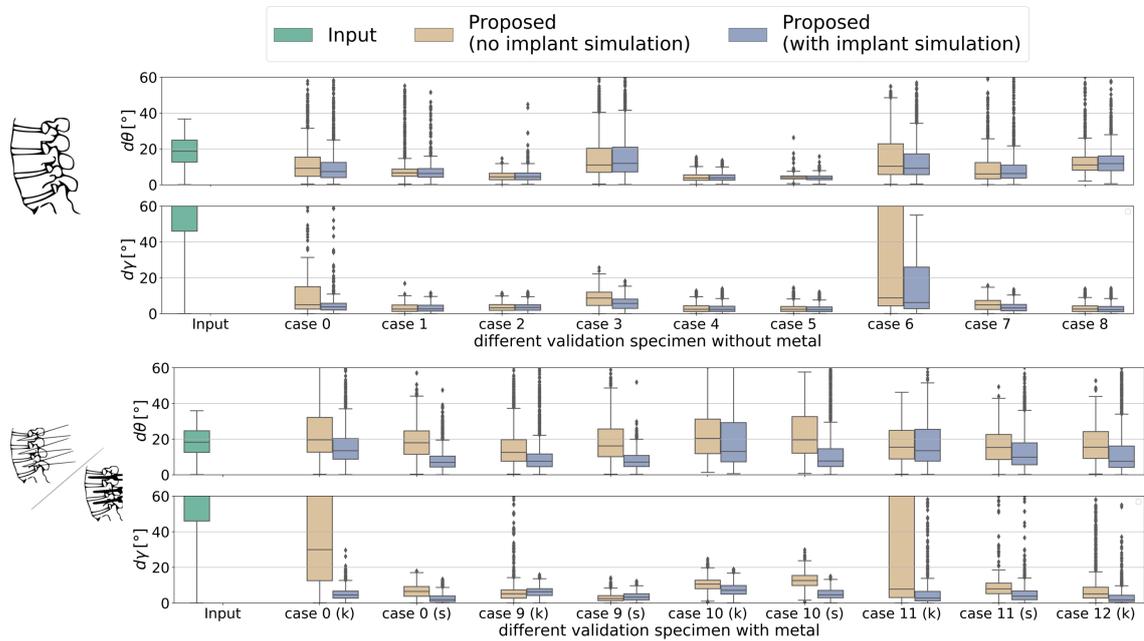


**Figure 5.33: Performance of 1-step pose regression after training on validation specimens.** The simulated training data was complemented with 137 X-rays equidistantly sampled from each validation specimen. The model was evaluated on views not seen during training. However, the anatomy and close by views were present during training.

**Exp. 2.3: Importance of proposed implant simulation strategies** The proposed approach was trained without and with the proposed implant simulation strategies (k-wire/screw simulation). Quantitative results are shown in Fig. 5.34 for the a-p. and lateral standard. The performance on real X-rays without metal is not largely affected for the majority of cases by the integration of implant simulation strategies. In contrast, for the real X-rays with metal accuracy and robustness is significantly improved for most cases.



(a) a.-p. standard



(b) lateral standard

**Figure 5.34: Importance of proposed implant simulation strategies.** The pipeline was trained without and with the proposed k-wire and screw simulation.

The discussion is structurally aligned with the previous chapters: The first part discusses the results concerning pedicle screw planning for CT-navigated spinal surgery (Section 6.1) and the second part discusses the results regarding C-arm positioning for fluoroscopy-guided spinal surgery (Section 6.2).

## 6.1 Pedicle screw planning for CT-navigated spinal surgery

This chapter discusses the proposed method for automatic pedicle screw planning with respect to the performed experiments and results (Sec. 5.1). Sec. 6.1.1 interprets the planning variabilities observed with the current manual screw planning procedure. Sec. 6.1.2 discusses the proposed automatic pedicle screw planning approach. It is divided into five discussion topics: (1) *training data* discusses the influence of learning from surgical planning data, (2) *quantitative results* discusses the automatic planning results in relation to computed manual screw planning variations, *qualitative results* discusses the clinical acceptability of the automatic planning results in comparison to the state of the art, (4) *comparison to related work* places the proposed deep learning-based approach in the state of the art, (5) *postoperative screw accuracy* discusses the observed deviations between intraoperative screw plans and actual postoperative screw locations. Sec. 6.1.3 assesses the proposed approach with respect to the requirements stated in the objectives. Limitations and future work directions are discussed in Sec. 6.1.4.

A deep learning-based approach for automated pedicle screw planning in intraoperative computed tomography (CT) was proposed and evaluated for navigated lumbosacral instrumentation. Instead of relying on manual defined geometric and structural components (Knez et al., 2016; Xiaozhao et al., 2016), explicitly defined constraints (Knez et al., 2018; Li et al., 2019), or atlas-based modeling (Goerres et al., 2017; Vijayan et al., 2019), deep learning techniques were employed to implicitly learn the screw placement from a diverse surgical planning dataset (155 cases, 1052 screws) in a patch-based approach; thereby introducing robustness to structural variations frequently observed in degenerative spine or scoliosis. Building upon the current successes of the U-Net architecture for segmentation tasks (Ronneberger et al., 2015), the screw planning task was interpreted as a segmentation task, assuming that although the intraoperative CT does not contain

the screws yet, their potential location can be deduced based on the spatial context. The proposed instance-based screw segmentation combined with a manual or automatic vertebra localization results in a semi- or fully automatic pipeline for pedicle screw planning. Evaluated on 24 CT covering the lumbosacral levels, the proposed method demonstrated clinically acceptable screw plans in 96.2% of screw trajectories (125/130 screws) that were entirely positioned within the pedicle (Gertzbein-Robbins Grade A). The 5 revisions were caused by proximal facet violations affecting the fifth lumbar vertebra (L5). In comparison to manual planning, time efficiency was improved by 8-fold on average.

### 6.1.1 Manual pedicle screw planning

The use of navigation systems and robotic assistant devices in spinal surgery is increasing (Rawicki et al., 2021) due to the improved accuracies and reduced revision rates of navigated pedicle screw placement compared to fluoroscopy, which makes the technique particularly valuable for complex surgical cases (Luther et al., 2015). Pedicle screw planning is a prerequisite necessary task to achieve the full potential of computer-assisted navigation. The planning with available software tools on the navigation workstations involves manual interaction, is time-consuming, and is expert-dependent. A gold standard for screw placement does not exist, and every surgeon develops their preferences within the anatomical constraints and screw placement criteria. The expert-dependence of planning results was confirmed by the inter-/intra-rater analysis of manual screw planning variability (Sec. 5.1.1), where significantly higher inter-rater than intra-rater deviations were observed across all screw parameters. The highest manual screw planning deviations were observed for the first sacral vertebra (S1) level. This reflects the increased freedom for screw positioning in the sacral segment caused by the increased vertebral space and reduced anatomical constraint compared to the lumbar segment (Vijayan et al., 2019). Manual screw planning for one CT took on average 5 *min* 38 *s*.

### 6.1.2 Automatic pedicle screw planning

The proposed automatic pedicle screw planning method was evaluated on an external test set (45 cases, 328 screws). Since a gold standard for pedicle screw placement does not exist, manual and automatic screw plans were additionally set into context to inter- and intra-rater variabilities observed after manual planning. Regarding the lumbosacral vertebra level, a high agreement between manual and automatic planning results was demonstrated quantitatively and qualitatively, indicating clinically acceptable results for a large majority of cases comparable to the human error rate. The inter-/intra-rater variability study confirmed that automatically generated plans were non-inferior to manual plans. Furthermore, the proposed automation led to a significant speed-up of 8-fold compared

to manual pedicle screw planning (4.9 s per vertebra level). Integration of the proposed approach into navigation systems can assist operators, facilitate screw planning, reduce errors, standardize the quality, and reduce the procedural time.

**Training data** The training data resulted from real-world surgical planning data. The data was retrospectively derived from a registry of CT-guided lumbosacral instrumentations. The screw plans were created intraoperatively during the respective procedure. Using existing clinical annotations made a large, heterogeneous dataset available. However, screw plans did not represent classic ground truth. Instead, it took individual anatomical conditions into account, as well as the surgeon's preference. The proposed approach was trained on 155 CTs containing 1052 intraoperatively manually labeled screws. As shown in Fig. 4.1a the training dataset was not balanced across different vertebra levels. Unbalanced training data may lead to non-optimal results for underrepresented classes (L1: 1.1%, L2: 8.4%, L3: 20.0%, L4: 26.6%, L5: 27.4%, S1: 16.5%). Test data for the first lumbar vertebra (L1) level was not available. Regarding the lumbar segment, increased deviations were observed from cranial to caudal (Fig. 5.2) while available training data increased. This contrary effect indicates that unbalanced training data does not bias the performance within the lumbar segment, where vertebra shapes are very similar. However, training data balancing becomes more relevant when extending the approach to other vertebra segments, e.g., thoracic or cervical. There, a higher inter-segment than intralumbar vertebra shape variation (Schünke et al., 2005) can bias the performance. For the S1 level, increased quantitative deviations were observed compared to the lumbar segment, but those differences were still comparable to inter-rater variances. The observed increased deviations can be attributed to the anatomical underrepresentation of the S1 level compared to third lumbar vertebra (L3)-L5 and increased anatomical freedoms for trajectory planning in the sacrum. Its clinical acceptability is discussed in more detail in the qualitative results.

**Quantitative results** First, the automatic screw planning performance will be discussed based on the whole test set covering the lumbosacral segment (45 cases, 328 screws). Then, the computed automatic variances are additionally set into context to observed variations after manual screw planning based on a randomly selected subset of 24 cases, 130 screws.

For evaluation of the qualitative results, it has to be considered that reference data resulted from surgical planning data and a gold standard for pedicle screw placement does not exist. Of clinical interest were the screw *head* and *axis* point along with the screw *direction* where a high agreement of manual and automatic plans was demonstrated quantitatively (Fig. 5.2). Robust performance was observed on a variety of different

patient anatomies ( $head=4.1 \pm 2.0 \text{ mm}$ ,  $axis=4.2 \pm 2.2 \text{ mm}$ ,  $direction=5.9 \pm 3.8^\circ$ ). Deviations of  $4.1 \pm 3.1 \text{ mm}$ ,  $0.5 \pm 0.3 \text{ mm}$  were computed for the screw *length* and *diameter*. In clinical practice, these parameters are additionally constrained by the screw sizes provided by the implant manufacturers.  $5 \text{ mm}$  in screw *length* and  $1 \text{ mm}$  in screw *diameter* reflect common implant increments. Setting the observed deviations of *length* and *diameter* into context, observed mean deviations fall below this threshold and would not have caused a different implant selection in terms of *diameter*. The suitable implant *length* can be chosen at the discretion of the surgeon within these bounds without clinical relevance. The vertebra level-specific evaluation illustrates higher variabilities for the S1 vertebra, reflecting more options for screw placement regarding screw convergence and cranial/caudal angulation.

While analyzing the quantitative screw positioning differences between manual and automatic planning, the acceptable planning variability of pedicle screw has to be considered. For quantitative analysis of the clinical impact of deviations, automatic vs. manual screw plan deviations were compared to intra- and inter-rater manual screw plan variations. Due to different surgical planning preferences within the guidelines for pedicle screw planning, intra-rater variability is lower than inter-rater variability in clinical practice. Since training data contained the screw planning preferences of several experts, intra-rater deviations are expected to be lower than automatic vs. manual deviations. This was confirmed by our experiments across all screw parameters (Tab. 5.1). The deviations of automatic vs. manual plans were comparable to manual inter-rater variances, indicating non-inferiority of automatic planning compared to manual planning. The variations of screw parameters *head* and *direction* were comparable to manual planning and *axis* and *overlap* significantly less than manual planning. The results suggest that automatic planning shows sufficient accuracy and could replace the current manual procedure performed by an expert while drastically improving the time efficiency by 8-fold (Fig. 5.7). The clinical acceptability is further discussed below in the qualitative results.

The proposed automatic screw planning took on average  $41.8 \pm 20.2 \text{ s}$  per CT comprising 2-10 trajectories per patient, in contrast to  $338.7 \pm 186.3 \text{ s}$  by manual expert planning. Screw planning for one vertebra level took  $4.91 \pm 0.11 \text{ s}$  with the proposed approach. In comparison, state of the art approaches based on 3D geometrical and structural properties reported computations times of  $2 \text{ mins}$  for vertebral structure modeling plus  $20 \text{ s}$  planning time for each screw trajectory (Knez et al., 2016). Atlas-based approaches (Vijayan et al., 2019) reported runtimes of  $312.1 \pm 104.0 \text{ s}$  for planning five-level bilateral screws and  $49.4 \pm 39.9 \text{ s}$  per vertebra-level, where the greatest proportion can be attributed to the shape model registration ( $41.2 \pm 39.9 \text{ s}$ ). This comparison indicates significantly speed-up benefits of the proposed patch-based deep learning approach. The proposed approach

shows robust performance towards poor image quality and anatomical variations like scoliosis (Fig. 5.3 (C)) indicating that the patch-based design choice combined with the collected training data and augmentation strategies is able to cover the necessary anatomical and image quality variations.

The screw *length* and *diameter* were predicted significantly shorter and thinner by the automatic approach compared to manual expert annotations, resulting in an average offset of 3.3 mm in length and 0.3 mm in width. Since these offsets are below the threshold of commonly used implant increments, the clinical impact is insignificant.

The vertebra-level specific Dice evaluation revealed significantly increased screw variations for the S1 level compared to the lumbar segment for manual planning and automatic planning, while overlaps within the lumbar segments were similar. More significant screw placement variations for S1 result from fewer anatomical constraints within the sacrum. Nonetheless, observed automatic vs. manual planning deviations for S1 agreed with inter-rater variations, indicating clinically acceptable performance also for the sacrum.

**Qualitative results** On the randomly selected subset of 24 cases, 130 screws, the acceptability of the automatically planned screws was clinically assessed by a spinal neurosurgeon using the Gertzbein-Robbins (GR) classification. The GR classification grades the screw placement according to the extent of cortical pedicle breach. 96.2% of the automatically planned screw were classified as GR Grade A and the remaining 3.8% as GR Grade B with minor clinically acceptable breaches. In comparison, state of the art approaches reported GR Grade A in 88.7% with geometrical and structural modeling (Knez et al., 2019) and 86.7% with atlas-based modeling (Vijayan et al., 2019). Though the methods were not evaluated on identical datasets, the results indicate improved robustness of the proposed deep learning-based approach, which directly learns from clinical planning data.

**Comparison to related work** The proposed approach via instance-based screw segmentation was shown to outperform direct convolutional neural network (CNN) landmark regression, indicating that the 3D mask representation contains valuable spatial information for the model. A comparison to other state of the art approaches is challenged by different validation datasets or missing quantitative evaluations. Compared to the approach by Knez et al. (2016), the proposed approach is on average 1 mm more accurate in *length* and 1° more accurate in *direction*, while both approaches differ in vertebra level and test size. Its parametric modeling assumption may limit the application to locally deformed vertebrae. The atlas-based approach (Vijayan et al., 2019) resulted on average

on increased accuracies, which could be attributed to the underlying ideal reference annotation specifically defined by a spinal neurosurgeon. In contrast, the proposed deep learning-based approach relied on intraoperative planning data, which can include variable surgical preferences, case-specific planning adjustments, and might be influenced by time restrictions. In addition, atlas-based approaches are targeted to represent the healthy anatomical shape. They may be limited in performance in altered, deformed, degenerated, or fractured bones, as commonly observed in patients treated with spinal instrumentation. Similarly, Vijayan et al. (2019) reported 9 outlier cases caused by shape model initialization collapsing, which is sensitive towards the initialization. In contrast, the proposed approach showed robust performance in a diverse test set, including anatomical variation, e.g., scoliosis and poor image contrast. Further, a shape model needs to be initialized and trained separately for each vertebra level.

**Postoperative screw accuracy** The postoperative screw accuracy after CT navigation was evaluated by comparing lumbosacral screw trajectories in the postoperative *control* CT to intraoperative screw plans, including 24 cases, 140 screws. Computed deviations ( $head=5.2 \pm 2.4 \text{ mm}$ ,  $axis=5.5 \pm 2.7 \text{ mm}$ ,  $direction=6.3 \pm 3.6 \text{ mm}$ ) are in the same range as automatic vs. manual deviations and surpass the mean registration error of the navigation system. Additionally, the measured screw deviations can be influenced by inaccuracies in the rigid registration between *control* and *navi* CT. Successful registration was verified by manual inspection; however, small offsets can remain caused by anatomical alternations between the intraoperative CT with fractures and without screws and the postoperative CT with repositioned fractures and inserted implants. Due to missing ground truth *navi* to *control* registrations, the rigid registration error could not be assessed.

### 6.1.3 Conclusion

In the current clinical routine, pedicle screw planning is performed manually, which requires a high level of expertise and in-depth knowledge of the anatomy and pedicle orientation. The screw planning task was interpreted as a segmentation task that represents valuable knowledge for the training process, leading to more accurate and robust performance than direct CNN landmark regression. The deep learning-based proposed approach for pedicle screw planning fulfills all the requirements that were stated in the objectives.

**Clinical applicability:** The proposed automatic screw planning tool can be easily integrated into the clinical routine.

**Performance requirements:** The approach was evaluated for the lumbosacral level on an unseen test set, showing robust performance on various different patient anatomies and

leading to clinically acceptable results in 96.2% of screw trajectories.

**Benefits for the patient:** The proposed automatic screw planning led to a significant speed-up of 8-fold compared to manual expert planning, which can decrease the overall procedural time. Further, automatic planning would lead to standardized results independent of the surgical experience.

**Benefits for the surgeon:** The proposed approach can assist the surgeon in pedicle screw planning and thereby standardize the quality and reduce the expertise dependence while significantly reducing the procedural time.

**Representation of anatomical variability:** The approach directly learns from surgical planning data of the lumbosacral segment, representing a variety of patient anatomies and surgical planning preferences.

**Addressing intraoperative requirements:** Intraoperative requirements are implicitly addressed by learning from surgical planning data.

**Generalization:** With available training data, the method can be extended to other vertebra levels.

#### 6.1.4 Limitations and future work

The proposed approach was trained and evaluated for the lumbosacral segment. With available training data, the approach could be transferred to the thoracic and cervical segments. As already mentioned in the training data section, the influence of inter-segment unbalanced training data should be considered. Different techniques for addressing imbalanced training data in deep learning have been proposed and can be clustered into data-driven, algorithm-driven, and hybrid techniques (Krawczyk, 2016). Data-driven approaches address the class imbalance by data sampling strategies, whereas algorithm-driven methods employ class weight balancing and directly modify the learning algorithm, and hybrid approaches combine both. Furthermore, the effect of the fixed patch window size needs to be analyzed for different vertebra segments with significantly smaller vertebra shapes (Schünke et al., 2005).

Further analysis of the proposed approach for automatic screw planning in pathological and degenerated scenarios, e.g., dislocated fractures and severe scoliosis, is required. If necessary, the representation of these challenging anatomical conditions needs to be reinforced in the training data.

Overall, 96.2% of automatically planned screws were rated as clinically acceptable. All revisions (5 screws, 3.8%) were caused by proximal facet violations that constitute a significant risk factor for adjacent segment destabilization (Sakaura et al., 2019). All

facet violations affected the L5 level, which is likely caused by case-specific trajectory modifications present in intraoperative training data. The L5 vertebra level requires steeper insertion trajectories (Fig. 2.4a) which might be compromised at the discretion of the surgeon to facilitate screw insertion. If data-driven learning algorithms are trained on noisy labels, this can result in deteriorated performance. Strategies to address this issue could be investigated in future work. Song et al. (2020) present a survey on robust training techniques, including low-quality annotations. Nevertheless, the rate of facet violations was still in line with those reported in the literature after navigated screw placement (Ohba et al., 2016).

The significant underestimation of length and width could be addressed by adapting the post-processing step of the screw mask to parameter conversion. However, the clinical impact is potentially neglectable since absolute differences fall below standard implant increments.

The proposed automatic patch initialization based on vertebra-instance localization has the drawback that it does not detect vertebrae at the image boundaries because the large scale vertebrae segmentation challenge (VerSe) training data (Löffler et al., 2020) only contained annotations for vertebrae fully contained inside the field-of-view (FoV). Further, nnU-Net inference is not optimized to speed but to accuracy, which is not decisive for the task of vertebra centroid initialization. For speeding up the automatic patch initialization, the instance-based segmentation could be trained at reduced resolution. Alternatively, detection algorithms can increase the speed of the automatic patch initialization. Levine et al. (2019) compared different detection approaches for vertebra centroid localization and reported best performance with Faster R-CNN (Ren et al., 2016) with a runtime of 8.72 s per detection.

In future work, the proposed automatic screw planning tool can be seamlessly integrated into navigation systems. The automatic screw planning tool provides assistance by suggesting automatically derived screw trajectories to the surgeon for final verification. Integration into the clinical workflow allows investigation of its clinical usability and performance further. Observed failure cases can be reported, and manual corrections can be included in the training process to improve the robustness further (Ramadan et al., 2020). In case of performance difficulties, additional geometric constraints could be integrated into the segmentation task by restricting the screw masks to cylindrical shapes (Bohlender et al., 2021). In the proposed approach, the combined Dice and cross-entropy loss (Isensee et al., 2021) has been used for weight optimization. However, the Dice Loss may not be optimal for small structures, where a single pixel difference can already have a

significant influence on the metric (Reinke et al., 2021). Analyzing the effect of different metrics, e.g., focal loss, on the performance can be considered. Furthermore, a systematic comparative study of the proposed approach with commercially available atlas-based software released by Brainlab (Riem, Germany) is in progress. Screw navigation accuracies should be evaluated further to assure the agreement of implanted screws with CT-based screw plans.

## 6.2 C-arm positioning for fluoroscopy-guided spinal surgery

This chapter discusses the proposed method for automatic C-arm positioning with respect to the performed experiments and results (Sec. 5.2). Sec. 6.2.1 interprets and discusses the derived measures for manual C-arm positioning accuracy assessment, along with manual C-arm positioning efficiency. Sec. 6.2.2 discusses the proposed automatic C-arm positioning approach. It is divided into five discussion topics: (1) *training data* discusses the design choices during simulation, (2) *simulated experiments* discusses the results obtained on simulated test data, *validation data* discusses the setup for validation data acquisition, (4) *specimens experiment* discusses the results obtained on real X-rays, (5) *pipeline design* discusses the influence of individual pipeline steps on the overall performance. Sec. 6.2.3 assesses the proposed approach with respect to the requirements stated in the objectives. Limitations and future work directions are discussed in Sec. 6.2.4.

A novel approach for automatic C-arm positioning for spinal standard projections was proposed that seamlessly integrate into the clinical workflow. It estimates the required pose update for C-arm repositioning from one initial X-ray. In contrast to many state of the art approaches, no preoperative computed tomography (CT) scan or other technical equipment is required. Instead of atlas- or model-based representation, deep learning is employed, and robust learning is enabled by the development and combination of various methods that address aspects like data scarcity, intraoperative confounding factors, anatomical variation, including learning from simulations, advanced data augmentation, and integration of task-specific expert knowledge derived from surgical routine. The proposed automatic C-arm positioning method was evaluated for the a.-p. and lateral standard for the fourth lumbar vertebra (L4). It significantly increased accuracy, robustness, and generalization properties compared to direct intensity-based pose regression. A large specimens study was performed to validate the approach on real X-rays covering different intraoperative settings. The method can handle inter-patient anatomical variation and was capable of generalizing from *in silico* data to *ex vivo* human cadaver data.

### 6.2.1 Manual C-arm positioning

Currently, the C-arm is positioned manually, involving repeated X-ray acquisitions, to derive the desired view onto the anatomy for medical verification. The manual positioning process is time-consuming, expert-dependent, and increases the radiation exposure of patients and surgeons. Though guidelines for correct anatomical standard projections exist (Sec. 2.2.1), inherent positioning variations depend on the surgical experience, surgical preference, anatomical conditions. Furthermore, manual positioning requires correct spatial interpretation of 2D projection images with superimposed anatomical structures and missing depth information. The variability of standard projections has not been assessed in a clinical study yet. In Sec. 5.2.1, 5.2.2, 5.2.3 four different accuracy measures were introduced and assessed to measure the acceptable variability of standard projections for setting the automatic positioning results into context. All accuracy measures were assessed on different data sources, which limits the direct comparison. The obtained results are discussed separately in the following:

**Training data accuracy:** 3D-3D deviation (retrospective)

The inter-rater deviations in viewing direction derived from the 3D training annotations (47 CTs, L4) (Sec. 5.2.1) were comparable for the a.-p. and lateral standard, since the standard planes were defined under the orthogonality constraint. The orthogonality constraint was introduced to increase the accuracy of the training data. Comparing the out-of-plane rotations, lower deviations were observed for the orbital rotation  $\alpha$  than for the angular rotation  $\beta$ . The orbital rotation was strictly constrained by the spinal symmetry axis, while the angular rotation, constraint by the parallelism condition on the vertebra endplates, allowed for more variation, e.g., if the ground and cover plate are not parallel, observed in an osteoporotic or degenerated vertebra. The derived measure can be interpreted as an upper bound for the reachable accuracy ( $\theta_{a.-p.} = 2.50^\circ$ ,  $\theta_{lateral} = 1.85^\circ$ ). However, due to the orthogonality constraint and the available 3D image information, the measure is not representable for manual C-arm positioning accuracies, where the optimal orientation must be deduced from 2D summation images with much higher expected variance.

**Validation data accuracy:** 2D-3D deviation (retrospectively)

The retrospectively assessed 2D-3D deviations on the validation dataset (Sec. 5.2.2) served as a more realistic measure for manual C-arm positioning accuracies. Variances were significantly decreased for the specimens with spinal implants because vertebrae were visible after soft tissue dissection, which facilitated orientation. Consequently, only the 7 specimens without metal were considered as references. The a.-p. and lateral deviations

( $d\theta_{a-p.} = 5.96^\circ$ ,  $d\theta_{lateral} = 4.7^\circ$ ) are in a similar range and exceed the 3D-3D deviations. The 2D-3D deviation was used as clinical accuracy bound. The retrospective assessment of 2D-3D deviations after manual standard projection acquisition required the acquisition of a corresponding 3D volume. Since 3D acquisitions involve additional radiation, they are only performed in complicated anatomical scenarios requiring 3D verification. Consequently, the study could only be performed in a specimens experiment, explaining the limited number of cases.

### **Standard projection accuracy:** 2D-2D deviations via *Swipe App*

The *Swipe App* experiment was performed to assess the clinical pose variability of accepted standard projections on a random subset of the simulated training data (Sec. 5.2.3). The rotational degrees-of-freedom (dof)  $\alpha$ ,  $\beta$  were assessed separately. The decision if the presented image is accepted or rejected as standard projection was based on a single image, the same input used by the proposed automatic pipeline for C-arm positioning. Thus, it allowed a direct comparison of human vs. algorithm performance on simulated data under identical conditions. Regarding the a.-p. standard, higher variations were observed in angular rotation  $\beta$  than orbital rotation  $\alpha$ , which was stronger constrained by the spinal coronal symmetry axis. This observation was reversed for the lateral standard due to the lack of a sagittal symmetry axis. The proposed method performed as well as a human or even more accurately and estimated all rotational and translational dof simultaneously (Tab. 5.4). Sometimes the *Swipe App* decisions were challenged by poor simulation quality, which impeded the bone assessment. Further, the surgeons expressed the lack of a continuous sequence to determine correct views, which hindered their decisions. However, the single-frame 2D input and image quality are the same information the proposed automatic method received. That indicates that the network is able to learn visual clues for pose estimation from a single image, which are difficult to deduce for a human expert.

### **C-arm positioning accuracy:** 2D-2D deviation (inter-/intra-rater)

The inter-/intra-rater 2D-2D manual C-arm positioning deviations were assessed on 2 specimens across all vertebra levels (Th1-L5) The results in Sec. 5.2.3, Fig. 5.15 indicate that intra-rater manual positioning variation is lower than inter-rater, reflecting the effect of surgical preference and missing standardization. Fig. 5.15 confirms the 2D-2D retrospective results in terms of rotation-specific deviation contributions. The computed deviations are higher for the a.-p. standard than for the lateral standard as observed in the previous measures (cp. Tab. 5.4). The manual C-arm positioning variances are much higher than retrospective 3D deviations due to the prior orthogonality assumption and additional spatial information in the 3D standard plane definition in contrast to 2D standard projections.

The derived inter-rater 2D-2D variances can serve as a threshold for acceptable deviations in clinical practice. However, it needs to be considered that the evaluation is only based on 2 specimens and averaged across all vertebra levels (Th1-L5). The measure delivers a rough estimate of expected variations in manual C-arm positioning. Repetition of the performed study on additional specimens would be required to ensure its generalization.

**Manual C-arm positioning efficiency:** Besides the manual C-arm positioning accuracy, also the positioning efficiency was assessed on the 2 specimens across all vertebra levels (Th1-L5) (Sec. 5.2.3, Fig. 5.16). The efficiency was measured in terms of the number of X-rays acquisitions, dose, and time. All assessed parameters correlated with the level of experience of the surgeon. The positioning for lateral standard projections required more dose due to the automatic exposure control (AEC), which increases the exposure factors (kV, mA) when the X-ray beam passes through thicker, stronger attenuating tissue. Regarding the a.-p. standard 6.5 X-rays were acquired on average for positioning (min/-max: 2-29) and for the lateral standard 5 X-rays (min/max: 1-30). The a.-p. positioning also includes the localization, whereas the lateral positioning was initialized from the a.-p. standard and only involves the orientation. With the proposed automatic approach, the number of positioning X-rays can be reduced to only 1 initial X-ray, assuming it is acquired in the capture range of the algorithm.

## 6.2.2 Automatic C-arm positioning

Due to the flexible acquisition setup with the mobile C-arm, intraoperative X-rays with annotated pose labels do not exist, and learning was based on simulations from retrospective CT scans. Real X-rays with k-wire and screws were acquired in a specimens study to validate the transferability from simulated training to intraoperative X-rays.

**Training data:** The training data was simulated from 47 CT scans with spatial dimensions of  $256 \times 256$  pixel. With a detector resolution of  $30 \times 30$  cm, this resulted in a squared resolution of  $1.17$  mm. Detector rotation and translation were approximated by image transformations, and translation along the beam direction was modeled by image scaling. Thereby, it was assumed that at reduced image resolutions, non-isotropic radiation patterns due to collimation and cone-beam radiation could be neglected. The proposed pipeline was trained for 5 dof, including the three rotational parameters and the 2 translations in the detector plane. The network was trained to be invariant towards image scaling for two reasons: 1) Translation along the beam direction is irrelevant since it does not influence the projection plane. In clinical practice, the detector is positioned as close to the patient as practical, constrained by patient anatomy or patient positioning. 2) The

scale is influenced by the patient size and can only be measured concerning a reference when a preoperative CT is available. However, a 3D scan is only acquired for complicated scenarios that require additional 3D verification. To acquire accurate standard projections, prediction of orientation and translation parallel to the detector plane is sufficient. To ensure sufficient quality of the simulations, CT volumes with a spatial resolution of up to 1 mm were collected and preprocessed to prevent extremities from superposition.

**Simulated experiments:** On simulated data, not seen during training, the proposed approach ( $d\theta_{a.-p.} = 3.48 \pm 2.48^\circ$ ,  $d\theta_{lateral} = 4.83 \pm 6.00^\circ$ ) was significantly more accurate in beam direction than the 2-step direct intensity-based pose regression ( $d\theta_{a.-p.} = 5.03 \pm 3.97^\circ$ ,  $d\theta_{lateral} = 9.60 \pm 8.68^\circ$ ), while only requiring a single prediction step, which further decreases the necessary radiation. The quantitative accuracies obtained with the fully automatic pipeline almost reached the upper baseline of ground truth landmark-based pose regression ( $d\theta_{a.-p.} = 3.26 \pm 1.83^\circ$ ,  $d\theta_{lateral} = 4.36 \pm 6.27^\circ$ ), indicating that the 28 CT volumes underlying the training data simulation covered sufficient anatomical variation to generalize to new unseen cases (Fig. 5.28). The translational offset with  $dx_{a.-p.} = 2.56 \pm 1.85mm$ ,  $dy_{a.-p.} = 3.59 \pm 7.95mm$  and  $dx_{lateral} = 3.56 \pm 2.75mm$ ,  $dy_{lateral} = 3.32 \pm 7.28mm$  was below the average vertebra height of 21.2 – 31.1 mm (Kim et al., 2013), so that the correct vertebra of interest was always centered in the projection. The patient-specific evaluation revealed that the outliers resulted from specific patients (cf. Sec. 5.2.6, Fig. 5.24). This concerned the *case 6*; for the a.-p. standard the image quality was decreased due to intraabdominal shadowing air, and for the lateral standard, the bones of the upper extremities superimposed the projection. Further, *case 5* showed superposition of abdominal air pockets in lateral projections, decreasing the performance. Excluding these cases from the evaluation decreased the computed standard deviations to  $d\theta_{a.-p.} = 2.96 \pm 1.89^\circ$ ,  $d\theta_{lateral} = 3.74 \pm 2.63^\circ$ . The out-of-distribution scenario with the overlaid extremities is of little clinical relevance since the extremities are positioned outside the field-of-view (FoV) before image acquisition. The outliers resulting from decreased image contrast caused by superimposed air pockets can probably be addressed by increasing the representation of these challenging conditions in the training data. The pose regression performance was shown to be independent of the initial offset to the desired standard pose (cf. Fig. 5.26), indicating that the vertebra segmentation and landmark detection modules are not biased toward specific views. It was experimentally confirmed that an iterative procedure does not decrease the deviations on average. Nonetheless, a 2-step procedure with separate coarse and fine positioning, like proposed for the intensity-based pose regression, could be transferred to the proposed approach and correct small pose offsets. However, the computed performance with the 1-step approach already complies with the derived clinical accuracy bound of  $d\theta_{a.-p.}^{2D-3D} = 5.96 \pm 3.98^\circ$ ,  $d\theta_{lateral}^{2D-3D} = 4.70 \pm 2.18^\circ$ .

**Validation data:** Validation data was acquired in a specimens study which allowed C-arm acquisitions from arbitrary many poses. However, the cadaveric X-rays may include scenarios not resembled in clinical data, e.g., intraabdominal air caused by cadaver decay, projection artifacts caused by the cadaver bag, or extremities in the FoV caused by cadaver placement.

The validation data consisted of acquisitions from the orbital projection sequence acquired at high quality (30 s scan with 400 images). The benefit of the conducted acquisition setup is that the pose annotations could be directly derived from the metadata of the C-arm because the orbital and angular positioning angles are stored for each sequence. In contrast, during manual C-arm positioning, iterative single images are acquired which do not contain any pose reference and thus could not serve as validation data without an external tracking system. The single images result from higher doses than the acquisitions of the orbital projection sequence, improving the image contrast. Moreover, the single images are internally post-processed for contrast enhancement. Integrating these internal processing steps in the proposed pipeline and evaluation on single projection inputs could improve the robustness further.

**Specimens experiments:** The proposed sequential framework for C-arm pose estimation for standard projections was solely trained on simulated data, which uniquely provides ground truth pose annotations. The domain gap to real X-rays containing superimposed spinal implants was addressed by the proposed k-wire augmentation and realistic screw simulation technique. It was demonstrated for the a.-p. and lateral standard projection that the proposed sequential approach combined with the implant simulation strategies increases accuracy, robustness, and generalization properties compared to direct intensity-based pose regression. For validation, a large specimens study was performed covering different stages of a real clinical procedure. Overall, the approach was capable of generalizing from *in silico* data to *ex vivo* human cadaver data and was applicable to different patient anatomies. Failure cases could be attributed to scenarios not seen during training (cadaveric bags, air pocket superpositions caused by cadaveric decay, lateral arm placement) and low image contrast. Out-of-distribution failure cases can be addressed by further increasing the variety of the CT volumes used for training data generation. Further, the simulated training data can be complemented in future work with real X-rays from clinical routine to represent intraoperative artifacts additionally. In general, the lateral standard projections showed lower image contrast than a.-p. standard projection, which could be explained by the increased attenuation in lateral projections due to higher tissue penetration depths (cf. Fig. 5.13). These failure cases could be addressed by acquiring higher quality initial X-rays in the form of single projections with the help of collimation

windows. Validation data acquisition with single images would require a different setup for pose tracking, e.g., integrating an external tracking system.

Under ideal conditions, pose regression from manual reference annotations resulted in performances that were in agreement with the derived clinical accuracy bound for all specimens, except *case 8* in the lateral standard (Fig. 5.29). The failure case could be attributed to severe lumbar scoliosis that was not covered in the training data. The performances for the a.-p. and lateral standard were comparable under ideal conditions, indicating that the selected landmarks are equally suitable for both standards and in between specimens without and with metal. This suggests that if the landmarks are correctly detected, this enables successful subsequent pose regression.

Training on a pose subset of the real X-rays led to superior performances than with ground truth landmarks. Since all validation specimens were present during training, it was not ensured that the network only memorized from close-by examples. A leave-one-out training on real X-rays did not improve the performance, indicating that failure cases were not caused by the domain gap between simulated and real X-rays but rather results from patient-specific anatomical conditions or artifacts that were not covered in the training data (Fig. 5.19).

The visual intermediate landmark detection adds transparency to the decision process. Since spinal neurosurgeons also use the detected features to identify correct standard projections, clinical interpretability of the predictions is explicitly promoted, which can improve the acceptability of deep learning algorithms in the medical environment. Further, wrong region-of-interest (ROI) localization and landmark detection can indicate potential failure cases of subsequent pose regression (cf. Fig. 5.22). This suggests that the intermediate steps of landmark detection can be employed to estimate the confidence of the final pose regression.

The proposed approach was designed to be seamlessly integrated into the clinical workflow and has the potential to decrease the number of necessary acquisitions and consequently the radiation dose for patients and clinical staff drastically. Furthermore, automation can help to improve and standardize the quality of standard projections in clinical practice.

**Pipeline design:** Different factors that influence the final performance of the pose estimation were analyzed and are discussed in the following.

**Implant simulation:** K-wires were augmented randomly in 2D and screws were forward projected after realistic modeling in 3D. Since k-wires are flexible and can occur at arbitrary poses during the insertion process, a random 2D augmentation strategy was selected instead of a realistic 3D modeling technique. Although the proposed 2D structure augmentation method does not model divergent beam effects, it shows to increase the robustness to highly attenuated, elongated structures effectively.

The proposed k-wire and implant simulation strategies were each applied to 1/3 of the data such that 1/3 of the training data did not contain metal. Consequently, the performance of the method with implant simulations performed at least as good as the one without implant simulations on the real X-rays without metal (cf. Fig. 5.34). Some cases without metal benefited from the metal simulation, especially *case 0* and *case 6* in the lateral standard. These cases are out-of-distribution due to projection artifacts resulting from the cadaver bag and upper extremity superposition. The metal augmentation shows to increase the robustness for these cases. This was confirmed by the evaluation on real X-rays with metal. Across all specimens, an increased pose offset was observed for the approach without implant simulations that was significantly decreased by incorporating the proposed k-wire and screw simulation into the training pipeline. The results highlight the importance of the proposed implant simulation strategies to bridge the domain gap from simulation-based training to intraoperative X-rays with spinal implants. Further, it suggests that additional augmentation strategies, e.g., collimation windows, surgical tool simulations, or random pixel dropout, can improve the performance for challenging cases even further.

**ROI localization:** The ROI detection module reflects that standard projections are only defined by the bone structures and not by the surrounding. By removing and replacing individual modules of the proposed pipeline, it was experimentally confirmed on simulated and real X-rays that the ROI detection module improves the overall performance (Fig. 5.32). The ROI localization prevents confounding factors present in the surrounding tissue or external surgical equipment from falsifying the prediction. It reflects the use of collimators in real clinical procedures, and further range normalization within such ROIs leads to better image contrast that benefits the subsequent task of landmark detection. The experiments showed that a bounding box ROI detection performed superior to a segmentation mask ROI detection for the real X-rays. The ROI detection via segmentation masks depends on accurate vertebra segmentation to ensure that the relevant image information for subsequent landmark detection still exists. In contrast, the bounding box ROI detection tolerates minor segmentation inaccuracies without eliminating relevant image features.

**Landmark detection:** In consultation with a spinal neurosurgeon, the following clinically meaningful anatomy-specific landmarks were selected: (1) endpoints of the transverse and spinous processes, (2) pedicles, (3) vertebra base and cover plate. According to the importance analysis of the single landmark channel on simulated data (Fig. 5.27), the endplates have the strongest influence on the accuracy because the 2D plates should ideally project to a line that poses the strongest orientation constraint. This was followed by the pedicles, which should project symmetrically for the a.-p. standard and superimpose for the lateral standard. The low influence of the endpoints landmarks was clinically expected since the transverse processes generally do not serve as a strict clinical reference for standard projections. However, the analysis showed that they still have a small beneficial effect, and the best overall performances were observed by combining all proposed landmark types.

The landmark detection module clearly improved the performance compared to direct intensity-based pose regression on simulated and real X-rays (Fig. 5.32). The results suggest that guiding the decision process by explicit anatomical landmarks that spinal neurosurgeons consider in diagnosis improves the generalization capabilities and the reliability of the model, especially in challenging intraoperative settings.

### 6.2.3 Conclusion

The current manual C-arm positioning procedure under repeated acquisitions or even continuous fluoroscopy is time-consuming and radiation-intense. In this thesis, a novel automatic C-arm positioning pipeline was proposed which fulfills all the requirements that were stated in the objectives.

**Clinical applicability:** The proposed approach directly works on the initial 2D projection, it does not require any additional technical equipment or 3D image information, enabling seamless clinical workflow integration.

**Performance requirements:** A large specimens study was performed to investigate the performance of the proposed method in different intraoperative settings. To set the computed performances into clinical context, different experiments were designed to assess the acceptable variability of standard projections under manual positioning. The approach showed strong performance that met the derived clinical accuracy bound for many cases. Failure cases could be attributed to out-of-distribution scenarios and can be addressed in the future by increasing the training data variation.

**Benefits for the patient:** Manual C-arm positioning can involve 1-30 X-ray acquisitions (6 on average), whereas the proposed automatic approach improved the positioning efficiency by only requiring a single initial X-ray in the capture range of the algorithm.

Additionally, the positioning time can be decreased from 5 – 474 s (68.8 s on average) under iterative radiation to a runtime of 0.12 s with the automatic approach.

**Benefits for the surgeon:** Spinal neurosurgeons are exposed to the highest amount of radiation among all trauma surgeons, which could be drastically decreased by the proposed computer-assisted positioning support. The average execution time of 0.12 s would drastically decrease the C-arm positioning and thereby the intervention time. Further, automatic C-arm positioning can standardize the quality of acquired standard projections and decrease the expertise dependence of results.

**Representation of anatomical variability:** The proposed approach is solely trained on simulations from CT scans. Anatomical variation can be easily increased during training by incorporating additional CT scans.

**Addressing intraoperative requirements:** Standard projections are used to monitor the implant or wire position. This means that bone fragments are already relocated to their initial position by inserting wires or implants. K-wire augmentation and realistic screw simulation techniques were proposed to bridge the domain gap from simulations to intraoperative X-rays with spinal implants.

**Generalization:** The developed pipeline is transferrable to other standard projections, vertebra levels, or anatomies, as further discussed in future work.

#### 6.2.4 Limitations and future work

The proposed pipeline for automatic C-arm positioning was evaluated in a specimens study including 16 specimens. Observed failure cases can be addressed in future work by increasing the anatomical variation during training by (1) collecting more high-resolution isotropic CT volumes for training data simulation, (2) introducing additional augmentation strategies, e.g., collimation window or random pixel dropout for improved generalization properties, or (3) generalizing the proposed implant simulation strategies to also model other surgical instruments present during spinal interventions.

Clinical translation of the proposed method would require further validation in different scenarios and on different patient anatomies. Besides specimens studies, the collection of intraoperative datasets for validation and training would be necessary to evaluate the approach in real clinical conditions. This would require datasets with consistent projection sequence and reconstructed volume. An intraoperative validation dataset would have the advantage of only including clinically relevant artifacts. However, clinical datasets will represent only a limited angular range which could be used for validation but will introduce a bias in training. Integration of intraoperative datasets during training has the potential to improve the landmark prediction in challenging scenarios.

Further, it would be interesting to extend the approach for application on (1) other vertebra levels and (2) other spinal standard projections, e.g., bull's eye or oblique views, (3) other anatomies. Application of the approach to arbitrary vertebra levels would require localizing and classifying the respective vertebra in the initial X-ray. This could be performed either automatically (Manbachi et al., 2018; Cina et al., 2021) or by manual user input. Transfer to other anatomies would require the definition of corresponding anatomy-specific relevant landmarks.

The intermediate step of visual landmark detection could be employed for confidence estimation of pose regression for automatic outlier identification and warning. Further, it can be analyzed if a multi-step approach that takes the predictions of the failed cases as input can improve upon the initial performance.

Device integration of the proposed automatic positioning method would require a robotic C-arm that is ideally designed to rotate around the isocenter for all rotational degrees of freedom. Current C-arm devices require a C-arm base movement to achieve pure gantry rotation perpendicular to the beam direction for the a.-p. projection and to vary angulation for the lateral projection. Additionally, for the Cios Spin<sup>®</sup> only the up translation is motorized, and forward and side motion can only be performed manually. Further, clinical integration of a robotized C-arm with automatic positioning would require environmental collision handling, which was not in the scope of this thesis.

The proposed method can be integrated into the clinical workflow without additional external hardware while reducing the time and dose consumption compared to the current manual positioning procedure.



## Summary

Spinal fusion using pedicle screws is the gold standard technique to treat spinal instabilities. Accurate screw placement is essential due to the proximity of the pedicle to the spinal cord and the related complications in case of co-injury. However, a high level of anatomical knowledge and expertise is required. Different image-assisted techniques enable the visualization of the internal anatomy and facilitate pedicle screw insertion, e.g., fluoroscopy guidance or computed tomography (CT) navigation. This thesis proposes two computer-assisted methods to support the surgeon during both techniques.

During fluoroscopy guidance, repeated anatomy-specific standard projections are acquired. Standard projections are X-rays acquired from patient-specific C-arm poses and allow assessing the fracture reduction and implant placement. In the current clinical routine, the C-arm is positioned manually under iterative or continuous fluoroscopy, involving a high radiation dose and time consumption. CT navigation gives an alternative approach for image guidance of spinal interventions. It involves the acquisition of an intraoperative CT in which screw trajectories are manually planned for subsequent 3D navigation. Both image guidance techniques require manual interventions, which are highly expert-dependent, require in-depth knowledge of the anatomy, understanding of anatomical orientation, and increase the procedural time.

The methods developed in this thesis should support the surgeon with C-arm positioning during fluoroscopy guidance and in pedicle screw planning during CT-navigation. Many state of the art approaches that propose computer assistance for the manually performed steps make restrictive prior assumptions about modeling or acquisition settings or require external hardware that impedes clinical workflow integration and limits clinical applicability until today. In this thesis, deep learning techniques are employed that learn the anatomical variation from retrospective CT datasets and additional simulations complemented with expert annotations without requiring any other technical equipment.

Specifically, for automatic screw planning, a method was developed that simulates screw positions using surgical planning data in retrospective CT datasets and learns them in a patch-based approach based on the image context; instead of relying on manually defined geometric and structural properties or atlas-based modeling. The proposed approach via an auxiliary segmentation task outperformed direct convolutional neural network-based landmark regression. Further, it was integrated as a plugin into the open-source frame-

work Medical Imaging Interaction Toolkit. The prototype implementation demonstrated clinically acceptable screw plans in 96.2% of screw trajectories. The automatic planning performance complied with the manual inter-rater variances, suggesting that the proposed method can replace the current manual expert planning. In addition, the proposed automatic planning method drastically increased the time efficiency by 8-fold (4.9 s per vertebra level), which improved upon state of the art computing times.

For automatic C-arm positioning for standard projections, anatomical variations were learned likewise from a retrospective diverse CT dataset using deep learning techniques. The C-arm pose update was estimated directly from a first X-ray without the need for a pre-operative CT or additional technical equipment, in contrast to many state of the art approaches. Compared to the current manual, iterative positioning method, the proposed approach would dramatically reduce radiation dose and time. Three different challenges were addressed. Firstly, real X-rays with annotated pose labels do not exist due to the lack of a constant reference frame in intraoperative fluoroscopy. Therefore, learning was based on simulations generated from CT scans. The second challenge is related to the intraoperative requirements, where surgical implants can be present and partly overlay the anatomy. This was addressed by complementing the synthetic X-rays with simulated k-wires and screws. To further achieve a transfer from simulated to real X-rays, the pipeline was designed by mimicking the clinical decision-making processes of spinal neurosurgeons. The inclusion of visible anatomical landmarks within the pipeline provides insights into the decision process. It increases clinical interpretability, which addresses the third challenge of deep learning algorithms, often being considered as black boxes. Evaluated on a large specimens study covering different steps of a real clinical procedure, the proposed extensions increased accuracy and robustness compared to direct intensity-based pose regression. The method can handle inter-patient anatomical variation and generalizes from *in silico* to *ex vivo* human cadaver data. Failure cases could be related to specific scenarios like the cadaver bag, cadaver decay, lateral arm superpositions, which could be addressed by increasing the training data variety. Further clinical evaluation is necessary to investigate the robustness and usability in an intraoperative setting. The presented approach was evaluated for the fourth lumbar vertebra level and the anterior-posterior and lateral standard projection. It can be transferred to other vertebra levels, standard projections, or anatomies with available training data.

In conclusion, the two proposed computer-assistance systems for spinal surgery offer support for the surgeon during two distinct image-guided techniques. Seamless integration of the proposed methods in the clinical workflow is facilitated. It can significantly decrease the manual interaction and expertise-dependence of results, which is especially beneficial for inexperienced surgeons, and also reduce the radiation exposure, and procedural time, both for patients and clinical staff.

## Zusammenfassung

Die Wirbelfusion über Pedikelverschraubung ist der Goldstandard für die Behandlung instabiler Wirbelsäulenerkrankungen. Durch die Nähe der Pedikel zum Rückenmark und der damit verbundenen Komplikationen im Falle von Verletzungen ist eine genaue Schraubenplatzierung essenziell, erfordert jedoch ein hohes Maß an anatomischem Wissen und Erfahrung. Verschiedene bildgestützte Verfahren ermöglichen die Visualisierung der inneren Anatomie und erleichtern die Schraubenplatzierung, z. B. Durchleuchtung oder CT-Navigation. In dieser Arbeit werden zwei computergestützte Methoden zur Unterstützung des Chirurgen bei beiden Verfahren vorgeschlagen.

Bei der Durchleuchtung werden wiederholt anatomiespezifische Standardprojektionen aufgenommen. Standardprojektionen sind Röntgenbilder, die aus patientenspezifischen C-Bogen Posen aufgenommen werden und die Beurteilung der Frakturposition und der Implantatplatzierung ermöglichen. In der derzeitigen klinischen Routine wird der C-Bogen unter iterativer oder kontinuierlicher Durchleuchtung manuell positioniert, was eine hohe Strahlendosis und einen hohen Zeitaufwand mit sich bringt. Die CT-Navigation bietet einen alternativen Ansatz für die bildgestützte Navigation von Wirbelsäuleneingriffen. Dabei wird intraoperativ ein CT aufgenommen, in dem die Schraubentrajektorien für die anschließende 3D-Navigation manuell geplant werden. Beide bildgestützten Verfahren erfordern manuelle Interventionen, die in hohem Maße expertenabhängig sind, tiefe Kenntnisse der Anatomie und der anatomischen Ausrichtung erfordern und die Verfahrensdauer verlängern.

Die in dieser Arbeit entwickelten Methoden sollen den Chirurgen bei der C-Bogen-Positionierung während der Durchleuchtungsführung und bei der Pedikelschraubenplanung während der CT-Navigation unterstützen. Viele aktuelle Ansätze, die eine Computerunterstützung für die manuell ausgeführten Schritte vorschlagen, gehen von restriktiven Annahmen bezüglich der Modellierung oder der Aufnahmetechnik aus oder erfordern externe Hardware, was die Integration in den klinischen Arbeitsablauf erschwert und die klinische Anwendbarkeit bis heute einschränkt. In dieser Arbeit werden Deep Learning Techniken verwendet, die die anatomische Variation aus retrospektiven CT Datensätzen und zusätzlichen Simulationen lernen, die durch Expertenannotationen ergänzt werden, ohne dass weitere technische Geräte erforderlich sind.

Konkret wurde für die automatische Schraubenplanung eine Methode entwickelt, die

die Schraubenpositionen anhand von chirurgischen Planungsdaten in retrospektiven CT Datensätzen simuliert und in einem patchbasierten Ansatz auf der Grundlage des Bildkontextes erlernt, anstelle von manuell definierten geometrischen und strukturellen Eigenschaften oder atlasbasierter Modellierung. Der vorgeschlagene Ansatz über eine zusätzliche Segmentierungsaufgabe zeigte genauere und robustere Performance als die direkte convolutional neural network-basierte Landmarkenregression. Die Methode wurde als Prototyp Plugin in das Open-Source-Framework Medical Imaging Interaction Toolkit integriert und führte in 96.2% der Fälle zu klinisch akzeptablen Schraubenplänen. Die automatische Planungsleistung war vergleichbar mit der inter-rater Varianz unter manueller Planung, sodass die vorgeschlagene Methode die derzeitige manuelle Expertenplanung ersetzen kann. Darüber hinaus erhöhte die vorgeschlagene automatische Planungsmethode die Zeiteffizienz drastisch um das 8-fache (4.9 s pro Wirbelebene), was eine Verbesserung gegenüber dem Stand der Technik darstellt.

Für die automatische C-Bogen Positionierung für Standardprojektionen wurde eine Methode entwickelt, die anatomische Variationen ebenfalls aus einem retrospektiven diversen CT-Datensatz mit Deep Learning Techniken lernt. Die gewünschte Positionierung des C-Bogens wurde direkt aus einem ersten Röntgenbild geschätzt, ohne die Notwendigkeit eines präoperatives CT oder zusätzlichen technischen Equipments, im Gegensatz zu vielen anderen state of the art Ansätzen. Im Vergleich zum derzeitigen manuellen, iterativen Positionierungsverfahren würde der Ansatz die Strahlendosis und den Zeitaufwand drastisch reduzieren. Dabei wurden drei verschiedene Herausforderungen angegangen. Erstens gibt es keine echten Röntgenbilder mit annotierten Pose Labels, da es in der intraoperativen Fluoroskopie keinen konstanten Referenzrahmen gibt. Daher basierte das Lernen auf Simulationen, die anhand von CT Scans erstellt wurden. Die zweite Herausforderung bezieht sich auf die intraoperativen Anforderungen, wo vorhandene chirurgische Implantate die Anatomie teilweise überlagern können. Dieses Problem wurde durch die Ergänzung der synthetischen Röntgenbilder mit simulierten K-Drähten und Schrauben adressiert. Um einen Transfer von simulierten zu realen Röntgenbildern zu erreichen, wurde die Pipeline so konzipiert, dass sie die klinischen Entscheidungsprozesse von Neurochirurgen nachahmt. Die Einbeziehung sichtbarer anatomischer Orientierungspunkte in die Pipeline bietet Einblicke in den Entscheidungsprozess. Sie erhöht die klinische Interpretierbarkeit, was die dritte Herausforderung von Deep Learning Algorithmen adressiert, die oft als Blackboxen betrachtet werden. Die erweiterte Methode wurde in einer großen Studie evaluiert, wobei verschiedene Schritte eines realen klinischen Eingriffs simuliert wurden. Im Vergleich zur direkten intensitätsbasierten Posenregression konnte die Genauigkeit und Robustheit erhöht werden. Die Methode kann mit anatomischer Variation umgehen und lässt sich von *in silico* auf *ex vivo* Daten menschlicher Kadaver übertragen. Fehler konnten auf spezifische Szenarien wie den Leichensack, den Leichenzerfall

---

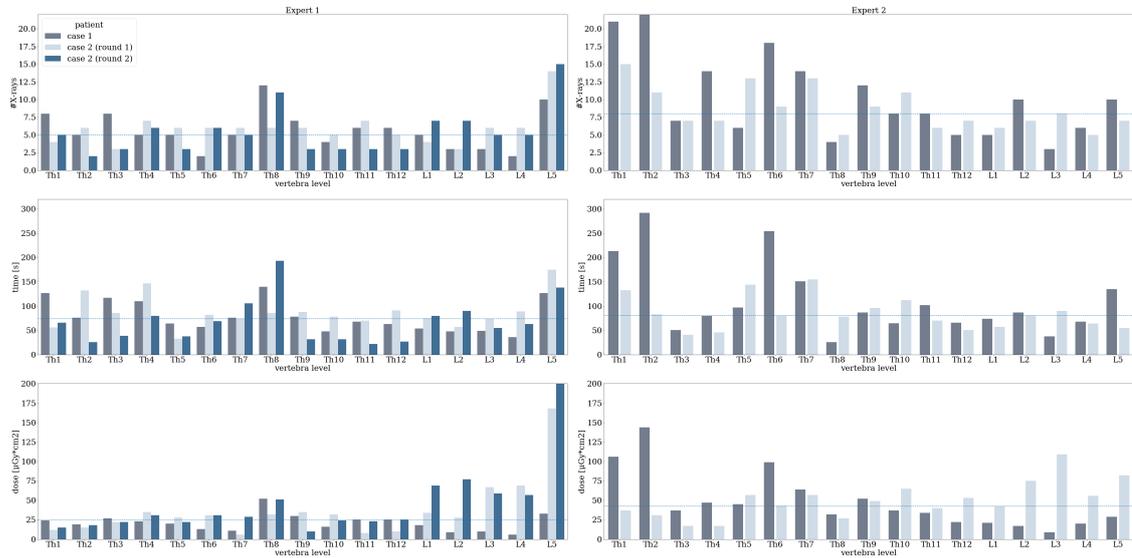
oder seitliche Armüberlagerungen zurückgeführt werden, was durch eine größere Vielfalt an Trainingsdaten behoben werden könnte. Weitere klinische Evaluation ist erforderlich, um die Robustheit und Anwendbarkeit in einem intraoperativen Umfeld zu untersuchen. Der vorgestellte Ansatz wurde für den vierten Lendenwirbel und die anterior-posteriore und laterale Standardprojektion evaluiert. Er kann auf andere Wirbelniveaus, Standardprojektionen oder Anatomien übertragen werden, wenn entsprechende Trainingsdaten zur Verfügung stehen.

Zusammenfassend bieten die beiden vorgeschlagenen Computer-Assistenzsysteme für die Wirbelsäulen Chirurgie eine Unterstützung für den Chirurgen während zwei unterschiedlicher bildgebender Verfahren. Die vorgeschlagenen Methoden können nahtlos in den klinischen Arbeitsablauf integriert werden. Dadurch kann die manuelle Interaktion und die Abhängigkeit von Expertenwissen erheblich verringert werden, was insbesondere für unerfahrene Chirurgen von Vorteil ist. Auch die Strahlenbelastung und die Eingriffsdauer kann sowohl für Patienten als auch für das Klinikpersonal reduziert werden.

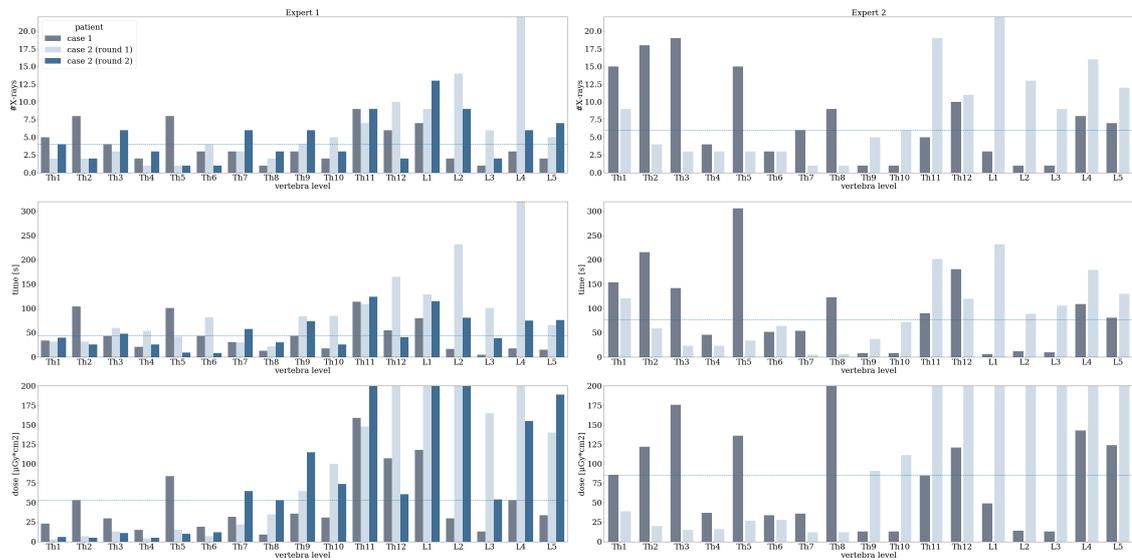


*A*

**Appendix**

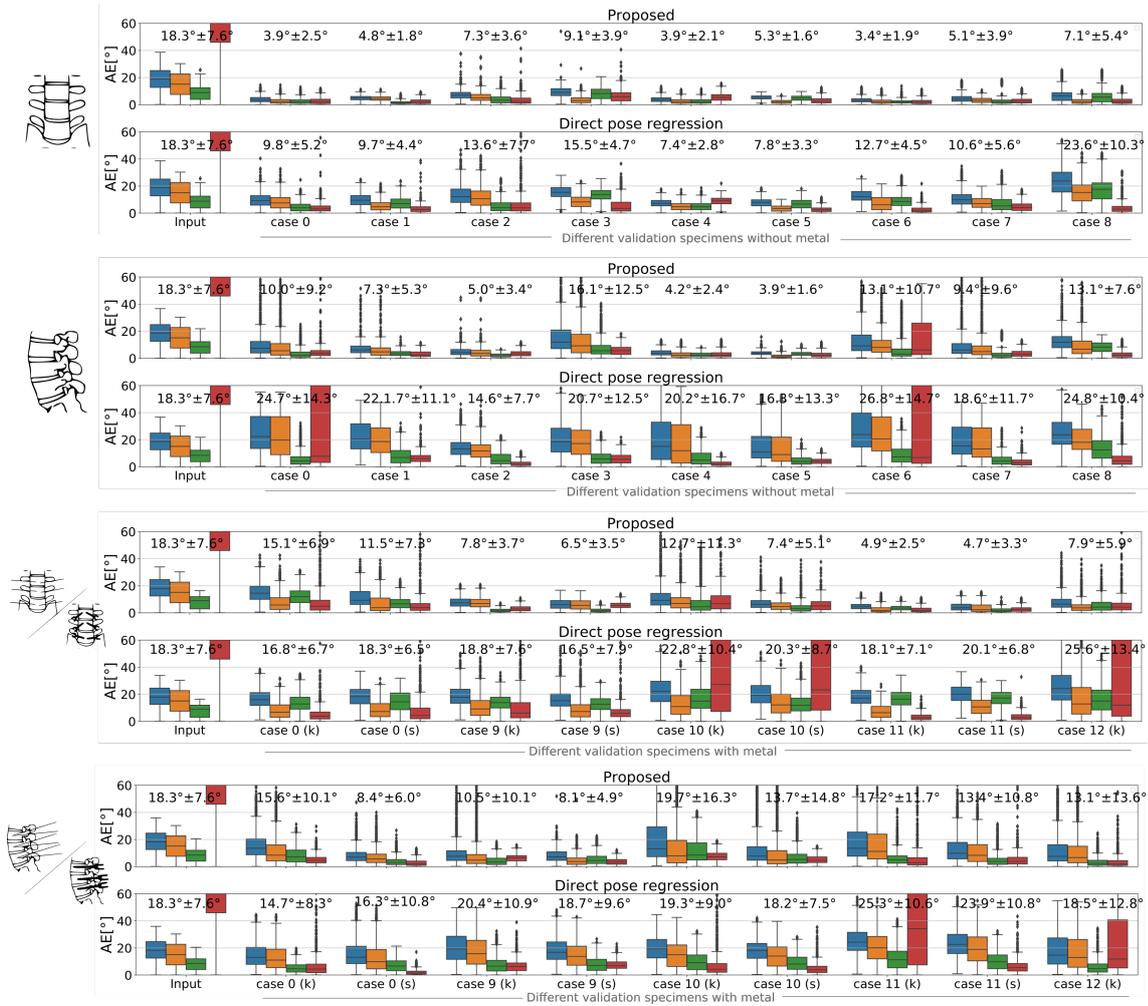


(a) a-p. standard



(b) lateral standard

**Figure A.1: Vertebra-level specific evaluation of manual C-arm positioning efficiency.** The number of X-rays, time and required doses are shown for each specimen and standard projection assessed for each rater separately.



**Figure A.2: Quantitative evaluation of proposed pose regression approach on real X-rays.** Evaluation of the proposed pose regression method for the **a-p. standard** and **lateral standard** on real X-rays **without / with metal** with  $MAE_{d\theta} \pm SD_{d\theta}$  indicated for each specimen case 0 - case 12, k and s indicating the presences of k-wires or screws. The proposed method shows improved generalization capabilities compared to direct intensity-based regression. A-p. figures taken and extended from Kausch et al. (2021b).

(a) inter-rater

$\mu \pm \sigma$	a.-p. standard					lateral standard				
	$d\theta$ [°]	$d\alpha$ [°]	$d\beta$ [°]	$tx$ [mm]	$ty$ [mm]	$d\theta$ [°]	$d\alpha$ [°]	$d\beta$ [°]	$tx$ [mm]	$ty$ [mm]
case 1	9.21	3.69	7.56	9.69	21.69	8.55	7.45	3.10	8.81	20.49
	±	±	±	±	±	±	±	±	±	±
case 2	5.63	4.45	5.90	5.85	17.31	2.88	3.53	2.15	4.23	17.99
	±	±	±	±	±	±	±	±	±	±
	8.57	2.44	7.86	14.80	16.01	4.17	1.71	3.07	14.85	17.99
	±	±	±	±	±	±	±	±	±	±
	6.05	2.60	5.90	4.60	12.41	3.15	3.31	2.09	3.45	16.36

(b) intra-rater

$\mu \pm \sigma$	a.-p. standard					lateral standard				
	$d\theta$ [°]	$d\alpha$ [°]	$d\beta$ [°]	$tx$ [mm]	$ty$ [mm]	$d\theta$ [°]	$d\alpha$ [°]	$d\beta$ [°]	$tx$ [mm]	$ty$ [mm]
case 2	4.79	1.68	4.12	13.38	20.49	3.51	2.17	1.97	11.30	18.81
	±	±	±	±	±	±	±	±	±	±
	4.74	1.15	4.99	11.25	14.62	4.20	3.88	2.62	10.15	12.56

**Table A.1: Specimen-specific quantitative evaluation of inter/intra-rater study** for manual C-arm positioning accuracy: (a) inter-rater, (b) intra-rater. Angular deviations ( $d\theta$ ,  $d\alpha$ ,  $d\beta$ ) and detector plane translations ( $tx$ ,  $ty$ ) are reported for each specimen and standard projection.

## Bibliography

- Aichinger, H., Dierker, J., Joite-Barfuß, S., and Säbel, M. (2012). **Principles of X-ray imaging**. In *Radiation Exposure and Image Quality in X-Ray Diagnostic Radiology*, pages 3–7. Springer. doi: 10.1007/978-3-642-11241-6\_1.
- Albarqouni, S., Fotouhi, J., and Navab, N. (2017). **X-ray in-depth decomposition: Revealing the latent structures**. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 444–452. Springer. doi: 10.1007/978-3-319-66179-7\_51.
- Aoude, A. A., Fortin, M., Figueiredo, R., Jarzem, P., Ouellet, J., and Weber, M. H. (2015). **Methods to determine pedicle screw placement accuracy in spine surgery: A systematic review**. *European Spine Journal*, 24(5):990–1004. doi: 10.1007/s00586-015-3853-x.
- Attar, A., Ugur, H. C., Uz, A., Tekdemir, I., Egemen, N., and Genc, Y. (2001). **Lumbar pedicle: Surgical anatomic evaluation and relationships**. *European Spine Journal*, 10(1): 10–15. doi: 10.1007/s005860000198.
- Azwan, M. S. and Abd Rahim, I. (2011). **Recent studies on the pullout strength behavior of spinal fixation**. *Journal of Developmental Biology and Tissue Engineering*, 3(4):48–54.
- Badal, A. and Badano, A. (2009). **Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit**. *Medical Physics*, 36(11):4878–4880. doi: 10.1118/1.3231824.
- Berger, M., Qiao, Y., and Maier, A. (2018). **X-ray imaging**. In *Medical Imaging Systems: An Introductory Guide.*, Lecture notes in computer science Image processing, computer

- vision, pattern recognition, and graphics, pages 119–145. Springer. doi: 10.1007/978-3-319-96520-8\_7.
- Bhatia, R. G. and Bowen, B. C. (2007). **Thoracolumbar spine trauma**. In *Spinal Imaging*, pages 325–357. Springer. doi: 10.1007/978-3-540-68483-1\_13.
- Bier, B., Aschoff, K., Syben, C., Unberath, M., Levenston, M., Gold, G., Fahrig, R., and Maier, A. (2018)a. **Detecting anatomical landmarks for motion estimation in weight-bearing imaging of knees**. In *International Workshop on Machine Learning for 3D Reconstruction*, pages 83–90. Springer. doi: 10.1007/978-3-030-00129-2\_10.
- Bier, B., Unberath, M., Zaech, J.-N., Fotouhi, J., Armand, M., Osgood, G., Navab, N., and Maier, A. (2018)b. **X-ray-transform invariant anatomical landmark detection for pelvic trauma surgery**. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 55–63. Springer. doi: 10.1007/978-3-030-00937-3\_7.
- Bier, B., Goldmann, F., Zaech, J.-N., Fotouhi, J., Hegeman, R., Grupp, R., Armand, M., Osgood, G., Navab, N., Maier, A., et al. (2019). **Learning to detect anatomical landmarks of the pelvis in x-rays from arbitrary views**. *International Journal of Computer Assisted Radiology and Surgery*, 14(9):1463–1473. doi: 10.1007/s11548-019-01975-5.
- Binder, N., Bodensteiner, C., Matthäus, L., Burgkart, R., and Schweikard, A. (2006). **Image guided positioning for an interactive c-arm fluoroscope**. In *Computer Assisted Radiology and Surgery*, pages 5–7.
- Blattert, T. R., Fill, U. A., Kunz, E., Panzer, W., Weckbach, A., and Regulla, D. F. (2004). **Skill dependence of radiation exposure for the orthopaedic surgeon during interlocking nailing of long-bone shaft fractures: a clinical study**. *Archives of Orthopaedic and Trauma Surgery*, 124(10):659–664. doi: 10.1007/s00402-004-0743-9.
- Bodart, L. E., Ciske, B. R., Wagner, M., Raval, A. N., and Speidel, M. A. (2018). **Clinical feasibility of x-ray based pose estimation of a transthoracic echo probe using attached fiducials**. In *Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10576, page 1057626. International Society for Optics and Photonics. doi: 10.1117/12.2293875.
- Bohlender, S., Oksuz, I., and Mukhopadhyay, A. (2021). **A survey on shape-constraint deep learning for medical image segmentation**. *arXiv preprint arXiv:2101.07721*.
- Bosco, C. (2015). **Freebie - principle card swipe template**. URL <https://dribbble.com/shots/2219454-Freebie-Principle-Card-Swipe-Template>. [Online; accessed 2021-09-09].

- 
- Bott, O. J., Dresing, K., Wagner, M., Raab, B.-W., and Teistler, M. (2011). **Use of a C-arm fluoroscopy simulator to support training in intraoperative radiography.** *Radiographics*, 31(3):E31–E41.
- Bratschitsch, G., Leitner, L., Stücklschweiger, G., Guss, H., Sadoghi, P., Puchwein, P., Leithner, A., and Radl, R. (2019). **Radiation exposure of patient and operating room personnel by fluoroscopy and navigation during spinal surgery.** *Scientific Reports*, 9(1):1–5. doi: 10.1038/s41598-019-53472-z.
- Bühren, V. and Josten, C. (2012). *Chirurgie der verletzten Wirbelsäule: Frakturen, Instabilitäten, Deformitäten.* Springer-Verlag.
- Bui, M., Albarqouni, S., Schropp, M., Navab, N., and Ilic, S. (2017). **X-ray PoseNet: 6 dof pose estimation for mobile x-ray devices.** In *IEEE Winter Conference on Applications of Computer Vision*, pages 1036–1044. IEEE. doi: 10.1109/WACV.2017.120.
- Bundesärztekammer. (2007). **Leitlinie der Bundesärztekammer zur Qualitätssicherung in der Röntgendiagnostik.** URL <https://www.bundesaerztekammer.de/aerzte/qualitaetssicherung/richtlinien-leitlinien-empfehlungen-stellungnahmen/richtlinien-leitlinien-empfehlungen-zur-qualitaetssicherung/bildgebende-verfahren/roentgendiagnostik/>. [Online; accessed 2021-08-03].
- Buzug, T. M. (2011). *Einführung in die Computertomographie: mathematisch-physikalische Grundlagen der Bildrekonstruktion.* Springer.
- Cina, A., Bassani, T., Panico, M., Luca, A., Masharawi, Y., Brayda-Bruno, M., and Galbusera, F. (2021). **2-step deep learning model for landmarks localization in spine radiographs.** *Scientific Reports*, 11(1):1–12. doi: 10.1038/s41598-021-89102-w.
- Daemi, N., Ahmadian, A., Mirbagheri, A., Ahmadian, A., Saberi, H., Amidi, F., and Alirezaie, J. (2015). **Planning screw insertion trajectory in lumbar spinal fusion using pre-operative CT images.** In *IEEE Engineering in Medicine and Biology Society*, pages 3639–3642. IEEE. doi: 10.1109/EMBC.2015.7319181.
- De Silva, T., Punnoose, J., Uneri, A., Goerres, J., Jacobson, M., Ketcha, M. D., Manbachi, A., Vogt, S., Kleinszig, G., Khanna, A. J., et al. (2017). **C-arm positioning using virtual fluoroscopy for image-guided surgery.** In *Medical Imaging: Image-guided Procedures, Robotic Interventions, and Modeling*, volume 10135, page 101352K. International Society for Optics and Photonics.

- Dhont, J., Verellen, D., Mollaert, I., Vanreusel, V., and Vandemeulebroucke, J. (2020). **RealDRR – Rendering of realistic digitally reconstructed radiographs using locally trained image-to-image translation.** *Radiotherapy and Oncology*, 153:213–219. doi: 10.1016/j.radonc.2020.10.004.
- Dice, L. R. (1945). **Measures of the amount of ecologic association between species.** *Ecology*, 26(3):297–302. doi: 10.2307/1932409.
- Diethelm, L., Joettgen, G., Lentz, W., and Voelkel, L. (1956). **An X-ray fluoroscope with image intensifier.** *Roentgen-Blaetter*, 9:215.
- Esfandiari, H., Anglin, C., Guy, P., and Hodgson, A. J. (2018). **A deep learning-based approach for localization of pedicle regions in preoperative CT scans.** *Computer Assisted Orthopaedic Surgery*, 2:46–50. doi: 10.29007/j56f.
- Esteban, J., Grimm, M., Unberath, M., Zahnd, G., and Navab, N. (2019). **Towards fully automatic x-ray to ct registration.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 631–639. Springer. doi: 10.1007/978-3-030-32226-7\_70.
- Fallavollita, P., Winkler, A., Habert, S., Wucherer, P., Stefan, P., Mansour, R., Ghotbi, R., and Navab, N. (2014). **Desired-view controlled positioning of angiographic C-arms.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 659–666. Springer. doi: 10.1007/978-3-319-10470-6\_82.
- Fan, Y., Du, J. P., Liu, J. J., Zhang, J. N., Liu, S. C., and Hao, D. J. (2018). **Radiological and clinical differences among three assisted technologies in pedicle screw fixation of adult degenerative scoliosis.** *Scientific Reports*, 8(1):1–9. doi: 10.1038/s41598-017-19054-7.
- Fotouhi, J., Unberath, M., Song, T., Gu, W., Johnson, A., Osgood, G., Armand, M., and Navab, N. (2019). **Interactive flying frustums (IFFs): Spatially aware surgical data visualization.** *International Journal of Computer Assisted Radiology and Surgery*, 14(6): 913–922. doi: 10.1007/s11548-019-01943-z.
- Gao, C., Unberath, M., Taylor, R., and Armand, M. (2019). **Localizing dexterous surgical tools in X-ray for image-based navigation.** *arXiv preprint arXiv:1901.06672*.
- Gelalis, I. D., Paschos, N. K., Pakos, E. E., Politis, A. N., Arnaoutoglou, C. M., Karageorgos, A. C., Ploumis, A., and Xenakis, T. A. (2012). **Accuracy of pedicle screw placement: A systematic review of prospective in vivo studies comparing free hand, fluoroscopy**

- 
- guidance and navigation techniques.** *European Spine Journal*, 21(2):247–255. doi: 10.1007/s00586-011-2011-3.
- Ghafurian, S., Metaxas, D. N., Tan, V., and Li, K. (2016). **Fast generation of digitally reconstructed radiograph through an efficient preprocessing of ray attenuation values.** In *Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9786, page 97860C. International Society for Optics and Photonics. doi: doi.org/10.1117/12.2217756.
- Goerres, J., Uneri, A., De Silva, T., Ketcha, M., Reaungamornrat, S., Jacobson, M., Vogt, S., Kleinszig, G., Osgood, G., Wolinsky, J., et al. (2017). **Spinal pedicle screw planning using deformable atlas registration.** *Physics in Medicine & Biology*, 62(7):2871. doi: 10.1088/1361-6560/aa5f42.
- Gong, R. H., Jenkins, B., Sze, R. W., and Yaniv, Z. (2014). **A cost effective and high fidelity fluoroscopy simulator using the image-guided surgery toolkit (IGSTK).** In *Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 9036, page 903618. International Society for Optics and Photonics. doi: 10.1117/12.2044112.
- Grupp, R. B., Unberath, M., Gao, C., Hegeman, R. A., Murphy, R. J., Alexander, C. P., Otake, Y., McArthur, B. A., Armand, M., and Taylor, R. H. (2020). **Automatic annotation of hip anatomy in fluoroscopy for robust and efficient 2D/3D registration.** *International Journal of Computer Assisted Radiology and Surgery*, 15(5):759–769. doi: 10.1007/s11548-020-02162-7.
- Haiderbhai, M., Turrubiates, J. G., Gutta, V., and Fallavollita, P. (2019). **Automatic C-arm positioning using multi-functional user interface.** *Canadian Medical and Biological Engineering Society Proceedings*, 42.
- Harris, E. B., Massey, P., Lawrence, J., Rihn, J., Vaccaro, A., and Anderson, D. G. (2008). **Per-cutaneous techniques for minimally invasive posterior lumbar fusion.** *Neurosurgical Focus*, 25(2):E12. doi: 10.3171/FOC/2008/25/8/E12.
- Hou, B., Alansary, A., McDonagh, S., Davidson, A., Rutherford, M., Hajnal, J. V., Rueckert, D., Glocker, B., and Kainz, B. (2017). **Predicting slice-to-volume transformation in presence of arbitrary subject motion.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 296–304. Springer. doi: 10.1007/978-3-319-66185-8\_34.
- Hubbell, J. H. and Seltzer, S. M. (1995). **Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z= 1 to 92 and**

**48 additional substances of dosimetric interest.** Technical report, National Inst. of Standards and Technology-PL, Gaithersburg, MD (United States). Ionizing Radiation Div.

International Atomic Energy Agency. (2021). **Good practices in interventional procedures.** URL <https://www.iaea.org/resources/rpop/health-professionals/interventional-procedures/good-practices-in-interventional-fluoroscopy#2>. [Online; accessed 2021-08-06].

Isensee, F., Jäger, P., Wasserthal, J., Zimmerer, D., Petersen, J., Kohl, S., Schock, J., Klein, A., Roß, T., Wirkert, S., Neher, P., Dinkelacker, S., Köhler, G., and Maier-Hein, K. (2020). **batchgenerators - a python framework for data augmentation.**

Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2021). **nnu-net: a self-configuring method for deep learning-based biomedical image segmentation.** *Nature Methods*, 18(2):203–211. doi: 10.1038/s41592-020-01008-z.

Ishak, B., Younsi, A., Wieckhusen, C., Slonczewski, P., Unterberg, A. W., and Kiening, K. L. (2019). **Accuracy and revision rate of intraoperative computed tomography point-to-point navigation for lateral mass and pedicle screw placement: 11-year single-center experience in 1054 patients.** *Neurosurgical review*, 42(4):895–905. doi: 10.1007/s10143-018-01067-z.

Kausch, L., Thomas, S., Kunze, H., Privalov, M., Vetter, S., Franke, J., Mahnken, A. H., Maier-Hein, L., and Maier-Hein, K. (2020). **Toward automatic C-arm positioning for standard projections in orthopedic surgery.** *International Journal of Computer Assisted Radiology and Surgery*, 15(7):1095–1105. doi: 10.1007/s11548-020-02204-0.

Kausch, L., Scherer, M., Thomas, S., Klein, A., Isensee, F., and Maier-Hein, K. (2021)a. **Automatic image-based pedicle screw planning.** In *Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 11598, page 115981I. International Society for Optics and Photonics. doi: 10.1117/12.2582571.

Kausch, L., Thomas, S., Kunze, H., Norajitra, T., Klein, A., El Barbari, J., Privalov, M., Vetter, S., Mahnken, A. H., Maier-Hein, L., and Maier-Hein, K. (2021)b. **C-arm positioning for spinal standard projections in different intra-operative setting.** *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 352–362. doi: 10.1007/978-3-030-87202-1\_34.

Kazeminia, S., Baur, C., Kuijper, A., van Ginneken, B., Navab, N., Albarqouni, S., and Mukhopadhyay, A. (2020). **GANs for medical image analysis.** *Artificial Intelligence in Medicine*, page 101938. doi: 10.1016/j.artmed.2020.101938.

- 
- Keil, H., Beisemann, N., Swartman, B., Vetter, S. Y., Grützner, P. A., and Franke, J. (2018). **Intra-operative imaging in trauma surgery.** *EFORT open reviews*, 3(10):541–549. doi: 10.1302/2058-5241.3.170074.
- Kim, K. H., Park, J. Y., Kuh, S. U., Chin, D. K., Kim, K. S., and Cho, Y. E. (2013). **Changes in spinal canal diameter and vertebral body height with age.** *Yonsei medical journal*, 54(6):1498–1504. doi: 10.3349/ymj.2013.54.6.1498.
- Kingma, D. P. and Ba, J. (2014). **Adam: A method for stochastic optimization.** *arXiv preprint arXiv:1412.6980*.
- Klein, A., Wasserthal, J., Greiner, M., Zimmerer, D., and Maier-Hein, K. H. (2019). **basic\_unet\_example (v2019.01).**
- Knez, D., Likar, B., Pernuš, F., and Vrtovec, T. (2016). **Computer-assisted screw size and insertion trajectory planning for pedicle screw placement surgery.** *IEEE Transactions on Medical Imaging*, 35(6):1420–1430. doi: 10.1109/TMI.2016.2514530.
- Knez, D., Nahle, I. S., Vrtovec, T., Parent, S., and Kadoury, S. (2018). **Computer-assisted pedicle screw placement planning: Towards clinical practice.** In *IEEE International Symposium on Biomedical Imaging*, pages 249–252. IEEE. doi: 10.1109/ISBI.2018.8363566.
- Knez, D., Nahle, I. S., Vrtovec, T., Parent, S., and Kadoury, S. (2019). **Computer-assisted pedicle screw trajectory planning using ct-inferred bone density: A demonstration against surgical outcomes.** *Medical physics*, 46(8):3543–3554. doi: 10.1002/mp.13585.
- Kordon, F., Maier, A., Swartman, B., and Kunze, H. (2020). **Font augmentation.** In *Bildverarbeitung für die Medizin*, pages 176–182. Springer. doi: 10.1007/978-3-658-29267-6\_36.
- Kosmopoulos, V. and Schizas, C. (2007). **Pedicle screw placement accuracy: A meta-analysis.** *Spine*, 32(3):E111–E120. doi: 10.1097/01.brs.0000254048.79024.8b.
- Krawczyk, B. (2016). **Learning from imbalanced data: open challenges and future directions.** *Progress in Artificial Intelligence*, 5(4):221–232. doi: 10.1007/s13748-016-0094-0.
- Kügler, D., Sehring, J., Stefanov, A., Stenin, I., Kristin, J., Klenzner, T., Schipper, J., and Mukhopadhyay, A. (2020). **i3PosNet: Instrument pose estimation from x-ray in temporal bone surgery.** *International Journal of Computer Assisted Radiology and Surgery*, 15:1137–1145. doi: 10.1007/s11548-020-02157-4.

- Laina, I., Rieke, N., Rupprecht, C., Vizcaíno, J. P., Eslami, A., Tombari, F., and Navab, N. (2017). **Concurrent segmentation and localization for tracking of surgical instruments**. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 664–672. Springer. doi: 10.1007/978-3-319-66185-8\_75.
- Lee, J., Kim, S., Kim, Y. S., and Chung, W. K. (2012). **Optimal surgical planning guidance for lumbar spinal fusion considering operational safety and vertebra – screw interface strength**. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 8(3):261–272. doi: 10.1002/rcs.1413.
- Lehman Jr, R. A., Polly Jr, D. W., Kuklo, T. R., Cunningham, B., Kirk, K. L., and Belmont Jr, P. J. (2003). **Straight-forward versus anatomic trajectory technique of thoracic pedicle screw fixation: A biomechanical analysis**. *Spine*, 28(18):2058–2065. doi: 10.1097/01.BRS.0000087743.57439.4F.
- Levine, M., De Silva, T., Ketcha, M. D., Vijayan, R., Doerr, S., Uneri, A., Vedula, S., Theodore, N., and Siewerdsen, J. H. (2019). **Automatic vertebrae localization in spine ct: a deep-learning approach for image guidance and surgical data science**. In *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10951, page 109510S. International Society for Optics and Photonics. doi: 10.1117/12.2513915.
- Li, M., Liao, S., Reda, F. A., Zhan, Y., Zhou, X. S., and Kleinszig, G. (2019). **Image-based pedicle screw positioning**. US Patent 10,390,886.
- Liao, H., Lin, W.-A., Zhang, J., Zhang, J., Luo, J., and Zhou, S. K. (2019). **Multiview 2d/3d rigid registration via a point-of-interest network for tracking and triangulation**. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 12638–12647. doi: 10.1109/CVPR.2019.01292.
- Liu, W., Wang, Y., Jiang, T., Chi, Y., Zhang, L., and Hua, X.-S. (2020). **Landmarks detection with anatomical constraints for total hip arthroplasty preoperative measurements**. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 670–679. Springer. doi: 10.1007/978-3-030-59719-1\_65.
- Löffler, M. T., Sekuboyina, A., Jacob, A., Grau, A.-L., Scharr, A., El Husseini, M., Kallweit, M., Zimmer, C., Baum, T., and Kirschke, J. S. (2020). **A vertebral segmentation dataset with fracture grading**. *Radiology: Artificial Intelligence*, 2(4):e190138. doi: 10.1148/ryai.2020190138.
- Luther, N., Iorgulescu, J. B., Geannette, C., Gebhard, H., Saleh, T., Tsiouris, A. J., and Härtl, R. (2015). **Comparison of navigated versus non-navigated pedicle screw placement**

- 
- in 260 patients and 1434 screws. *Journal of Spinal Disorders and Techniques*, 28(5):E298–E303. doi: 10.1097/BSD.0b013e31828af33e.
- Maier, A., Syben, C., Lasser, T., and Riess, C. (2019). **A gentle introduction to deep learning in medical image processing.** *Zeitschrift für Medizinische Physik*, 29(2):86–101. doi: 10.1016/j.zemedi.2018.12.003.
- Maier, J., Aichert, A., Mehringer, W., Bier, B., Eskofier, B., Levenston, M., Gold, G., Fahrig, R., Bonaretti, S., and Maier, A. (2018). **Feasibility of motion compensation using inertial measurement in C-arm CT.** In *IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings*, pages 1–3. IEEE. doi: 10.1109/NSSMIC.2018.8824463.
- Manbachi, A., De Silva, T., Uneri, A., Jacobson, M., Goerres, J., Ketcha, M., Han, R., Aygun, N., Thompson, D., Ye, X., et al. (2018). **Clinical translation of the levelcheck decision support algorithm for target localization in spine surgery.** *Annals of biomedical engineering*, 46(10):1548–1557. doi: 10.1007/s10439-018-2099-2.
- Marathakis, P. (2013). **Model pedicle screw.** URL <https://grabcad.com/library/model-pedicle-screw-solidworks-iges>. [Online; accessed 2021-08-30].
- Mason, A., Paulsen, R., Babuska, J. M., Rajpal, S., Burneikiene, S., Nelson, E. L., and Villavicencio, A. T. (2014). **The accuracy of pedicle screw placement using intraoperative image guidance systems: a systematic review.** *Journal of Neurosurgery: Spine*, 20(2): 196–203. doi: 10.3171/2013.11.SPINE13413.
- Matityahu, A., Duffy, R. K., Goldhahn, S., Joeris, A., Richter, P. H., and Gebhard, F. (2017). **The great unknown — a systematic literature review about risk associated with intraoperative imaging during orthopaedic surgeries.** *Injury*, 48(8):1727–1734. doi: 10.1016/j.injury.2017.04.041.
- Matthäus, L., Binder, N., Bodensteiner, C., and Schweikard, A. (2007). **Closed-form inverse kinematic solution for fluoroscopic C-arms.** *Advanced Robotics*, 21(8):869–886. doi: 10.1163/156855307780851957.
- Matthews, F., Hoigne, D. J., Weiser, M., Wanner, G. A., Regazzoni, P., Suhm, N., and Messmer, P. (2007). **Navigating the fluoroscope’s C-arm back into position: An accurate and practicable solution to cut radiation and optimize intraoperative workflow.** *Journal of Orthopaedic Trauma*, 21(10):687–692. doi: 10.1097/BOT.0b013e318158fd42.
- Miao, S., Wang, Z. J., and Liao, R. (2016). **A CNN regression approach for real-time 2D/3D registration.** *IEEE Transactions on Medical Imaging*, 35(5):1352–1363. doi: 10.1109/TMI.2016.2521800.

- Miao, S., Piat, S., Fischer, P., Tuysuzoglu, A., Mewes, P., Mansi, T., and Liao, R. (2018). **Dilated FCN for multi-agent 2D/3D medical image registration**. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.
- Mobbs, R. J., Sivabalan, P., and Li, J. (2011). **Technique, challenges and indications for percutaneous pedicle screw fixation**. *Journal of Clinical Neuroscience*, 18(6):741–749. doi: 10.1016/j.jocn.2010.09.019.
- Molinari, R. W. (2016). **Lumbar pedicle screw placement**. In Vaccaro, A. R. and Todd, A. J., editors, *Spine surgery: tricks of the trade*, pages 129–135. Thieme. doi: 10.1055/b-0034-83489.
- Mrabet, Y. (2012). **Human anatomy planes**. URL [https://commons.wikimedia.org/wiki/File:Human\\_anatomy\\_planes.svg](https://commons.wikimedia.org/wiki/File:Human_anatomy_planes.svg). [Online; accessed 2021-11-07].
- Newell, A., Yang, K., and Deng, J. (2016). **Stacked hourglass networks for human pose estimation**. In *European Conference on Computer Vision*, pages 483–499. Springer. doi: 10.1007/978-3-319-46484-8\_29.
- Nolden, M., Zelzer, S., Seitel, A., Wald, D., Müller, M., Franz, A. M., Maleike, D., Fangerau, M., Baumhauer, M., Maier-Hein, L., et al. (2013). **The Medical Imaging Interaction Toolkit: Challenges and advances**. *International Journal of Computer Assisted Radiology and Surgery*, 8(4):607–620. doi: 10.1007/s11548-013-0840-8.
- Norris, B. L., Hahn, D. H., Bosse, M. J., Kellam, J. F., and Sims, S. H. (1999). **Intraoperative fluoroscopy to evaluate fracture reduction and hardware placement during acetabular surgery**. *Journal of Orthopaedic Trauma*, 13(6):414–417. doi: 10.1097/00005131-199908000-00004.
- Ohba, T., Ebata, S., Fujita, K., Sato, H., and Haro, H. (2016). **Percutaneous pedicle screw placements: accuracy and rates of cranial facet joint violation using conventional fluoroscopy compared with intraoperative three-dimensional computed tomography computer navigation**. *European Spine Journal*, 25(6):1775–1780. doi: 10.1007/s00586-016-4489-1.
- Otake, Y., Armand, M., Armiger, R. S., Kutzer, M. D., Basafa, E., Kazanzides, P., and Taylor, R. H. (2011). **Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: Incorporation of fiducial-based C-arm tracking and GPU-acceleration**. *IEEE Transactions on Medical Imaging*, 31(4):948–962. doi: 10.1109/TMI.2011.2176555.

- 
- Pally, E. and Kreder, H. J. (2013). **Survey of terminology used for the intraoperative direction of C-arm fluoroscopy**. *Canadian Journal of Surgery*, 56(2):109. doi: 10.1503/cjs.015311.
- Patel, V. V., Burger, E., and Brown, C. W. (2010). **Part iii: Thoracolumbar and lumbar**. In *Spine Trauma: Surgical Techniques*. Springer Science & Business Media.
- Payer, C., Urschler, M., Bischof, H., and Štern, D. (2020). **Uncertainty estimation in landmark localization based on gaussian heatmaps**. In *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, pages 42–51. Springer. doi: 10.1007/978-3-030-60365-6\_5.
- Pelozo, J. H. (2016). **Percutaneous lumbar pedicle screw fixation**. In Vaccaro, A. R. and Todd, A. J., editors, *Spine surgery: tricks of the trade*, pages 129–135. Thieme. doi: 10.1055/b-0034-83520.
- Perna, F., Borghi, R., Pilla, F., Stefanini, N., Mazzotti, A., and Chehrassan, M. (2016). **Pedicle screw insertion techniques: An update and review of the literature**. *Musculoskeletal Surgery*, 100(3):165–169. doi: 10.1007/s12306-016-0438-8.
- Rajae, S. S., Bae, H. W., Kanim, L. E., and Delamarter, R. B. (2012). **Spinal fusion in the United States: Analysis of trends from 1998 to 2008**. *Spine*, 37(1):67–76. doi: 10.1097/BRS.0b013e31820cccfb.
- Rajasekaran, S. and Shetty, A. P. (2021). **Section 11, chapter 14: Navigation in spine surgery**. In *Lumbar Spine Textbook*.
- Ramadan, H., Lachqar, C., and Tairi, H. (2020). **A survey of recent interactive image segmentation methods**. *Computational visual media*, pages 1–30. doi: 10.1007/s41095-020-0177-5.
- Rawicki, N., Dowdell, J. E., and Sandhu, H. S. (2021). **Current state of navigation in spine surgery**. *Annals of Translational Medicine*, 9(1). doi: 10.21037/atm-20-1335.
- Reinke, A., Eisenmann, M., Tizabi, M. D., Sudre, C. H., Rädtsch, T., Antonelli, M., Arbel, T., Bakas, S., Cardoso, M. J., Cheplygina, V., Farahani, K., Glocker, B., Heckmann-Nötzel, D., Isensee, F., Jannin, P., Kahn, C. E., Kleesiek, J., Kurc, T., Kozubek, M., Landman, B. A., Litjens, G., Maier-Hein, K., Menze, B., Müller, H., Petersen, J., Reyes, M., Rieke, N., Stieltjes, B., Summers, R. M., Tsaftaris, S. A., van Ginneken, B., Kopp-Scheider, P., Annette amd Jäger, and Maier-Hein, L. (2021). **Common limitations of image processing metrics: A picture story**. *arXiv preprint arXiv:2104.05642*.

- Reiser, E. W., Desai, R., Byrd, S. A., Farber, H., Chi, D., Idler, C. S., and Isaacs, R. E. (2017). **C-arm positioning is a significant source of radiation in spine surgery.** *Spine*, 42(9): 707–710. doi: 10.1097/BRS.0000000000001869.
- Ren, S., He, K., Girshick, R., and Sun, J. (2016). **Faster r-cnn: towards real-time object detection with region proposal networks.** *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149. doi: TPAMI.2016.2577031.
- Riemer, A. (2017). **Scanparameter und Bildqualität.** In *Computertomografie für MTRA/RT*, pages 18–20. doi: 10.1055/b-0037-143974.
- Rikli, D., Goldhahn, S., Blauth, M., Mehta, S., Cunningham, M., Joeris, A., Group, P. S., et al. (2018). **Optimizing intraoperative imaging during proximal femoral fracture fixation – a performance improvement program for surgeons.** *Injury*, 49(2):339–344. doi: 10.1016/j.injury.2017.11.024.
- Rodas, N. L., Bert, J., Visvikis, D., De Mathelin, M., and Padoy, N. (2017). **Pose optimization of a C-arm imaging device to reduce intraoperative radiation exposure of staff and patient during interventional procedures.** In *2017 IEEE International Conference on Robotics and Automation*, pages 4200–4207. IEEE. doi: 10.1109/ICRA.2017.7989483.
- Rodts, G., Jr. (2003). **Percutaneous lumbar pedicle screws: Indications, technique, results.** *Progress in Neurological Surgery*, 16:204–212. doi: 10.1159/000072643.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). **U-net: Convolutional networks for biomedical image segmentation.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer. doi: 10.1007/978-3-319-24574-4\_28.
- Rüedi, T. P., Buckley, R., and Moran, C. (2008). **Wirbelsäule.** In *AO-Prinzipien des Frakturmanagements [AO principles of fracture management]*, pages 605–619. Georg Thieme.
- Russakoff, D. B., Rohlfing, T., Mori, K., Rueckert, D., Ho, A., Adler, J. R., and Maurer, C. R. (2005). **Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration.** *IEEE Transactions on Medical Imaging*, 24(11):1441–1454. doi: 10.1109/TMI.2005.856749.
- Sakaura, H., Ikegami, D., Fujimori, T., Sugiura, T., Mukai, Y., Hosono, N., and Fuji, T. (2019). **Early cephalad adjacent segment degeneration after posterior lumbar interbody fusion: a comparative study between cortical bone trajectory screw fixation and traditional trajectory screw fixation.** *Journal of Neurosurgery: Spine*, 32(2):155–159. doi: 10.3171/2019.8.SPINE19631.

- 
- Samdani, A. F., Ranade, A., Sciubba, D. M., Cahill, P. J., Antonacci, M. D., Clements, D. H., and Betz, R. R. (2010). **Accuracy of free-hand placement of thoracic pedicle screws in adolescent idiopathic scoliosis: how much of a difference does surgeon experience make?** *European Spine Journal*, 19(1):91–95. doi: 10.1007/s00586-009-1183-6.
- Sasso, R. C. and Garrido, B. J. (2007). **Computer-assisted spinal navigation versus serial radiography and operative time for posterior spinal fusion at L5-S1.** *Clinical Spine Surgery*, 20(2):118–122. doi: 10.1097/01.bsd.0000211263.13250.b1.
- Schaller, C., Rohkohl, C., Penne, J., Stürmer, M., and Hornegger, J. (2009). **Inverse C-arm positioning for interventional procedures using real-time body part detection.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 549–556. Springer. doi: 10.1007/978-3-642-04268-3\_68.
- Schatlo, B., Molliqaj, G., Cuvinciuc, V., Kotowski, M., Schaller, K., and Tessitore, E. (2014). **Safety and accuracy of robot-assisted versus fluoroscopy-guided pedicle screw insertion for degenerative diseases of the lumbar spine: A matched cohort comparison.** *Journal of Neurosurgery: Spine*, 20(6):636–643. doi: 10.3171/2014.3.SPINE13714.
- Scherer, M., Kausch, L., Ishak, B., Norajitra, T., Kiening, K., Unterberg, A., Maier-Hein, K., and Neumann, J.-O. (in submission). **Development and validation of an automated planning tool for navigated lumbosacral pedicle screws using a convolutional neural network.** *Journal of Neurosurgery - Spine*.
- Schreiber, J. J., Anderson, P. A., Rosas, H. G., Buchholz, A. L., and Au, A. G. (2011). **Hounsfield units for assessing bone mineral density and strength: A tool for osteoporosis management.** *Journal of Bone and Joint Structures*, 93(11):1057–1063. doi: 10.2106/JBJS.J.00160.
- Schünke, M., Schulte, E., Schumacher, U., Voll, M., and Wesker, K. (2005). **Prometheus: Allgemeine anatomie und bewegungssystem.** *LernAtlas der Anatomie. Stuttgart: Thieme.*
- Shao, Z., Guan, Y., and Tan, J. (2014). **Virtual reality aided positioning of mobile C-arms for image-guided surgery.** *Advances in Mechanical Engineering*, 6:943025. doi: 10.1155/2014/943025.
- Siemens AG. (2013). **The integration statement: Somatom emotion - computed tomography.** URL [https://cdn0.scrvt.com/39b415fb07de4d9656c7b516d8e2d907/1800000002385793/f87e05f60110/ihe\\_euro\\_emotion\\_vb42\\_12\\_13-02385793\\_1800000002385793.pdf](https://cdn0.scrvt.com/39b415fb07de4d9656c7b516d8e2d907/1800000002385793/f87e05f60110/ihe_euro_emotion_vb42_12_13-02385793_1800000002385793.pdf). [Online; accessed 2021-09-03].

- Siemens AG. (2019). **Cios Spin**. URL <https://www.siemens-healthineers.com/at/surgical-c-arms-and-navigation/mobile-c-arms/cios-spin>. [Online; accessed 2021-09-03].
- Simonyan, K. and Zisserman, A. (2014). **Very deep convolutional networks for large-scale image recognition**. *arXiv preprint arXiv:1409.1556*.
- Soltow, E. and Rosenhahn, B. (2015). **Automatic pose estimation using contour information from X-ray images**. In *Image and Video Technology*, pages 246–257. Springer. doi: 10.1007/978-3-319-30285-0\_20.
- Song, H., Kim, M., Park, D., Shin, Y., and Lee, J.-G. (2020). **Learning from noisy labels with deep neural networks: A survey**. *arXiv preprint arXiv:2007.08199*.
- Sun, K., Xiao, B., Liu, D., and Wang, J. (2019). **Deep high-resolution representation learning for human pose estimation**. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 5693–5703. doi: 10.1109/CVPR.2019.00584.
- Taubmann, O., Berger, M., Bögel, M., Xia, Y., Balda, M., and Maier, A. (2018). **Computed tomography**. In *Medical Imaging Systems: An Introductory Guide.*, Lecture notes in computer science Image processing, computer vision, pattern recognition, and graphics, pages 147–189. Springer. doi: 10.1007/978-3-319-96520-8\_8.
- Terunuma, T., Tokui, A., and Sakae, T. (2018). **Novel real-time tumor-contouring method using deep learning to prevent mistracking in X-ray fluoroscopy**. *Radiological Physics and Technology*, 11(1):43–53. doi: 10.1007/s12194-017-0435-0.
- Thies, M., Zäch, J.-N., Gao, C., Taylor, R., Navab, N., Maier, A., and Unberath, M. (2020). **A learning-based method for online adjustment of C-arm cone-beam CT source trajectories for artifact avoidance**. *International Journal of Computer Assisted Radiology and Surgery*, 15(11):1787–1796. doi: 10.1007/s11548-020-02249-1.
- Thomas, S., Isensee, F., Kohl, S., Privalov, M., Beisemann, N., Swartman, B., Keil, H., Vetter, S. Y., Franke, J., Grütznier, P. A., et al. (2019). **Computer-assisted intra-operative verification of surgical outcome for the treatment of syndesmotic injuries through contralateral side comparison**. *International Journal of Computer Assisted Radiology and Surgery*, 14(12):2211–2220. doi: 10.1007/s11548-019-02043-8.
- Toth, D., Cimen, S., Ceccaldi, P., Kurzendorfer, T., Rhode, K., and Mountney, P. (2019). **Training deep networks on domain randomized synthetic X-ray data for cardiac interventions**. In *International Conference on Medical Imaging with Deep Learning*, pages 468–482. PMLR.

- 
- Touchette, M., Newell, R., Anglin, C., Guy, P., Lefaiivre, K., Amlani, M., and Hodgson, A. (2021). **The effect of artificial X-rays on C-arm positioning performance in a simulated orthopaedic surgical setting.** *International Journal of Computer Assisted Radiology and Surgery*, 16(1):11–22. doi: 10.1007/s11548-020-02280-2.
- Tscherne, H. and Blauth, M. (1998). **4 Brust- und Lendenwirbelsäule.** In *Unfallchirurgie-Wirbelsäule*, volume 54, pages pp. 241–280. München: Hans Marseille Verlag.
- Unberath, M., Fotouhi, J., Hajek, J., Maier, A., Osgood, G., Taylor, R., Armand, M., and Navab, N. (2018)a. **Augmented reality-based feedback for technician-in-the-loop C-arm repositioning.** *Healthcare Technology Letters*, 5(5):143–147. doi: 10.1049/htl.2018.5066.
- Unberath, M., Zaech, J.-N., Lee, S. C., Bier, B., Fotouhi, J., Armand, M., and Navab, N. (2018)b. **DeepDRR – a catalyst for machine learning in fluoroscopy-guided procedures.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 98–106. Springer.
- Unberath, M., Zaech, J.-N., Gao, C., Bier, B., Goldmann, F., Lee, S. C., Fotouhi, J., Taylor, R., Armand, M., and Navab, N. (2019). **Enabling machine learning in X-ray-based procedures via realistic simulation of image formation.** *International Journal of Computer Assisted Radiology and Surgery*, 14(9):1517–1528. doi: 10.1007/s11548-019-02011-2.
- Vaccaro, A., Kandziora, F., Fehlings, M., and Shanmughanathan, R. (2021). **AO surgery reference: Percutaneous posterior approach for pedicle screw placement.** URL <https://surgeryreference.aofoundation.org/spine/trauma/thoracolumbar/approach/percutaneous-posterior-approach-for-pedicle-screw-placement#first-editio>. [Online; accessed 2021-07-15].
- Vaccaro, A. R., Kim, D. H., Brodke, D. S., Harris, M., Chapman, J., Schildhauer, T., Routt, M. C., and Sasso, R. C. (2003). **Diagnosis and management of thoracolumbar spine fractures.** *Journal of Bone and Joint Structures*, 85(12):2456–2470. doi: 10.2106/00004623-200312000-00029.
- Vijayan, R., De Silva, T., Han, R., Zhang, X., Uneri, A., Doerr, S., Ketcha, M., Perdomo-Pantoja, A., Theodore, N., and Siewerdsen, J. (2019). **Automatic pedicle screw planning using atlas-based registration of anatomy and reference trajectories.** *Physics in Medicine & Biology*, 64(16):165020. doi: 10.1088/1361-6560/ab2d66.
- Villard, J., Ryang, Y.-M., Demetriades, A. K., Reinke, A., Behr, M., Preuss, A., Meyer, B., and Ringel, F. (2014). **Radiation exposure to the surgeon and the patient during**

- posterior lumbar spinal instrumentation: A prospective randomized comparison of navigated versus non-navigated freehand techniques.** *Spine*, 39(13):1004–1009. doi: 10.1097/BRS.0000000000000351.
- Walthers, E. M. (2017). **11 wirbelsäule.** In *Skelettradiologie: Orthopädie, Traumatologie, Rheumatologie, Onkologie*, pages 418–483. Elsevier Health Sciences.
- Wang, L., Fallavollita, P., Zou, R., Chen, X., Weidert, S., and Navab, N. (2012). **Closed-form inverse kinematics for interventional C-arm X-ray imaging with six degrees of freedom: Modeling and application.** *IEEE Transactions on Medical Imaging*, 31(5): 1086–1099. doi: 10.1109/TMI.2012.2185708.
- Wei, S.-E., Ramakrishna, V., Kanade, T., and Sheikh, Y. (2016). **Convolutional pose machines.** In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4724–4732. doi: 10.1109/CVPR.2016.511.
- Wicker, R. and Tedla, B. (2004). **Automatic determination of pedicle screw size, length, and trajectory from patient data.** In *IEEE Engineering in Medicine and Biology Society*, volume 1, pages 1487–1490. IEEE. doi: 10.1109/IEMBS.2004.1403457.
- Wiesenfarth, M., Reinke, A., Landman, B. A., Eisenmann, M., Saiz, L. A., Cardoso, M. J., Maier-Hein, L., and Kopp-Schneider, A. (2021). **Methods and open-source toolkit for analyzing and visualizing challenge results.** *Scientific Reports*, 11(1):1–15.
- Würfl, T., Hoffmann, M., Christlein, V., Breininger, K., Huang, Y., Unberath, M., and Maier, A. K. (2018). **Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems.** *IEEE Transactions on Medical Imaging*, 37(6):1454–1463. doi: 10.1109/TMI.2018.2833499.
- Xiaozhao, C., Jinfeng, H., Baolin, M., Chongnan, Y., and Yan, K. (2016). **A method of lumbar pedicle screw placement optimization applied to guidance techniques.** *Computer Assisted Surgery*, 21(sup1):142–147. doi: 10.1080/24699322.2016.1240301.
- Zaech, J.-N., Gao, C., Bier, B., Taylor, R., Maier, A., Navab, N., and Unberath, M. (2019). **Learning to avoid poor images: Towards task-aware C-arm cone-beam CT trajectories.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 11–19. Springer. doi: 10.1007/978-3-030-32254-0\_2.
- Zhang, Y., Miao, S., Mansi, T., and Liao, R. (2018). **Task driven generative modeling for unsupervised domain adaptation: Application to X-ray image segmentation.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 599–607. Springer. doi: 10.1007/978-3-030-00934-2\_67.

- 
- Zhang, Z., Luo, P., Loy, C. C., and Tang, X. (2014). **Facial landmark detection by deep multi-task learning**. In *European Conference on Computer Vision*, pages 94–108. Springer. doi: 10.1007/978-3-319-10599-4\_7.
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). **Unpaired image-to-image translation using cycle-consistent adversarial networks**. In *IEEE International Conference on Computer Vision*, pages 2223–2232. doi: 10.1109/ICCV.2017.244.
- Zimmerer, D., Petersen, J., Köhler, G., Wasserthal, J., Adler, T., Wirkert, S., and Klein, A. (2017). **Trixi - Training & Retrospective Insights eXperiment Infrastructure**.



## Own Contributions

This chapter gives an overview of my contributions in distinction to team efforts.

### Own share in data acquisition and data analysis

This thesis was written in the division of Medical Image Computing (MIC) headed by Prof. Dr. Klaus Maier-Hein, who is the primary supervisor for this thesis. Throughout the entire time of my thesis, I was closely collaborating with members of Prof. Dr. Klaus Maier-Hein's group, medical, as well as industrial partners.

The thesis is divided into two parts: **1. Pedicle screw planning for CT-navigated spinal surgery**, **2. C-arm positioning for fluoroscopy-guided spinal surgery**. Part 1 was conducted in close cooperation with the Neurosurgical Clinic in Heidelberg, whereas part 2 was funded by a Siemens cooperation project (L-18563, L-23577) and was carried out in cooperation with Siemens Healthineers, the BG trauma center Ludwigshafen, and the university hospital Marburg.

#### Own share in Data Acquisition.

- 1. Pedicle screw planning for CT-navigated spinal surgery:** The surgical planning data for training and evaluating the proposed method was gathered at the university hospital Heidelberg (Sec. 4.1.1). The data was derived from a consecutive registry of navigated spinal instrumentations (n=1660) performed from 01/2010-12/1018. The dataset was retrospectively collected by PD Dr. med. Moritz Scherer. Expert

annotations were obtained from the surgical planning data, which contained screw planning preferences of several experts, depending on the surgeon who performed the respective intervention. A randomly selected subset of the test dataset (24 cases, 130 screws) was additionally annotated by a second specialized spine surgeon, PD Dr. med. Moritz Scherer, who performed manual screw planning twice at intervals of 8 weeks without considering previous planning results. The screw masks that were derived from the surgical planning data were created by myself, as described in the corresponding chapter. Data preprocessing was performed by myself.

The surgical planning data for the post-operative screw accuracy experiment was collected from the registry of post-operative control CT for screw placement verification between 02/2014 and 06/2016 by Arthur Gubian, who also retrospectively manually aligned screw trajectories with the screw locations in the post-operative control CTs using the commercial system *Stryker Spinemap 3D* (Sec. 5.1.4). Automatic rigid registrations between *navi* and *control* CT were performed by Arthur Gubian using *Medical Imaging Interaction Toolkit (MITK)*. The data was prepared for evaluation by myself.

- 2. C-arm positioning for fluoroscopy-guided spinal surgery:** The high-resolution CTs were collected by Prof. Dr. med. Andreas H. Mahnken from the university hospital Marburg (Sec. 4.2.1). Reference standard projection planes were annotated by myself and Dr. med. Maxim Privalov, validated by PD Dr. med. Sven Vetter from the BG trauma center Ludwigshafen using the *MITK Standard Plane Plugin with DRR Preview* developed by myself. All simulated data was generated by myself, relying on the state of the art DRR simulation framework *DeepDRR* (Unberath et al., 2018b). Reference 3D landmark annotations on a randomly selected subset of the CTs were generated by myself in consultation with PD Dr. med. Sven Vetter and Jan Siad El Barbari from the BG trauma center Ludwigshafen. Implant simulations and data preprocessing was performed by myself.

The specimens studies for validation data acquisition were facilitated by Siemens Healthineers, Erlangen and Rimasys, Köln (Sec. 4.2.2). The acquisitions without metal were performed by Dr.-Ing. Holger Kunze, Dr. sc. hum. Sarina Thomas and myself. The acquisitions with spinal implants, including the vertebra instrumentation, were performed by Dr. Björn Kreher and myself.

The *Swipe App* to measure the clinically acceptable deviations of standard projections was developed by myself (Sec. 5.2.3). The study was conducted by PD Dr. med. Sven Vetter, Dr. med. Maxim Privalov, and Jan Siad El Barbari from the BG trauma center Ludwigshafen. The data was prepared for evaluation by myself.

The inter-/intra-rater study to assess the manual C-arm positioning accuracy and

---

efficiency was designed by myself in consultation with Eric Mandelka from the BG trauma center Ludwigshafen (Sec. 5.2.3). The specimens study was facilitated by the BG trauma center Ludwigshafen and Rimasys, Köln. The manual C-arm positioning was and conducted by PD Dr. med. Sven Vetter and Jan Siad El Barbari and protocolled by Eric Mandelka and myself.

**Own share in Data Analysis.** All methods presented in this thesis were developed and implemented by myself with regular input from the MIC group, in particular Prof. Dr. Klaus Maier-Hein, Dr. sc. hum. Sarina Thomas and my external supervisor Dr.-Ing. Holger Kunze. Regarding part (2), the automatic screw planning and evaluation was performed by myself, whereas the qualitative assessment from a medical view was performed by PD Dr. med. Moritz Scherer.

## Own Publications

In this section, all publications that I was a part of and contributed to during my Ph.D. work are listed. It is subdivided into *First Authorships* and *Co-Authorships*.

### First Authorships - Peer Reviewed International Journal Publications

**Lisa Kausch**, Sarina Thomas, Holger Kunze, Maxim Privalov, Sven Vetter, Jochen Franke, Andreas Mahnken, Lena Maier-Hein, Klaus Maier-Hein (2020). *Toward automatic C-arm positioning for standard projections in orthopedic surgery*. In: International Journal of Computer Assisted Radiology and Surgery, 15 (7), 1095-1105. DOI: 10.1007/s11548-020-02204-0

Moritz Scherer\*, **Lisa Kausch\***, Basem Ishak, Tobias Norajitra, Karl Kiening, Andreas Unterberg, Klaus Maier-Hein\*, Jan-Oliver Neumann\*. *Development and Validation of an Automated Planning Tool for Navigated Lumbosacral Pedicle Screws Using a Convolutional Neural Network*. In submission to Journal of Neurosurgery - Spine.

### First Authorships - Peer Reviewed International Conference Publications

**Lisa Kausch**, Moritz Scherer, Sarina Thomas, André Klein, Fabian Isensee, Klaus Maier-Hein (2021). *Automatic image-based pedicle screw planning*. In: Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, 11598, 115981I. DOI: 10.1117/12.2582571

---

\*These authors contributed equally to this work

**Lisa Kausch**, Sarina Thomas, Holger Kunze, Tobias Norajitra, André Klein, Jan El Barbari, Maxim Privalov, Sven Vetter, Andreas Mahnken, Lena Maier-Hein, Klaus Maier-Hein (2021). *C-arm positioning for spinal standard projections in different intra-operative settings*. In: International Conference on Medical Image Computing and Computer Assisted Interventions, 352-362. DOI: 10.1007/978-3-030-87202-1\_34. MICCAI YOUNG SCIENTIST AWARD

### Co-Authorships

Sarina Thomas, **Lisa Kausch**, Holger Kunze, Maxim Privalov, André Klein, Jan El Barbari, Celia Martin Vicario, Jochen Franke, Klaus Maier-Hein (2021). *Computer-assisted contralateral side comparison of the ankle joint using flat panel technology*. In: International Journal of Computer Assisted Radiology and Surgery, 16 (5), 767-777. DOI: 10.1007/s11548-021-02329-w

Celia Martin Vicario, Florian Kordon, Felix Denzinger, Markus Weiten, Sarina Thomas, **Lisa Kausch**, Jochen Franke, Holger Keil, Andreas Maier, Holger Kunze (2020). *Automatic Plane Adjustment of Orthopedic Intraoperative Flat Panel Detector CT-Volumes*. In: International Conference on Medical Image Computing and Computer Assisted Intervention, 486-495. DOI: 10.1007/978-3-030-59713-9\_47

Darya Trofimova, Tim Adler, **Lisa Kausch**, Lynton Ardizzone, Klaus Maier-Hein, Ulrich Köthe, Carsten Rother, Lena Maier-Hein (2020). *Representing Ambiguity in Registration Problems with Conditional Invertible Neural Networks*. In: NeurIPs Medical Imaging.

Moritz Scherer, **Lisa Kausch**, Andreas Unterberg (2021). *Automated Planning of Navigated Lumbosacral Pedicle Screws Using a Convolutional Neural Network*. In: American Association of Neurological Surgeons. AANS BEST INTERNATIONAL ABSTRACT AWARD

# LISA KAUSCH



## PERSONAL INFORMATION

---

PLACE OF BIRTH: Wolfenbüttel  
DATE OF BIRTH: February 14, 1990  
NATIONALITY: German

## WORK EXPERIENCE

---

MAR 2018 - PRESENT	GERMAN CANCER RESEARCH CENTER, HEIDELBERG Ph.D. student
SEP 2015 - DEC 2017	FRAUNHOFER HEINRICH-HERTZ INSTITUTE, BERLIN Research Scientist
NOV 2014 - JUN 2015	PATTERN RECOGNITION COMPANY, LÜBECK Master Thesis Project
APR 2014 - OCT 2014 JUN 2013 - SEP 2013	PHILIPS MEDICAL SYSTEMS DMC Research Intern
SEP 2011	FRAUNHOFER MEVIS, LÜBECK Research Intern

## EDUCATION

---

MAR 2018 - PRESENT	PH.D. Medical Faculty, <b>Ruprecht-Karls-University Heidelberg</b>
AUG 2012 - JUN 2015	MASTER OF SCIENCE in MATHEMATICS IN MEDICINE AND LIFE SCIENCES Specialization: Image Processing, <b>Universität zu Lübeck</b> Final grade: 1.2
AUG 2012 - MAI 2013	10-month study visit at the <b>Technical University Tampere</b> , Finland
OCT 2009 - JUL 2012	BACHELOR OF SCIENCE in MATHEMATICS IN MEDICINE AND LIFE SCIENCES <b>Universität zu Lübeck</b> Final grade: 1.3
JUN 2009	ABITUR <b>Große Schule Wolfenbüttel</b> Final grade: 2.0

## SCHOLARSHIPS AND AWARDS

---

- 2021 | MICCAI YOUNG SCIENTIST AWARD  
Recognizes the highest quality papers that are first-authored by young scientists at the main MICCAI conference
- 2021 | MICCAI STUDENT TRAVEL AWARD  
Rewarding the best (such as highest scoring) papers at the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) first authored by early career researchers
- 2021 | AMERICAN ASSOCIATION OF NEUROLOGICAL SURGEONS (AANS) BEST INTERNATIONAL ABSTRACT AWARD
- SEP 2015 - SEP 2017 | FRAUNHOFER TALENTA START SCHOLARSHIP
- JAN 2013 | PHILIPS BEST-BACHELOR PRICE

## FIRST AUTHORSHIPS

---

- 2021 | L. Kausch, S. Thomas, H. Kunze, T. Norajitra, A. Klein, J. El Barbari, M. Privalov, S. Vetter, A.S. Mahnken, L. Maier-Hein, K. Maier-Hein. *C-arm positioning for spinal standard projections in different intra-operative settings*. International conference on medical image computing and computer assisted interventions, 352-362.
- 2021 | L. Kausch, M. Scherer, S. Thomas, A. Klein, F. Isensee, K. Maier-Hein. *Automatic image-based pedicle screw planning*. Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, 11598, 1159811.
- 2020 | L. Kausch, S. Thomas, H. Kunze, M. Privalov, S. Vetter, J. Franke, A. Mahnken, L. Maier-Hein, K. Maier-Hein. *Toward automatic C-arm positioning for standard projections in orthopedic surgery*. International journal of computer assisted radiology and surgery, 15 (7), 1095-1105.
- 2017 | L. Kausch, A. Hilsmann, P. Eisert: *Template-based 3D non-rigid shape estimation from monocular image sequences*. Proceedings of the conference on Vision, Modeling and Visualization, 37-44.

## LANGUAGES

---

- GERMAN: Native proficiency  
ENGLISH: Professional proficiency  
FRENCH: Elementary proficiency  
FINISH: Elementary proficiency

## PROGRAMMING

---

C++, PYTHON, MATLAB, MATHEMATICA, PYTORCH, GIT, MITK

## TRAINING

---

AGILE PROJECT MANAGEMENT, GIMP, BODY LANGUAGE, TIME MANAGEMENT

## Acknowledgement

There are many people who accompanied and supported me during the development of this thesis and whom I would like to acknowledge in this section. First of all, I'm very grateful to Prof. Dr. Klaus Maier-Hein for offering me the possibility to work in this very inspiring research environment, for his support, guidance, trust, and optimism throughout the entire time.

Further, I would like to thank the whole CAMIC team; it was a pleasure and honor for me to work in this multi-disciplinary team of scientists with plenty of opportunities to develop my personal and scientific skills.

Further, I would like to thank my thesis advisory committee members: Dr. sc. hum. Sarina Thomas and Dr.-Ing. Holger Kunze who accompanied me from the beginning. I'm really happy I had the possibility to discuss with both of you in our weekly meetings and learn from your experience which often helped me think in different directions when being stucked. Especially thankful I'm to Dr. Sc. Sarina Thomas, for accompanying me throughout the whole journey with endless support even after our paths started to separate. I'm very happy we met, enjoyed working together a lot, and hope our paths cross again in the future.

I'm thankful for the possibility and the experience I gained by working in an interdisciplinary project with industrial and medical partners. Therefore, I would like to acknowledge the involved partners: Dr.-Ing. Holger Kunze, Dr. Alexander Gemmel, Dr. Björn Kreher, and Dr. Gerhard Kleinszig from Siemens Healthineers for partly funding the project, providing technical support, and facilitating and accompanying the specimens studies. Further, the clinicians from the BG trauma center in Ludwigshafen, PD Dr. Sven Vetter, PD Dr. Jochen Franke, Dr. Maxim Privalov, Jan Siad El Barbari, Eric Mandelka, who supported me with medical advice, gave me insights into the OR, accompanied the

specimens study, and provided valuable feedback during our regular milestone meetings. In addition, Prof. Dr. med. Andreas Mahnken from the university hospital Marburg who agreed to support my project by providing medical datasets. Last but not least, PD Dr. Moritz Scherer and his colleagues from the neurosurgical clinic in Heidelberg for the cooperation and inspiring knowledge exchange between clinical and technical partners. Further, I want to thank all members of my internal thesis advisory board for their valuable, constructive feedback: Tobias, Matthias, David, Sarina, and especially to André Klein, for being my tutor on my first day, always offering support if needed, and making sure I stay on track. Thanks also to all my proofreaders who helped me improve the quality of my thesis, Sarina, André, Leo, David, Matze, Tobi, Gregor, Silvia, Sarah, Yannick, Klaus, Sven, Jan, Moritz, and also to my former colleagues Clemens, Philipp, and Anna from Fraunhofer HHI.

In the end, I would like to thank my family and friends for their support and trust in me. Thanks for listening and helping me out whenever needed! Special thanks go to Leo, my personal life coach and soul mate - at least most of the time :P - for always encouraging me and reminding me in stressful times what is really important in life; truly a very important person I wish everyone in life! And to my parents and sister Hannah without whom I would not be who I am today. Thanks for all the lovely memories and many more to come! :)

If I missed someone in this section, that is on me, but know that I appreciate every little piece of help.

## Eidesstattliche Versicherung

# Statutory Declaration

1. Bei der eingereichten Dissertation zu dem Thema *Robust Deep Learning for Computer-Assisted Spinal Surgery* handelt es sich um meine eigenständig erbrachte Leistung.

*I herewith formally declare that I have written the submitted dissertation *Robust Deep Learning for Computer-Assisted Spinal Surgery* independently.*

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

*I did not use any third party support except for the quoted literature and other sources mentioned in the text. Content from other work, either literally or in content, has been declared as such.*

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

*The thesis has not been submitted to any examination body in this, or similar, form.*

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

*I confirm the correctness of the aforementioned declarations.*

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

*I am aware of the legal consequences of this declaration. To the best of my knowledge I have told the pure truth and not concealed anything.*

Heidelberg, 09.11.2021

---

Lisa Kausch