

Article

Multi-Stage Platform for (Semi-)Automatic Planning in Reconstructive Orthopedic Surgery

Florian Kordon ^{1,2,3,*} , Andreas Maier ^{1,2} , Benedict Swartman ⁴, Maxim Privalov ⁴, Jan Siad El Barbari ⁴ and Holger Kunze ^{1,3} 

¹ Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nuremberg, 91058 Erlangen, Germany; andreas.maier@fau.de (A.M.); holger.hk.kunze@siemens-healthineers.com (H.K.)

² Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander University Erlangen-Nuremberg, 91052 Erlangen, Germany

³ Advanced Therapies, Siemens Healthcare GmbH, 91031 Forchheim, Germany

⁴ Department for Trauma and Orthopaedic Surgery, BG Trauma Center, Ludwigshafen, 67071 Ludwigshafen, Germany; benedict.swartman@bgu-ludwigshafen.de (B.S.); maxim.privalov@bgu-ludwigshafen.de (M.P.); jan.elbarbari@bgu-ludwigshafen.de (J.S.E.B.)

* Correspondence: florian.kordon@fau.de

Abstract: Intricate lesions of the musculoskeletal system require reconstructive orthopedic surgery to restore the correct biomechanics. Careful pre-operative planning of the surgical steps on 2D image data is an essential tool to increase the precision and safety of these operations. However, the plan's effectiveness in the intra-operative workflow is challenged by unpredictable patient and device positioning and complex registration protocols. Here, we develop and analyze a multi-stage algorithm that combines deep learning-based anatomical feature detection and geometric post-processing to enable accurate pre- and intra-operative surgery planning on 2D X-ray images. The algorithm allows granular control over each element of the planning geometry, enabling real-time adjustments directly in the operating room (OR). In the method evaluation of three ligament reconstruction tasks effect on the knee joint, we found high spatial precision in drilling point localization ($\epsilon < 2.9$ mm) and low angulation errors for k-wire instrumentation ($\epsilon < 0.75^\circ$) on 38 diagnostic radiographs. Comparable precision was demonstrated in 15 complex intra-operative trauma cases suffering from strong implant overlap and multi-anatomy exposure. Furthermore, we found that the diverse feature detection tasks can be efficiently solved with a multi-task network topology, improving precision over the single-task case. Our platform will help overcome the limitations of current clinical practice and foster surgical plan generation and adjustment directly in the OR, ultimately motivating the development of novel 2D planning guidelines.

Keywords: computer-assisted surgery; surgical planning; reconstructive orthopedic surgery; ligament reconstruction; deep learning; multi-task learning; X-ray images



Citation: Kordon, F.; Maier, A.; Swartman, B.; Privalov, M.; El Barbari, J.S.; Kunze, H. Multi-Stage Platform for (Semi-)Automatic Planning in Reconstructive Orthopedic Surgery. *J. Imaging* **2022**, *8*, 108. <https://doi.org/10.3390/jimaging8040108>

Academic Editors: Terry Peters and Elvis C.S. Chen

Received: 28 February 2022

Accepted: 8 April 2022

Published: 12 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Joint injuries and the following structural instability are challenging patterns in orthopedics, usually caused by events of dislocation or unnatural motion and tears of restraining ligaments [1–4]. The destabilizing event comes with various implications for the patient's health and well-being, causing significant joint pain, functional impairment, symptomatic giving way, osteochondral lesions, or damage to the articular cartilage [3,5–13]. For particularly severe lesions, the biomechanically correct joint function is restored by surgical reconstruction to relieve the patient's symptoms and restore long-term stability [1,7,12,14–20].

An essential instrument to elevate the safety and accuracy of surgical reconstruction is the careful planning and outcome verification on 2D image data. Fundamentally, a surgical plan involves interrelating several salient features of the patient anatomy, such as osseous

landmarks and anatomical axes, and performing a geometric construction for the underlying task. For example, reconstruction surgery of the Medial Patellofemoral Ligament (MPFL) necessitates a precise fixation of the substitute graft on the femoral surface. Specifically, the physiological correct ligament insertion point that guarantees joint isometry can be approximated by the construction of the Schoettle Point [21] on a 2D true-lateral radiograph. This construction builds on two anatomical keypoints on the distal femur surface and a subsection of the posterior cortex of the femoral shaft to ultimately localize the optimal point for tunnel drilling. Another application is the careful placement and angulation of k-wires inserted into the patient anatomy for transtibial Posterior Cruciate Ligament (PCL) reconstruction surgery, guiding the subsequent tunnel drilling [22–25]. High precision in the tunnel's placement is essential to prevent any defects in joint biomechanics after reconstruction and unnecessary graft abrasion [26]. It also protects from early breakout of the bone cortex with the risk of damaging nearby vascular structures [27–31].

For such reconstructive applications, surgical planning promises elevated treatment safety and standardized assessment with a physiologically-grounded rationale. However, several clinical and technical challenges complicate the plan's intra-operative application.

1. Most planning methods rely on the combination of salient proxy structures. This indirect description implies a greater number of error sources, which translates to higher observer-variability if the planning is done manually;
2. Certain surgical planning can have a high level of geometric complexity. Sufficiently precise manual execution is only possible with tailored software tools or otherwise requires great amounts of time and labor;
3. The ability to register pre-operative planning and intra-operative live data is highly complex due to the variable configuration of the joint and arbitrary relation between patient, table, and imaging system;
4. Ad-hoc modifications of the surgical plan are essential to compensate for motion during the intervention;
5. Manual interaction with a computer-assisted planning system is undesirable due to the surgery's sterile setting. At the same time, the planning system should offer granular controls to correct each construction step with real-time visualization.

We propose a versatile platform for (semi-)automatic 2D planning and verification in reconstructive orthopedic surgery that addresses most of the above challenges. Our method involves three sequential stages: (i) First, a deep learning system detects the relevant anatomical structures for the underlying surgical task on 2D image data. The detection problem for each type of anatomical feature is solved in parallel using a multi-task learning (MTL) scheme; (ii) Second, the identified structures are transferred to geometric representations, such as keypoint coordinates and line segments. If needed, post-processing can be used to derive additional information required by the planning geometry; (iii) Third, the objects are combined and interrelated using a pre-defined blueprint describing the planning geometry and calculation of the planning result. By mimicking each manual step of the clinically established procedure, we align the automatic planning process to clinical practice as close as possible. This interpretable relation between anatomy and planning geometry reduces the hurdle of acceptance that the practitioner must overcome. Additionally, if needed, the segmentation of the relevant bone can be used for a registration of the planning result on subsequent live images. This allows an easier transfer of the approach into an intra-operative workflow.

We evaluate our approach on three kinds of ligament reconstruction surgery on the knee joint (Figure 1): (i) Localization of the optimal graft fixation point (Schoettle Point) for tunnel drilling in MPFL reconstruction surgery [21]; (ii) Localization of anteromedial (AM) and posterolateral (PL) bundle attachment points on the femur and tibia for Anterior Cruciate Ligament (ACL) reconstruction [32–34]; (iii) Finding the optimal guidewire orientation and insertion point for transtibial PCL reconstruction [23,24]. We study different strategies for separating the anatomical features into individual tasks for each application and evaluate different levels of parameter sharing using multi-head, multi-decoder, and

single-task network topologies. Furthermore, we analyze a shared inductive bias across all three applications and compare the performance to the single-application case.

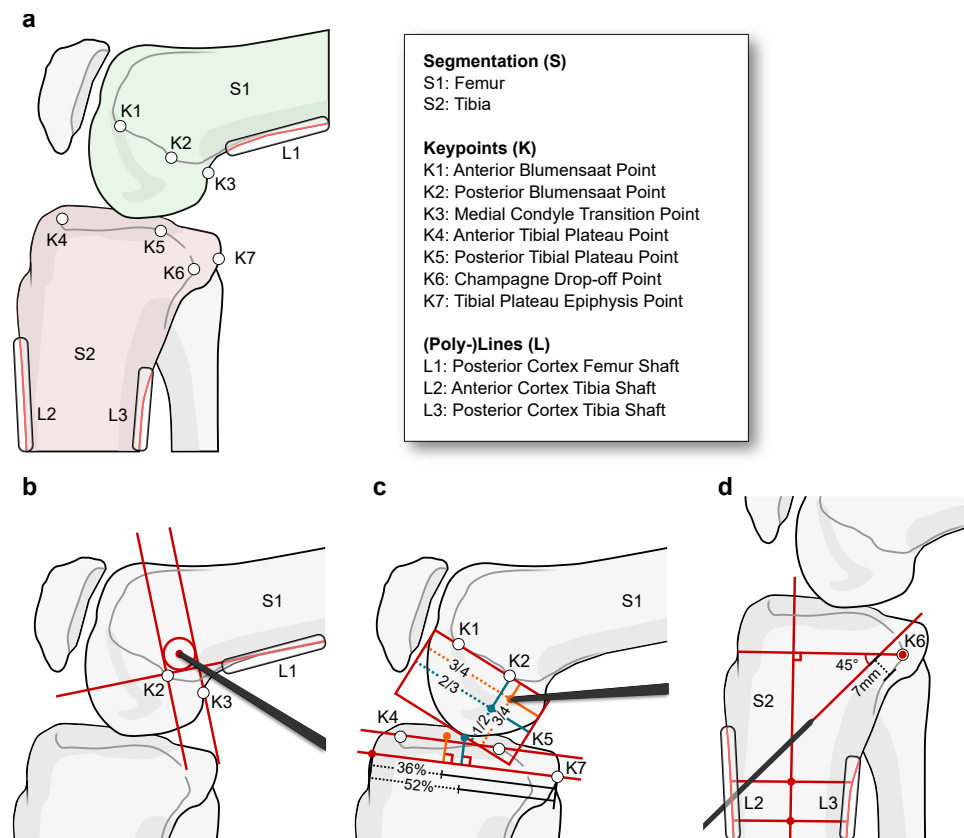


Figure 1. Illustration of the three considered surgical planning guidelines for ligament reconstruction on the knee joint. (a) Description and localization of relevant anatomical structures for all plans. (b) Localization of the anatomic femoral insertion site for Medial Patellofemoral Ligament (MPFL) reconstruction surgery. The construction builds upon the clinical study by Schoettle et al. [21]. (c) Identification of the surface attachment points for the anteromedial (AM; orange) and posterolateral (PL; teal) bundles for Anterior Cruciate Ligament (ACL) reconstruction surgery. The planning follows the quadrant method by Bernard et al. [32] and Stäubli/Rauschnig [34] for the femur and tibia, respectively. (d) Detection of the optimal drill path angulation and position for transtibial reconstruction of the Posterior Cruciate Ligament (PCL) according to Johannsen et al. [23].

2. Related Work and Contribution

2.1. Image-Based Surgical Planning in Orthopedics and Traumatology

Prior research generally puts image-based surgical planning as a central component of computer-assisted orthopedic surgery and a driver for increased safety and minimal invasiveness of surgical intervention [35–37]. We refer the reader to comprehensive reviews covering general concepts and applications in robotics [38–43], navigation [44–46], surgical training [47], and augmented reality in orthopedic surgery [48].

The overall process of image-based surgical planning can be structured into four main components: (1) (Interactive) Visualization, (2) Modeling, (3) Analysis, and (4) Plan generation [49]. Visualization addresses intuitive ways of enriching the original image with relevant numeric and geometric information for the application [48]. This information includes the rotational and positional relation of joints [50], anatomic landmarks [51,52], and optimal trajectories for instrumentation and implant positioning [53–57]. Modeling covers the selection and mathematical representations of these data, allowing subsequent optimization of the planning solution, i.e., numerical and geometric analysis. The academic community has explored these two steps mainly by conventional image analysis [36], statis-

tical shape modeling and machine learning [58], and, more recently, via deep learning [59]. The prevailing representation and analysis methods are semantic segmentation of the target surface in 3D, landmarking of salient regions using numerical keypoints, the fusion of image information with simulated joint kinematics [39,60], and registration using camera data, markers, or anatomic surface information [48].

Despite the wide adoption of pre-operative plans in various types of orthopedic surgery, their application to trauma interventions, such as ligament reconstruction, is limited [35,37,48]. For these cases, diagnostic 3D-scans are rarely used and require complex 3D/2D registration of localized intra-operative fluoroscopic data with the pre-operative volume. On the other hand, planning on diagnostic 2D radiographs has limited value in the operating room due to limitations in patient positioning and the arbitrary relation between the patient, table, and imaging device. In these scenarios, intra-operative planning on 2D image data become an increasingly attractive alternative [21,61].

2.2. Multi-Task Learning and Task Weighting

In MTL, data samples from multiple related tasks are used together to build a learning system with shared inductive bias [62]. By leveraging similarities in data and their underlying formation process, the goal is to discover more generalized and potent features that ultimately lead to better performance than independent single-task learning (STL) inference [62]. For feature-based MTL, the shared inductive bias is typically realized by soft or hard parameter sharing [62–65]. In the soft case, the parameters for each task are separated but constrained using a distance penalty. In hard parameter-sharing, the interaction between tasks is structured by using a meta learner with shared parameters and individual learners with task-specific parameters [66].

For most MTL problems, the update policy for the meta-learner parameters should be designed to prevent any dominance of certain tasks over others. For task balancing, several weighting heuristics have been proposed, most of them being based on a weighted summation of the task-specific losses [67]. Gradient normalization (GradNorm) dynamically approximates the gradient magnitude of the shared representation w.r.t. for each task and aims to level the individual training rates [68]. Liu et al. [69] simplify this idea by directly balancing the task losses' rate of change over training time without considering gradient magnitude. Kendall et al. [70] propose task uncertainty as a proxy for each task's weight, whereas Sener and Koltun [71] formulate MTL as a multi-objective optimization based on Pareto-optimal updates. Multiple solutions consider analysis and manipulation of individual gradient directions. Yu et al. [72] project the task gradients onto the normal plane of conflicting tasks, thus avoiding opposing gradient directions and establishing a stable update trajectory. Wang et al. [73] update this formulation by explicitly adapting gradients for cases of positive similarity to better capture the diverse and often highly complex task interactions. A different implementation of this rationale is probabilistic masking of gradients with a conflicting sign using a drop-out-style operation [74], or weighting the tasks such that their aggregated gradient has equal projections onto individual tasks [75].

We use GradNorm task balancing [68] in this work as it shows competitive performance for more related tasks and is easy to implement at little to no computational overhead.

2.3. Contribution

Our method builds on previous work [50,54,55] and extends them in several ways:

1. This work establishes a multi-stage workflow that covers all necessary steps for image-based surgical planning on 2D X-ray images. The workflow is designed to mimic the clinically-established manual planning process, enabling granular control over each anatomical feature contributing to the planning geometry;
2. We evaluate the model for three trauma-surgical planning applications on both diagnostic as well as intra-operative X-ray images of the knee joint. The numeric results match clinical requirements and encourage further clinical evaluation;

3. We empirically show that the detection of anatomical landmarks benefits from a MTL setting. We confirm that explicit task weighting significantly reduces the landmark localization error, and illustrate that a multi-head network topology achieves similar performance to task-specific decoders, which are computationally much more expensive.
4. Our study demonstrates that sharing tasks across anatomically related applications does not significantly improve performance compared to the single-application variant.

2.4. Article Structure

In Section 3, we present the multi-stage algorithm for (semi-)automatic orthopedic surgery planning and provide details on each step. Furthermore, we describe the different MTL schemes considered in this work, their training protocols, the pre- and intra-operative datasets used, and finally the evaluation metrics for reporting the results. In Section 4, we pose several research questions that are investigated through a series of experiments. After that, we discuss the observations and findings in Section 5 and summarize our contribution in Section 6.

3. Materials and Methods

3.1. (Semi-)Automatic Workflow for 2D Surgical Planning

3.1.1. Multi-Stage Planning Algorithm

In the following, we describe a multi-stage algorithm that automatically models, analyzes, and visualizes a 2D orthopedic surgery plan. The general workflow is illustrated in Figure 2.

Before any geometric processing, a planning blueprint for the target surgical application needs to be defined, describing the necessary sequential construction steps. Based on this blueprint, we can derive the relevant anatomical features and group them according to the underlying task type. Furthermore, any post-processing that builds upon these features can be defined, e.g., the calculation of intersection points, or the extraction of relevant parts of a bone outline. In this work, we focus on three main feature and spatial task types (excluding any secondary features generated by post-processing) that are employed by most surgical planning procedures:

1. Semantically coherent regions. Segmentation of connected regions that share certain characteristics, mostly bones and tools;
2. Anatomical keypoints. Point-like landmarks that pinpoint features of interest on the bone surface;
3. Elongated structures. Straight and curved lines that describe edges, ridges, or that refer to indirect features, such as anatomical axes.

After task configuration, in Stage A), an MTL algorithm extracts all relevant anatomical structures in parallel, sorted by their task type. Second, in Stage B), the features are translated from their original inference encoding to geometrically interpretable objects/descriptors. For example, the heatmap encoding of a keypoint is converted to its geometric equivalent of x and y coordinates, using the spatial argmax operation or a least-squares curve fitting. Stage C) finally describes the automatic processing of the individual planning steps and its overlay on the input image. The user can interact with the planning result to correct any misalignment, and the planning is updated in real-time. In the following, we provide details on the individual stages.

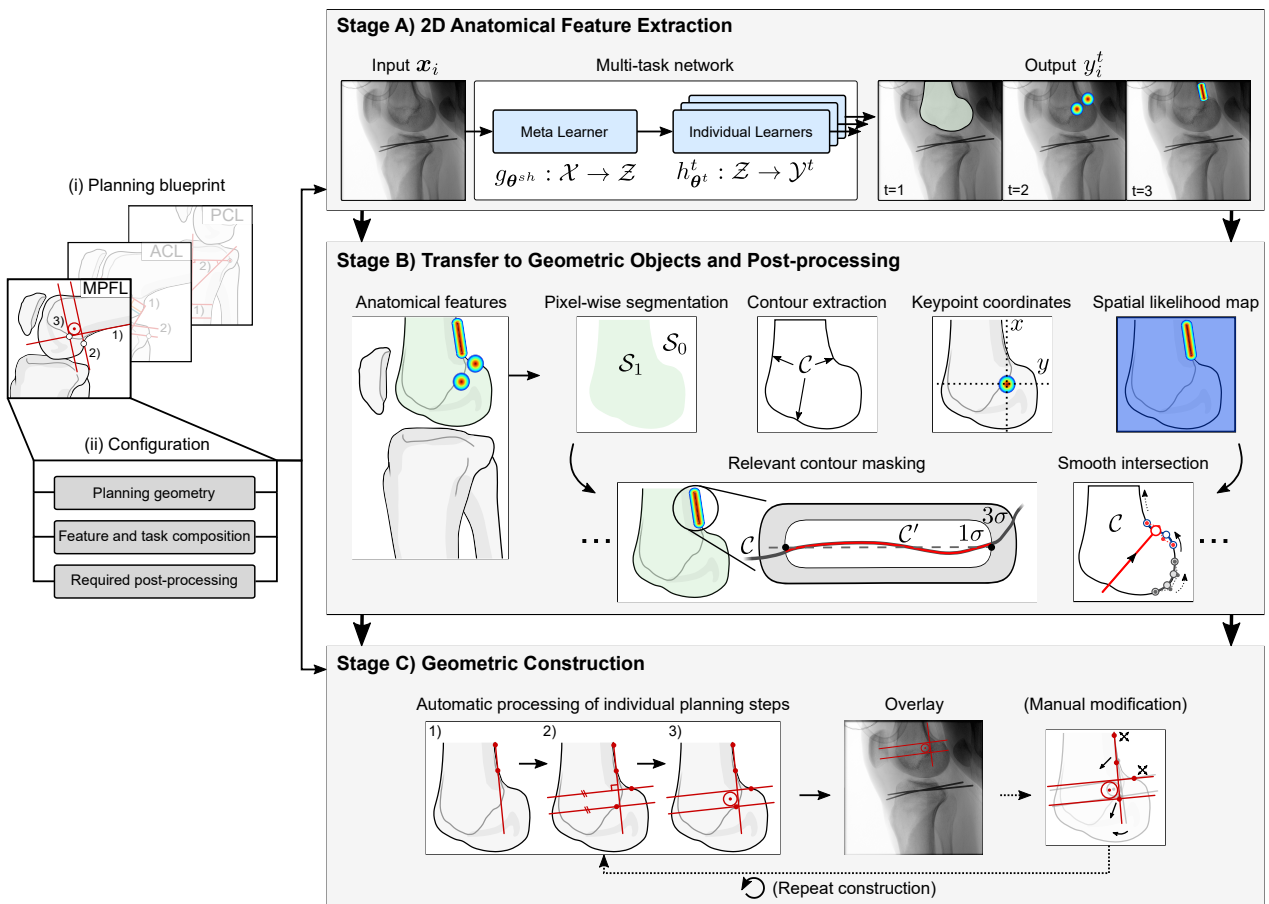


Figure 2. Overview of the automatic planning workflow. After selecting a planning blueprint and a corresponding configuration, the target X-ray image is processed in three consecutive stages. After extraction the 2D anatomical features with a multi-task network, they are transformed to their corresponding geometric descriptors. The individual construction steps defined in the planning blueprint are automatically calculated by interrelating the extracted geometry (here, steps 1–3 in Stage C)). If needed, the user can manually modify the planning by adjusting individual control points upon real-time updates of the complete plan.

3.1.2. Stage A) MTL for Joint Extraction of Anatomical Features

Following the MTL paradigm, we want to find a functional mapping $f_{\theta}^t : \mathcal{X} \rightarrow \mathcal{Y}^t$ from the input space \mathcal{X} to multiple task solution spaces $\{\mathcal{Y}^t\}_{t \in [T]}$ [62,71]. θ is a set of trainable parameters and T is the number of tasks. The data that are required to approximate this mapping in a supervised learning manner are characterized by M datapoints $\{x_i, y_i^1, \dots, y_i^T\}_{i \in [M]}$, where y_i^T is the individual label for task t .

We further consider a hard parameter-sharing strategy and split the parameters θ into a single set of shared parameters θ^{sh} and task-specific parameter sets $\{\theta^t\}_{t \in [T]}$ [66]. Specifically, we define a meta learner $g_{\theta^{sh}} : \mathcal{X} \rightarrow \mathcal{Z}$ and individual learners for each task $h_{\theta^t}^t : \mathcal{Z} \rightarrow \mathcal{Y}^t$ that by their composition $h_{\theta^t}^t \circ g_{\theta^{sh}}$ yield function

$$f_{\theta^{sh}, \theta^t}^t : \mathcal{X} \rightarrow \mathcal{Y}^t. \quad (1)$$

This learning scheme allows to leverage dependencies and similarities between the tasks, and to retain task-specific discriminative capacity. Specifically, for the extraction of anatomical features, the salient structures used for localizing certain landmarks may be of similar importance when trying to solve a segmentation task of the complete bone. In addition, most structures relevant for surgical planning share the same underlying anatomical processes. For the optimization of each task with gradient descent, a distinct

loss function $\mathcal{L}^t(\cdot, \cdot) : \mathcal{Y}^t \times \mathcal{Y}^t \rightarrow \mathbb{R}_0^+$ is used. Although each loss function dictates the updates for the individual learners, the parameter updates for the meta learner are derived by their weighted linear combination. The general optimization problem can be written as

$$\min_{\theta^{sh}; \theta^1, \dots, \theta^T} \sum_{t=1}^T w^t \left[\frac{1}{M} \sum_{i=1}^M \mathcal{L}^t \left(f_{\theta^{sh}, \theta^t}^t(x_i), y_i^t \right) \right]. \tag{2}$$

w_t is a scalar weight for the loss function for the t th task.

Although the choice for the task’s loss function in theory only depends on the task type, it should ideally be based on a spatial representation of the anatomical structure of interest. Such representation exploits dependencies in spatial appearance and a closer alignment of the feature topology. For example, semantic segmentation and heatmap-based landmark regression rely on a strong pixel-wise correlation between input and target. In contrast, a numerical regression task where the x and y coordinates of a landmark are approximated with a fully-connected layer typically exhibits a weaker spatial coupling between the learned features and the numerical coordinates. For this reason, we use spatial representations for the tasks considered in this work (Table 1).

Table 1. Mathematical representation of the relevant anatomical structures that allow subsequent optimization and geometric analysis.

Anatomical Structure	Spatial Representation
Semantically coherent regions	Pixel-wise and multi-label segmentation. The multi-label aspect allows for overlap-aware segmentation, e.g., in areas between bones or metal implants.
Anatomical keypoints	Individual heatmaps/activation maps. For that purpose, a multivariate Gaussian distribution with its mean at the keypoint coordinate and a predefined standard deviation is sampled.
Elongated structures	Line-symmetric heatmap/activation map. The distance to the line segment or the axis of interest is evaluated and transformed using a Gaussian function.

The derivation of the weights w_t , which ultimately steer the task’s relative priority, is a delicate yet decisive step. In general, we can distinguish two variants: (i) Constant weights, and (ii) Dynamic weights which are adapted during optimization. At first glance, constant weights are easily interpretable and can be used to incorporate prior knowledge to assign relative importance. However, since these weights act as multiplicative factors of the gradient magnitudes, which might vary significantly across subsequent backward passes, it is difficult to identify a suitable and stable weight composition [54]. In addition, different magnitudes of the individual loss functions need to be accounted for to prevent dominance of some tasks over others. These downsides can be mitigated by using dynamic weighting policies, which compensate for fluctuations in the tasks’ relative standings. They, however, come with the disadvantage of more elaborate update algorithms, or, depending on the particular policy and task composition, might lead to inferior total performance.

For our particular problem scenario, we consider the following weighting strategies.

1. Uniform (constant). All tasks are weighted uniformly: $w^t = 1$.
2. Balanced relative learning rates (dynamic). Gradient normalization by Chen et al. [68] to ensure balanced training rates of all tasks, i.e., approximately equally-sized update steps for each task: $w^t = \text{GradNorm}(t)$.

We further consider different variants of model topology that differ in proportion between shared and task-specific parameters. Indeed, the shared inductive bias’s capacity greatly influences the types of features learned in the meta learner. As the task-specific

complexity shrinks, more general features need to be extracted, leading to superior solutions for similar tasks that benefit from each other or inferior solutions for conflicting tasks. We analyze this trade-off using two model topologies that facilitate different levels of parameter sharing and compare them to a single-task learning baseline.

1. Single-task baseline. For comparison, a STL baseline of independent encoder-decoder structures is optimized. Here, no parameters are shared;
2. Multi-head topology. Both the encoder and decoder parameters are shared between the tasks in this variant. The output of the decoder is fed into task-specific prediction heads, which involve significantly less dedicated parameters and, thus, require a multi-purpose feature decoding. We argue that such a constrained decoder might benefit learning for highly similar tasks;
3. Multi-decoder topology. After feature extraction in a shared feature encoder, the latent representation is used as input for the task-specific decoders and prediction heads. In other words, an abstract representation has to be found that serves the reconstruction for different kinds of tasks.

3.1.3. Stage B) Extraction of Geometric Objects and Post-Processing

In Stage A), the multi-task network infers all anatomical features via specific spatial encodings. These encodings ensure an efficient optimization of the MTL scheme but cannot be directly used for the geometrical construction defined in the planning blueprint. For that reason, the feature representations are subsequently transformed to more general descriptors using a common set of operations shared between different planning applications.

For the three main tasks (Section 3.1.2), the representation transformations are as follows. The pixel-wise segmentation on the input image is interpreted as subsets \mathcal{S}_k with $k \in \{0, \dots, K\}$ that represent coherent regions of interest, e.g., bones, implants, soft tissue, or background. Due to the multi-label definition, the subsets do not have to be disjoint, which allows to identify region overlap. This is an important trait in the context of transmissive imaging, which assumes an additive model of the pixel intensities and, thus, perceivable superimposition. The segmentation results are further represented as closed polygons, facilitating further spatial analysis and shape relations. The inferred heatmaps directly encode the positional likelihood of the anatomical keypoint. The equivalent x and y coordinates are extracted by a spatial argmax operation or by a least-squares curve fitting in case predicted heatmaps suffer from large local intensity variation. The heatmap representation of an elongated structure is converted to a polyline by the following steps. First, the heatmap is interpreted as a 2D point cloud, and each point is assigned an importance weighting based on the predicted likelihood. By selecting the points with a weight above the threshold of 1σ , we extract only the most descriptive points for the target structure. Finally, we obtain the 2D polyline by interconnecting a set of equidistant points sampled from a smooth parametric B-spline which interpolates the selected points. If we are interested in a particular region of an object's contour \mathcal{C} , we alternatively interpret the structure's heatmap as a weighted mask that reduces the contour to a certain subsection $\mathcal{C}' \subset \mathcal{C}$.

3.1.4. Stage C) Geometric Construction of Individual Planning Steps

Once the geometric descriptions and groupings are created, they are interrelated based on the planning geometry provided in the planning blueprint (Appendix B). According to the manual workflow, this interrelation is distributed over several steps, each targeting a single anatomical structure. The final result is then overlaid on the input image to provide visual guidance for the user. After close inspection and visual alignment of the planning to the underlying anatomy, the user might want to modify some parts of the planning geometry retrospectively. The design choice of our multi-stage approach naturally supports this use case. After manually adjusting any of the contained geometric objects, the geometric calculations in Stage 3) can be repeated at little computational cost, enabling real-time visualization of the planning on live images.

3.2. Multi-Task Network Architecture

The deployed multi-task network architecture (Figures 3 and A3, Table A1) is a Fully Convolutional Network and builds on the highly symmetrical Hourglass architecture [54,76]. An Hourglass first captures and encodes the input morphology at various scales and abstraction levels in a contracting path. The encoded features are then processed in an expanding path and recombined with residual information from the encoder (via parameterized skip connections) to reconstruct a multi-channel representation of the input. This representation is the input to one or more prediction heads, each responsible for a parameterized mapping to the desired output for one of the tasks.

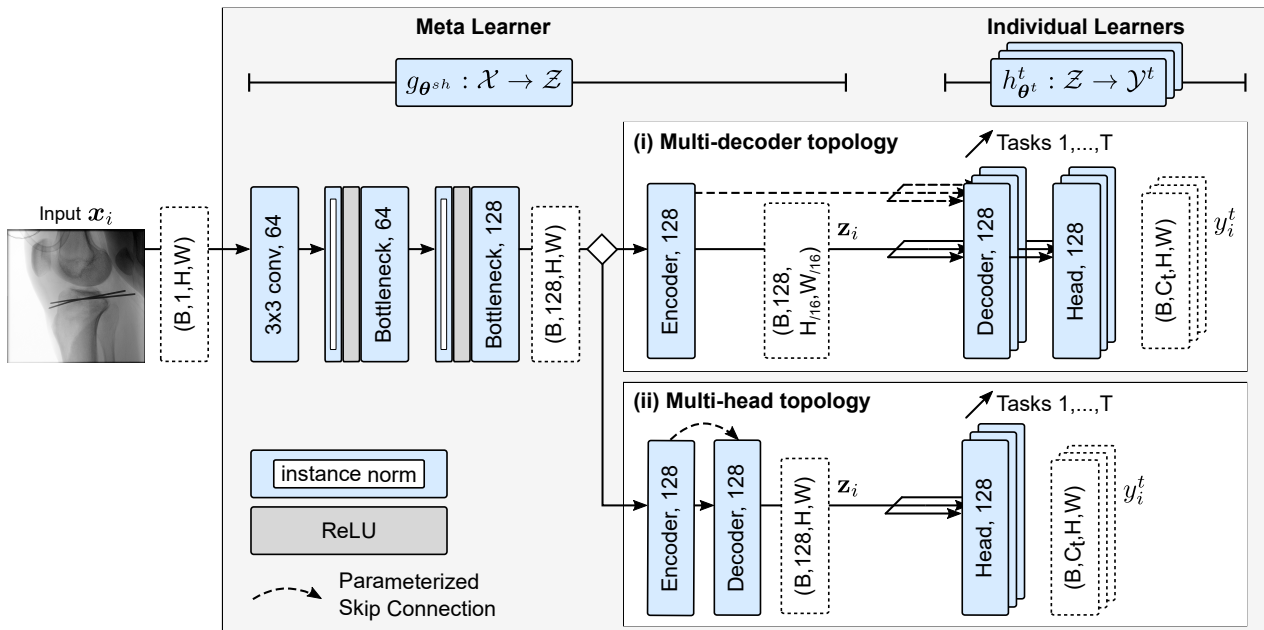


Figure 3. Architecture of the proposed multi-task network. After calculating a shared representation of the input image x_i with the meta learner, an individual learner for each task computes a corresponding prediction y_i^T in the task solution space \mathcal{Y}^t . Two strategies for the individual learners with different levels of parameter sharing are displayed: (i) Multi-decoder topology with independent decoders and prediction heads, (ii) Multi-head topology with shared decoder but independent prediction heads. Details of the architecture are provided in Figure A3, Table A1.

Every convolution block in the Hourglass, both in the contracting path, the expanding path(s), as well as the task-specific mappings, is constructed as a stack of residual pre-activation bottleneck units [76]. These residual units introduce an additional identity function to provide a direct and unchanged propagation path for the features and allow the preservation of the feature identity across multiple processing steps [77–79]. Based on a joint representation of feature information at every stage in the computation graph, the network can learn an optimal iterative refinement strategy and globally retain important information without sacrificing information gain [79]. We argue that this design choice is particularly advantageous in MTL scenarios, where some tasks may put more emphasis on global semantics provided by highly aggregated abstract features (e.g., semantic segmentation). In contrast, others may depend more on low-level feature representations (e.g., precise anatomical keypoint localization).

In comparison to the U-Net topology, where the number of convolutional kernels typically increases with decreasing feature size [80], the Hourglass assigns equal importance to all abstraction levels by using a constant number of kernels throughout the complete encoder–decoder architecture.

3.3. Dataset, Ground Truth, and Augmentation Protocol

3.3.1. Cohort 1: Diagnostic X-ray Images

First, we use diagnostic radiographs to analyze the effect of the MTL network topology, different loss/task weightings, and the amount of training data. We perform the ablation study on 211 (strictly) lateral diagnostic X-ray images of the knee joint acquired prior to reconstruction surgery. The images are of various spatial resolutions and occasionally show overlapping surgical implants. For each image, the study's first author annotated the outline polygon for the femur (S1) and tibia (S2) bone using the labelme annotation tool [81]. The anatomical keypoints (K1–K7) and lines (L1–L3) that are part of the planning geometries for MPFL, PCL, and ACL reconstruction surgery were annotated by the study's first author (training) an orthopedic surgeon (test) with a proprietary interactive tool by Siemens Healthcare GmbH. The data were separated into three subsets for training (167), validation (16), and testing (38) using stratified sampling. For the test split, all data that show a fixed-sized steel sphere with a 30 mm diameter was selected, allowing conversion between pixel and mm spaces.

3.3.2. Cohort 2: Intra-Operative X-ray Images

We study the applicability to intra-operative surgical planning using 89 lateral intra-operative fluoroscopic X-ray images from 43 distinct patients. The retrospectively collected data were originally acquired during reconstructive trauma surgery with a mobile C-arm system with flat panel detector (CIOS Spin, Siemens Healthcare GmbH). Compared to diagnostic radiographs, this cohort poses unique challenges to the automatic image analysis system. The images show a considerable overlap of the anatomy with implants and surgical tools, variance in administered dose, and in some cases, bilateral exposure of the knee joints due to spatial limitations in patient positioning. Additionally, the field of view of mobile C-arms is smaller than that of conventional X-ray machines, which leads to a more pronounced truncation of the bone shafts. The cohort was split into training (60), validation (14), and test (15) with no patient overlap across subsets. Similar to the diagnostic data, the planning geometry for MPFL, PCL, and ACL reconstruction surgery was annotated by the first author using a custom extension of labelme.

3.3.3. Augmentation and Ground Truth

For the training subset we applied online-augmentation with light spatial shearing ($s \in [-2^\circ, 2^\circ]$, $p = 0.5$), horizontal flipping ($p = 0.5$), in-plane rotation ($\alpha \in [-45^\circ, 45^\circ]$, $p = 1$), and scaling ($s \in [0.8, 1.2]$, $p = 1$) [82]. Any spatial transformation of the ground truth data were performed using projections of the numerical coordinate representations to prevent information loss by discrete interpolation. Before network propagation, the image intensities were normalized to the value range of $[0, 1]$ using min–max normalization given by $\tilde{x} = (x - \min(x)) / (\max(x) - \min(x))$. Since the original images are of different spatial resolutions and aspect ratios, they were brought to common dimensions of $[H:256 \times W:256]$ px by bi-cubic down-sampling. Thereby, the original aspect ratio was maintained, and the images were zero-padded to the target resolution.

After augmentation, the numerical ground truth was converted to its respective spatial representation. For that purpose, the bone outline polygons were converted to binary segmentation masks with the original image resolution. The numerical keypoint coordinates were translated to spatial activation maps/heatmaps by sampling a bivariate Gaussian function without normalization factor (i.e., $(x_0, y_0) = 1$ at keypoint coordinate) and a standard deviation of $\sigma = 6$ px. Similarly, the line heatmaps were generated by evaluating the point-to-line distance of the line-surrounding points in a locally-bounded hull. The line-orthogonal width of this hull was defined as $w = 6\sigma$. The point and line heatmaps were truncated at $\pm 3\sigma$, the remaining values were set to zero.

3.4. Training Policy and Implementation Details

The network parameters of the pre-operative models were optimized for 400 epochs using mini-batch gradient descent with a batch size of 2 and Nesterov momentum [83]. The learning rate and momentum were adjusted according to the cyclical learning rate policy (CLR) [84], annealing both values linearly between a lower and upper bound ($lr_{\min} = 0.001$, $lr_{\max} = 0.1$, $momentum_{\min} = 0.8$, $momentum_{\max} = 0.9$) within a full cycle length of 360 update iterations. For the intra-operative data, the training time was increased to 500. To overcome initially observed optimization instability, the cyclic gradient descent policy was replaced with RMSProp and a single learning rate decay with factor 0.1 at 350 epochs. A linear combination of Soft-Dice and Binary Cross Entropy $L_{\text{seg}} := \text{sDice} + \text{BCE}$ was used as loss function for the segmentation task. For keypoint and line regression, mean squared error losses $L_{\{\text{kpt}, \text{line}\}} := \text{MSE}$ were used. In the case of dynamic updates of the task weights using Gradient Normalization, the task weights were optimized using the same update policy and the asymmetry hyper-parameter was set to $\alpha = 1$, which was found to be the optimal value in a previous study ([54]). The model selection was based on the minimum inference error across all tasks on the validation set, which we evaluated after each training epoch. The selection criterion was defined as the minimum unweighted composite loss, i.e., $\mathcal{L} = \sum_{t=1}^T \mathcal{L}^t$. The network parameters were penalized with an L2 weight regularization term of $d = 0.00005$ to mitigate overfitting effects on the training subset.

All models were implemented in PyTorch v1.8 (Python v3.8.12, CUDA v11.0, cuDNN v7.6.5). It was ensured that all pseudo-random processes were initialized using the same random seed.

3.5. Evaluation Protocol

The sensitivity and precision of the predicted segmentation masks was measured using the Dice coefficient across both spatial dimensions, averaged over foreground and background. To also account for contour fidelity, we calculated the Average Symmetric Surface Distance (ASSD) [85]. For two contours C and C' with points $x \in C$ and $x' \in C'$, the ASSD is computed by evaluating the distance function $d(x, C') = \min_{x' \in C'} \|x - x'\|_2$.

$$\text{ASSD} = \frac{1}{|C| + |C'|} \left(\sum_{x \in C} d(x, C') + \sum_{x' \in C'} d(x', C) \right)$$

The keypoint localization error is measured as the Euclidean Distance (ED) between the ground truth coordinates and network predictions. For the lines, we are interested in the quality of fit between centroids and the orientation angles. For that purpose, we first truncate the heatmaps at the likelihood of 2σ to get rid of outliers and then calculate their centroid $c_{x,y}$ and orientation angle γ via raw image moments M_{pq} [86].

$$c_{x,y} = (\bar{x}, \bar{y}) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right)$$

$$\gamma = \frac{1}{2} \arctan \left(\frac{2\mu'_{11}}{\mu'_{20} - \mu'_{02}} \right), \quad \text{with } \mu'_{20} = \frac{M_{20}}{M_{00}} - \bar{x}^2$$

$$\mu'_{02} = \frac{M_{02}}{M_{00}} - \bar{y}^2$$

$$\mu'_{11} = \frac{M_{11}}{M_{00}} - \bar{x}\bar{y}$$

4. Results

4.1. Research Questions

As motivated in Section 3.1, we analyze the proposed MTL algorithm for optimization of several anatomical detection tasks for 2D surgical planning. Specifically, we answer the following research questions:

- Rq (1). How does the MTL network topology and task weighting strategy affect anatomical feature extraction?
- Rq (2). Does sharing tasks across anatomically related applications improve the feature extraction and target positioning compared to the single-application variant?
- Rq (3). How does the number of training data affect the planning accuracy?
- Rq (4). Can the performance on highly-standardized diagnostics images be applied to more complex imaging data in the intra-operative environment?

4.2. Rq (1) Network Topology and Task Weighting

We study the effect of a shared inductive bias via different levels of hard parameter-sharing on the optimization behavior and surgical planning quality. The two different MTL topologies, multi-head and multi-decoder, are optimized using either a uniform or dynamic GradNorm task weighting. The performances are compared to the respective STL variant.

The individual prediction errors per application and anatomical feature according to *Stage A* of the multi-stage algorithm are illustrated in Figure 4. For all three application types (Figure 4a,c), the segmentation quality does not benefit from a shared feature representation. The difference between the STL variant and the unweighted MTL variants is negligible but increases substantially for the dynamically weighted variants, which is reflected most notably in the average contour error. Interestingly, this trend is more pronounced for the applications with three tasks (MPFL, PCL) than it is for the ACL application with two tasks.

In contrast, the median precision of the predicted keypoints is significantly higher for the task-weighted MTL variants and increases with the number of landmarks. Starting with the PCL application ($n_{\text{kpt}} = 1$) where no substantial difference in the localization error between the variants can be observed, the weighted multi-head and multi-decoder models lead to a reduction of the position error by up to 50% for the ACL application ($n_{\text{kpt}} = 5$). In the latter ACL case, this particularly shows for the tibial keypoints K4–K6 that are anatomically more complex and have a larger positional variance. There, the STL variant and unweighted multi-decoder variant suffer significantly from a comparatively large task-specific inductive bias. The reduced keypoint precision for these variants is accompanied by a drastically reduced maximum intensity of the predicted heatmaps, indicating low intra-model confidence and leading to error-prone intensity-peak extraction (Table 2, Figure 5). For the estimation of the lines, the results do not show a clear trend, but the weighted multi-head variant generally achieves the most accurate and consistent results. The results further indicate that in most cases, separating the reconstruction with task-specific decoders leads to slight improvements of the segmentation, but in all other cases is comparable to or even performs worse than the significantly more sparse multi-head variant. This observation suggests that not only a shared encoding but also a shared decoding of aggregated features with only a small trailing task-specific part is preferable.

These insights are in large part supported by the applications' planning metrics calculated at *Stage C* of the multi-stage algorithm (Figure 6). Looking at the tibial aspects of the ACL planning, we see highly significant improvements of the approximated bundle attachments using a weighted MTL scheme compared to STL and unweighted MTL. However, the difference between the two levels of parameter-sharing is not significant, which means that a shared decoder neither benefits nor worsens the discriminative power. Interestingly, the planning metrics corresponding to the position of the femoral ligament attachment bundles are not positively affected by any MTL setting, with the single-task performance being among the best variants. This relation supports the previous findings (Figure 4) where anatomically intricate and ambiguous landmarks significantly benefit from more constrained and generalized features, and clearly visible ones, such as the Blumensaat keypoints K1 and K2, profit from a higher number of task-specific features, allowing to align much stronger to the most frequent anatomical pose.

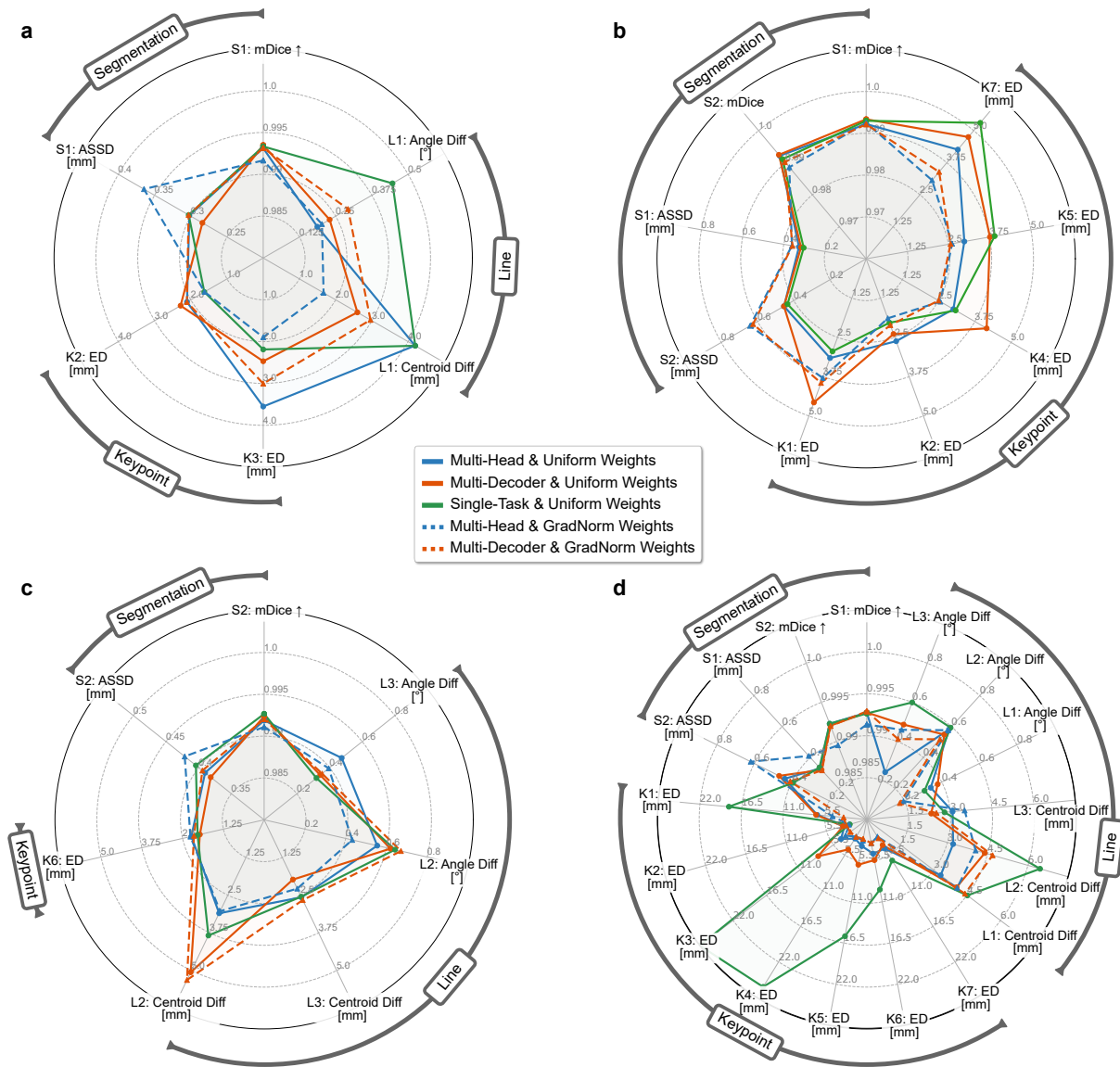


Figure 4. Radar charts of prediction errors per application and anatomical feature according to Stage A) (Section 3.1.2) and different model topologies and optimization policies. For all metrics but the mean Dice coefficient (mDice ↑) for segmentation tasks, lower is better. (a), MPFL. (b), ACL. (c), PCL. (d), combined (MPFL, ACL, PCL).

Table 2. Analysis of the maximum intensity of keypoint heatmaps for different model topologies and task weighting strategies. Each variant’s values for mean and standard deviation (mean ± std) are calculated across all application-specific keypoints on every test set image. The results of the best network variant for each planning methodology is marked in bold.

Planning	Multi-Head		Multi-Decoder		Single-Task
	Uniform	GradNorm	Uniform	GradNorm	Uniform
MPFL ($n = 2$)	0.64 ± 0.11	0.83 ± 0.09	0.76 ± 0.09	0.76 ± 0.11	0.80 ± 0.12
ACL ($n = 5$)	0.42 ± 0.08	0.71 ± 0.12	0.43 ± 0.07	0.73 ± 0.12	0.57 ± 0.16
PCL ($n = 1$)	0.90 ± 0.11	0.84 ± 0.11	0.83 ± 0.09	0.79 ± 0.10	0.84 ± 0.08
Comb. ($n = 7$)	0.41 ± 0.12	0.66 ± 0.12	0.38 ± 0.10	0.71 ± 0.12	0.23 ± 0.11

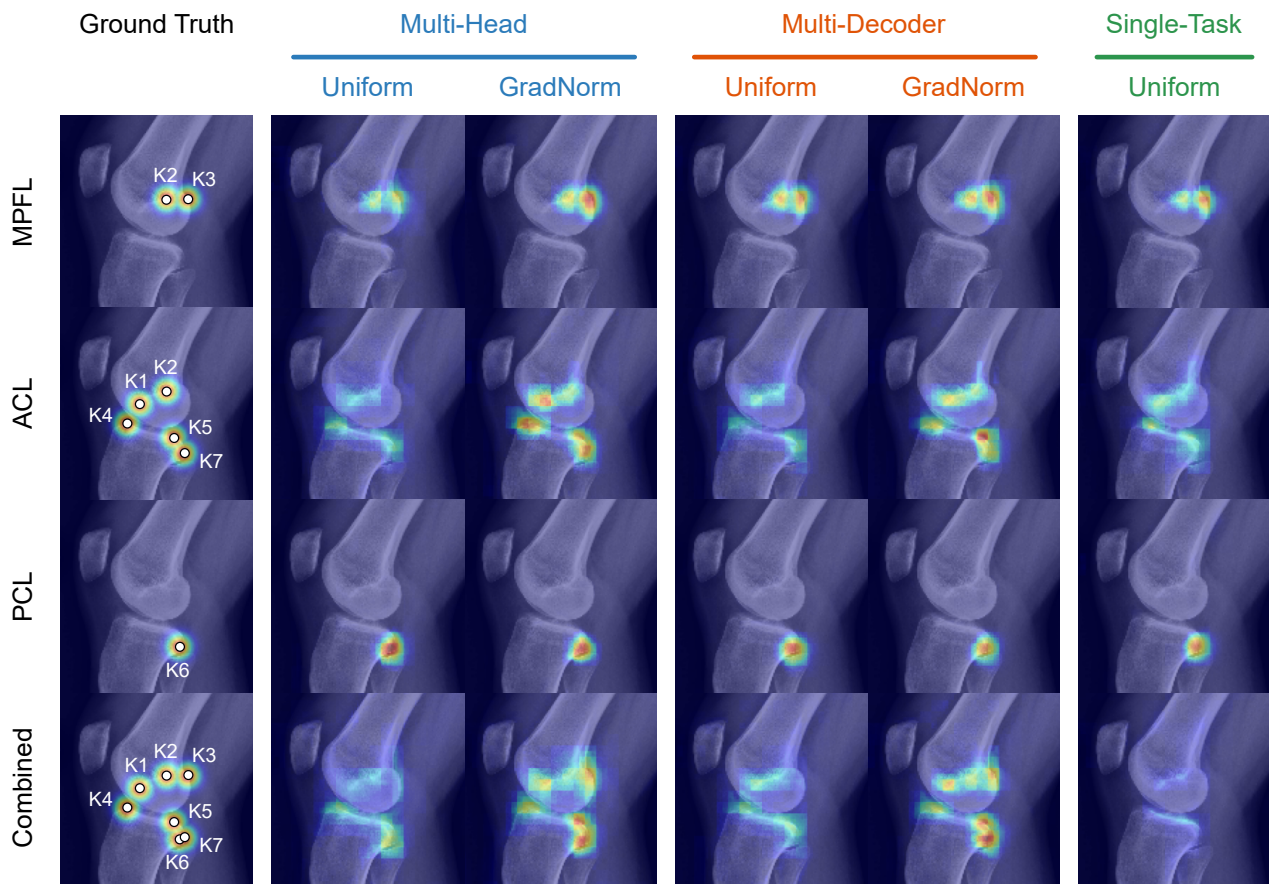


Figure 5. Visualization of heatmap prediction quality for different model topologies and task weighting strategies for a random sample from the test set.

In this context, it should be noted that the visible spatial layout of the proximal tibia head, where most considered landmarks are located, strongly varies with slight changes in the viewing direction (Figure A1). Since the complex 3D structure of the tibial articular surface cannot be fully assessed by a single projection to 2D, different viewing directions result in a greater positional variance of the landmarks. This variance leads to ambiguity in ground truth annotation and consequently limits the models’ ability to generalize well. The ambiguity is further increased by difficulties in interpreting depth information, caused by the general thickness of the tibial plateau, multiple intra-bone overlaps, and the resulting fuzzy contours. However, these uncertainties in landmark detection are less pronounced for the distal femur. There, the anatomy is less complex, and the depth in 2D projections is easier to assess. Moreover, the position of the femoral keypoints K1 and K2 located in the intercondylar notch is not affected by some rotation of the viewing direction around the Blumensaat line.

For the MPFL and PCL planning applications, the different network topologies and weighting variants do not significantly differ from each other. For MPFL planning in particular, all model variants meet the standard clinical requirements of (<2.5 mm) precision [21] and previously published inter-rater performance [54].

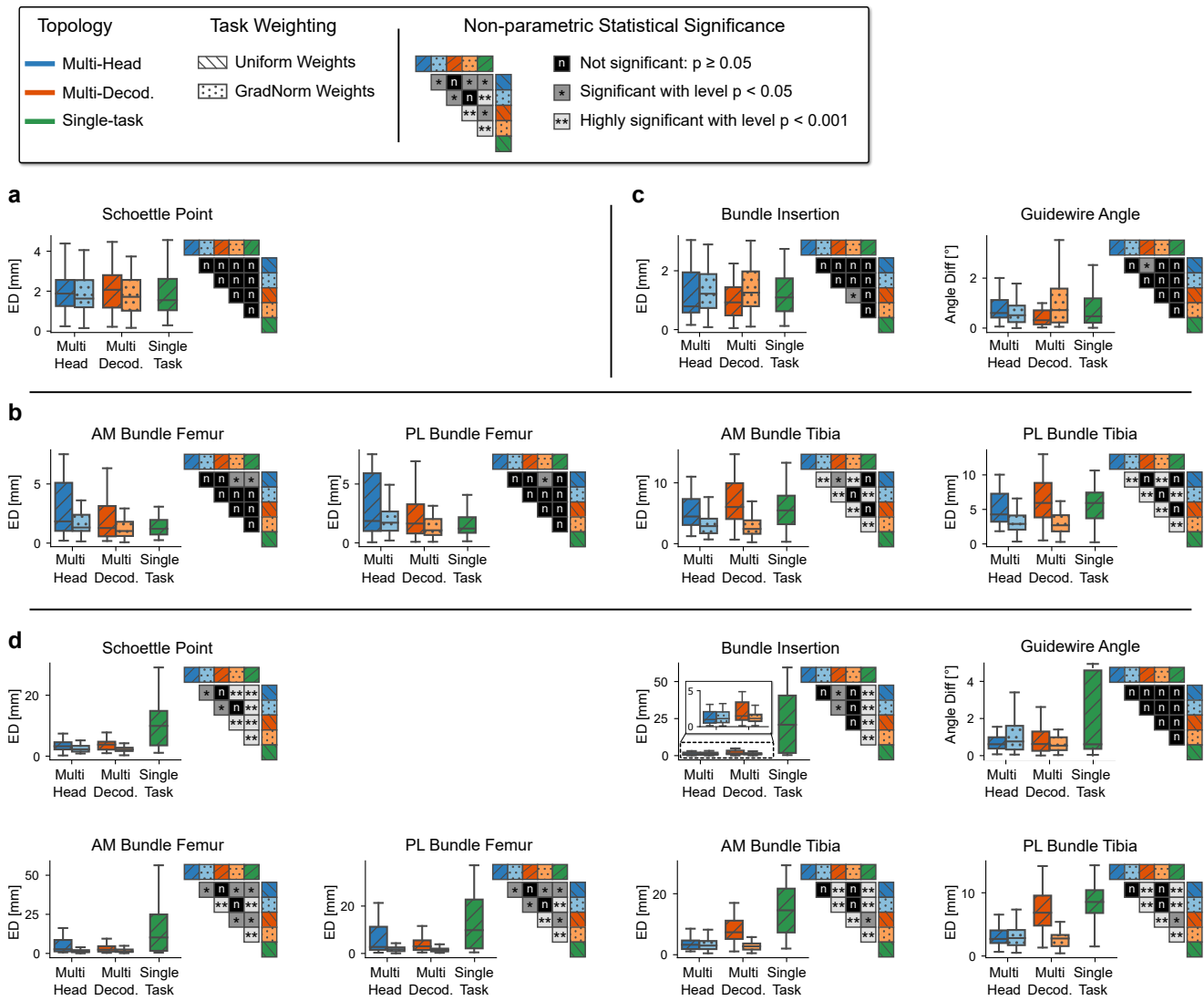


Figure 6. Boxplots of the planning metrics for each surgical application and different model topologies and optimization policies (without fliers/outliers). (a), MPFL. (b), ACL. (c), PCL. (d), combined (MPFL, ACL, and PCL). For each boxplot and topology pairing, we tested statistical significance with a two-sided Mann–Whitney U rank test ((n) not significant: $p \geq 0.05$, (*) significant: $p < 0.05$, (**) highly significant: $p < 0.001$).

4.3. Rq (2) Combining Tasks Across Multiple Surgical Applications

We analyze whether task sharing across different surgical planning types improves over the single-application case. For that purpose, we aggregate the anatomical features of a task for all considered planning applications and define a combined model that we analyze the same way as in Experiment (1).

As can be seen in Figures 4 and 5, the precision of the keypoints estimated with the combined model benefits from a generalized feature set. Especially for the keypoints K3–K7 that suffer from positional uncertainty in the 2D projections, the combined model suffers from extensive spatial smearing at the heatmap peak and severely reduced likelihood (Table 2). This imprecision of the intensity’s spatial peak negatively impacts the argmax-based coordinate extraction, resulting in large deviations from the ground truth. Furthermore, the combined model’s segmentation quality slightly improves when using task-specific decoders, matching the models’ behavior in the single-application case. However, this performance gain does not translate into better line and keypoint prediction accuracy, which ultimately does not warrant the additional computational cost of mul-

multiple decoders. The combined model’s planning metrics show no significant change in performance compared to their single-application counter-parts in ACL and PCL planning (Figures 6 and 7). However, the localization of the MPFL Schoettle Point suffers from a combined task definition which leads to larger angulation and general positioning errors of the approximated femoral shaft line (Figure 4).

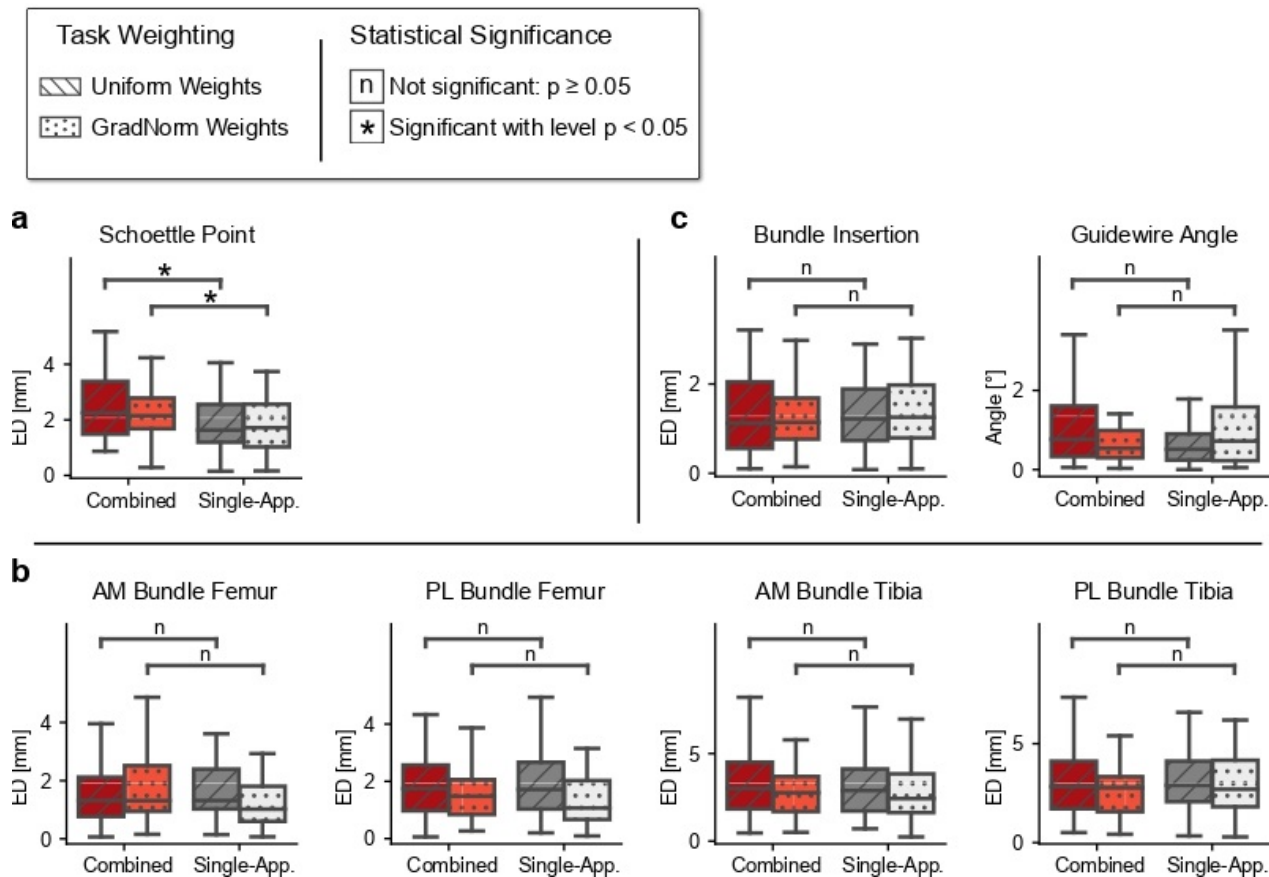


Figure 7. Comparison between single- and multi-application model performance on the test set. (a), MPFL. (b), ACL. (c), PCL. For each boxplot and pairing, we tested statistical significance with a two-sided Mann–Whitney U rank test ((n) not significant: $p \geq 0.05$, (*) significant: $p < 0.05$).

4.4. Rq (3) Effect of the Number of Training Data

We analyze how the amount of training data influences the choice of MTL topology design and loss weighting. For that purpose, the size of the training cohort is subsequently reduced by taking only a fraction of data $f \in \{1.0, 0.75, 0.5, 0.25\}$ (random selection, fixed between experiments).

Almost all single-application variants (Figure 8a,c) evidence a low negative correlation of the planning errors with increasing data count (strictly negative correlation for all weighted MTL models). Most noticeable is the significant positive impact (correlation probability $p < 0.05$) of the amount of data when applying the STL models to the ACL application (Figure 8b). However, the performance peaks are not always reached using the maximum amount of data. Although the femoral planning metrics, for example, benefit from a broader selection of data samples, the error rate is lowest at a reduced data factor of $f = 0.75$.

No clear trend of the effect of differing amounts of training data can be extrapolated for combined task representations of multiple planning applications (Figure 8). For the STL models, the performance drops significantly with increasing data points, particularly for the femoral planning metrics. Interestingly, the performance of the MTL models remains almost unchanged with increasing factor f , indicating that a higher variance in the training

distribution can not be exploited for learning a more generalized feature set. However, the direct comparison between STL and MTL in the multi-application scenario attests to the MTL variant a more effective extraction of the relevant features and thus a more stable learning process and better inference performance.

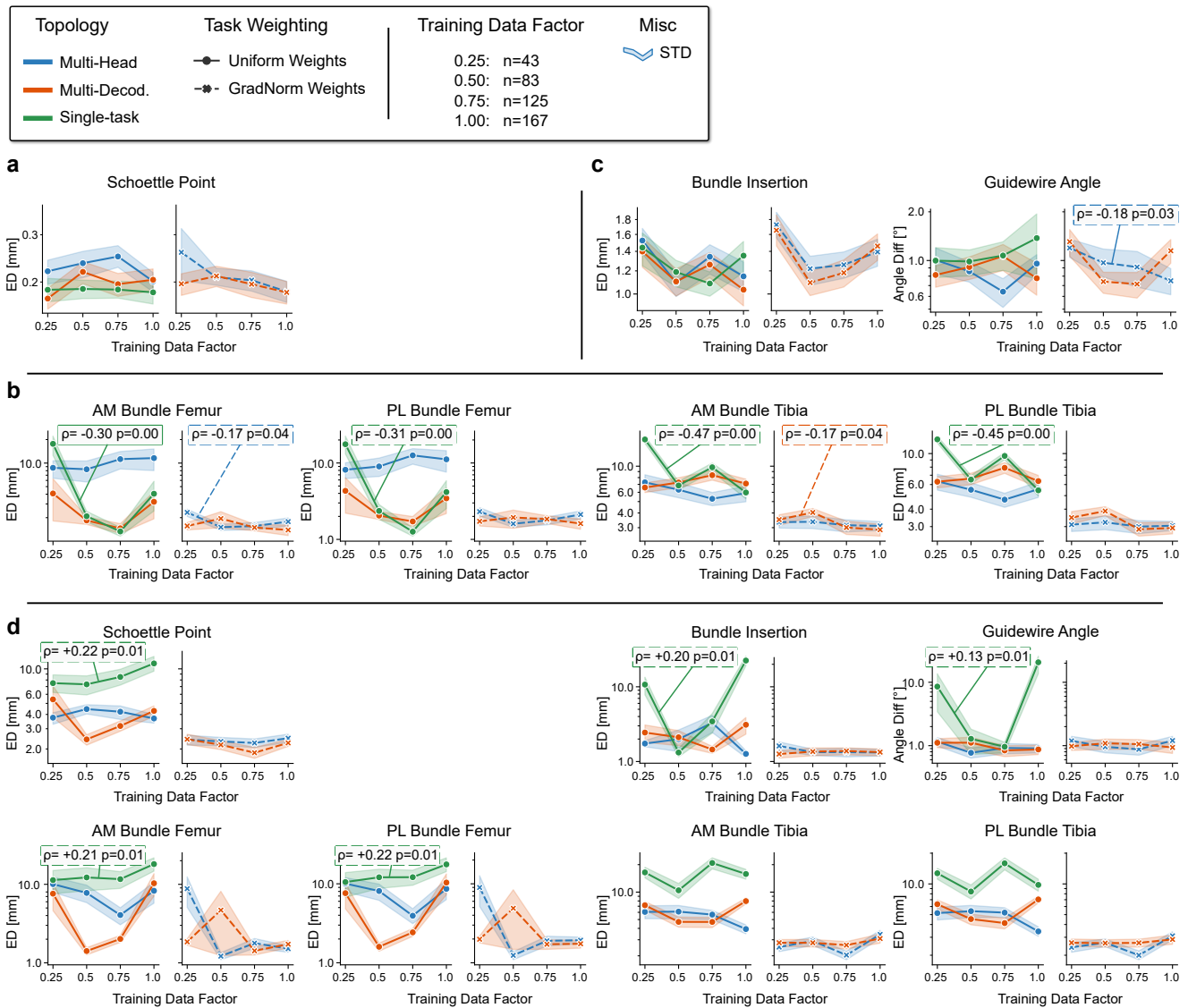


Figure 8. Effect of different amounts of training data on the planning metrics (logarithmic scale for y-axis for better visibility). (a), MPFL. (b), ACL. (c), PCL. (d), combined (MPFL, ACL, PCL). We report correlation coefficients ρ with rejection probability $p < 0.05$ based on two-sided Spearman’s rank correlation measure.

4.5. Rq (4) Application to Intra-Operative Data

We apply the weighted MTL algorithms to a complex intra-operative setting. The X-ray images show difficult trauma cases with fractured anatomy, strong instrumentation overlap, weak contrast, and bilateral exposure of the knee joints. We skip the evaluation of the PCL planning since it requires explicit knowledge of the pixel-to-mm relation.

For MPFL planning, the automatic planning achieves high accuracy for cases with clearly visible bone shaft (Table 3, Figure 9). In the case of strong shaft truncation, the estimated femoral cortex line suffers from dorsal shifting with increased deviation from the anatomical bone axis. The translation leads to malpositioning of the Schoettle Point. For extreme cases, the cortex line cannot be recovered. This dependency confirms previ-

ous findings highlighting the importance of an application-specific acquisition protocol, especially for the limited field of view in mobile C-arm imaging [50]. Although femoral ACL planning is highly accurate, tibial ACL planning suffers from ventral migration of the tibial proximal epiphysis corner towards the champagne drop-off point. This area is often superimposed by metallic plates, obscuring salient visual cues needed for precise localization. In 3 out of 15 test cases, the planning was cancelled due to fragmentation of the tibia segmentation mask and subsequent errors in contour extraction. These error modes can be reduced for three out of four bundle insertion points in ACL planning when using a multi-decoder model.

The increased complexity in the intra-operative setting could to be compensated by additional training data. In particular, more variation in instrument type and position must be covered to address the ambiguity in bone segmentation and the resulting high contour errors.

Table 3. Numerical planning results on intra-operative test data. Each metric value is reported w.r.t. the original resolution of [H:976 × W:976] px. The results of the best network variant for each planning methodology is marked in bold. In three ACL cases, planning was not possible due to poor segmentation quality.

Planning	Metric [px]	Multi-Head		Multi-Decoder	
		Median, CI _{95%}	Cnt.	Median, CI _{95%}	Cnt.
MPFL	Schoettle Pt.	9.08 , [6.35, 128.21]	15	14.02, [9.72, 22.46]	15
ACL	AM Femur	8.17, [5.34, 15.65]	12	4.80 , [3.79, 14.04]	12
	PL Femur	5.66, [4.17, 12.84]	12	5.04 , [2.85, 12.48]	12
	AM Tibia	22.13, [9.52, 82.04]	12	14.18 , [10.27, 42.87]	12
	PL Tibia	13.17 , [10.09, 38.74]	12	16.55, [12.85, 41.71]	12

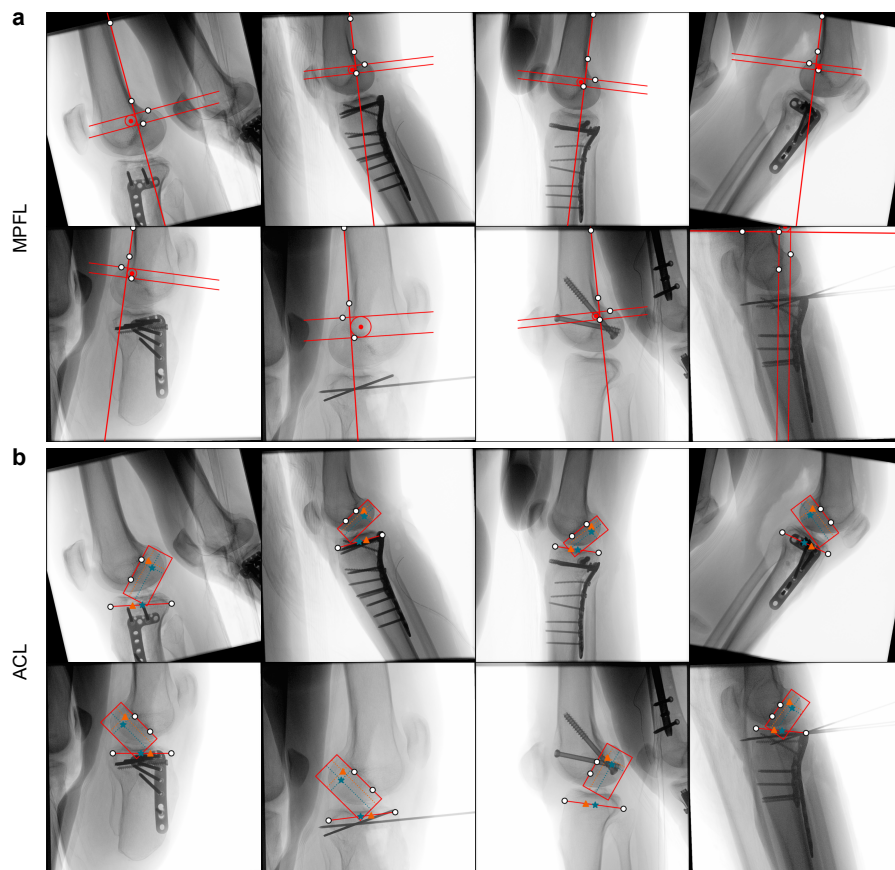


Figure 9. Planning predictions on intra-operative test data with multi-head topology and GradNorm task weighting. (a), MPFL. (b), ACL.

5. Discussion

Careful preparation of the surgical plan and its integration into the intra-operative workflow are key drivers for safety and low revision rates in reconstructive orthopedic surgery. Currently, available software mainly supports pre-operative planning on diagnostic 2D and 3D data and the following registration on intra-operative images. However, for some surgical interventions, such as ligament reconstruction and intramedullary nailing, pre-operative 3D scans are typically not acquired. Moreover, the usefulness of planning on static 2D radiographs remains questionable due to an unpredictable patient and device positioning during intra-operative management. A planning platform operating on intra-operative data would elevate surgical safety, increase the precision in critical positioning steps, and enable live outcome assessment in the operating room (OR).

In this study, we developed and analyzed a multi-stage algorithm that combines deep learning-based anatomical structure localization on 2D X-ray images and geometric post-processing to enable pre- and intra-operative surgery planning. The algorithm mimics the clinically established, manual planning rationale by separating feature detection and their geometric interrelation in a surgical plan. In the method evaluation on three trauma-surgical applications, we observed a high spatial precision in localizing optimal drill insertion points and low angulation errors for k-wire instrumentation and anatomical axis detection.

A fundamental characteristic of the multi-stage design is that the extrinsic positioning errors depend solely on the performance of automatic feature detection. While we cannot reduce the intrinsic error between the true anatomical position and approximation of the target by the planning blueprint, we approach extrinsic error reduction in two ways. First, users can manually adjust the position of the detected anatomical features, granting granular real-time control over the plan. Second, the feature extractor's learning policy can be improved using an MTL scheme. MTL's main rationale is to exploit similarities in data and underlying data formation by using a meta-learner that optimizes a common feature set. Our results substantiate the benefit of these generalized features for detecting anatomical structures on orthopedic image data. In particular, the precision of anatomically complex keypoints benefits from MTL compared to STL. We conclude that additional supervision of a keypoint's context area reduces positional uncertainty by exploiting strong spatial correlations between the keypoint and the bone outline on a 2D projection. This outline can be reliably detected and exhibits low variation between patients. Since most radiographic keypoints lie on this outline, boundary-aware tasks, such as bone segmentation, provide strong auxiliary cues during optimization.

Consistent with recent MTL literature, task balancing effectively prevents overfitting to the segmentation task, which is spatially more dominant than the keypoint heatmaps and contributes high-magnitude gradient signals to the composite loss function. Although balancing with GradNorm shows superior performance over unweighted optimization, it solely targets homogenization of the tasks' relative learning rates, not taking into account complex and potentially conflicting task interaction. We believe that it could be advantageous to combine this balancing with an active alignment of gradient direction and magnitude, e.g., by using Gradient Sign Dropout [74] or Gradient Vaccine [73]. Although such iterative improvements are attractive, their actual expediency is challenged by an upper performance limit imposed by the number of tasks, their concrete characteristics and relatedness, and the model complexity of the meta-learner.

This limitation is particularly evident when combining tasks across similar surgical applications. Faced with increased task complexity, the combined model is unable to find a common representation that improves over the single-application case. Without increasing model capacity—which comes at the expense of increased execution time and computational footprint—we advocate an MTL scheme on the application level rather than grouping multiple applications per anatomical region.

No clear conclusion can be made for the necessary amount of training data. Although the model performance increases slightly with more data samples, the magnitude of improvement neither warrants the additional labeling effort nor lets us define lower and

upper bounds for the cohort size. In addition, the ablation study was performed on pre-operative diagnostic X-ray images with a highly standardized protocol for acquisition and patient positioning. This little variation compares poorly to the types of variance that are observed in the inter-operative environment. There, intricate fracture patterns, a broad array of surgical tools and implants, and greater out-of-plane rotation angles caused by spatially-constrained patient positioning introduce novel types of image characteristics. In this scenario, additional data may help to reduce model bias.

We see several limitations of our study. At the present stage, our method does not consider the registration of an initial surgical plan on subsequent intra-operative live data. Although the lack of registration does not immediately impact the planning process, it hampers outcome evaluation and more advanced real-time guidance using fluoroscopy or periodical positioning checks with single shots. Furthermore, the method was only evaluated for surgical applications on the knee joint. Compared to the large femur and tibia, the bones of the wrist, elbow or foot are much smaller and have less pronounced contours in the 2D projection due to strong inter-bone overlap. A further restriction of our method is the disconnect between the optimization criterion and the actual planning target. The algorithm's first stage solves the anatomical feature detection problem independently from the planning geometry, which can be seen as optimizing a proxy criterion with less anatomic and geometric constraints. However, such constraints could help rule out implausible configurations and likely push the model to learn an abstract semantic representation of the actual planning target.

To address these shortcomings, future studies could investigate a combination of anatomical feature extraction and subsequent plan generation with direct regression of the planning target. The planning geometry can be implemented with differentiable functions, allowing gradient propagation from the actual target to the anatomical feature extractor. In this scenario, a consistency term between the geometric and direct target estimates can help align both estimates during optimization. Furthermore, previous studies have shown that planning methods are highly susceptible to rotational deviations between the underlying anatomy and the standard projection [87,88]. We believe that extending our method with an automatic assessment [89,90] of this mismatch could assist in preventing target malpositioning and improve the planning's reliability and clinical acceptance. Another desirable extension is to combine our method with surgical tool recognition. In particular, knowledge of the k-wire direction and tip position would allow software-side security controls and advanced guidance in the intra-operative execution of the plan [56].

However, already in its current stage, the presented approach offers a precise automatic tool to support surgical intervention planning directly on intra-operative 2D data. Such intra-operative plan generation is particularly relevant for trauma surgery, where pre-operative 3D images are often unavailable, limiting the application of many analysis and navigation solutions operating on MRI [91–94] and CT [95–97] data. The proposed multi-stage planning algorithm was designed to separate automatic image analysis and geometric plan generation, allowing the user fine-grained control over each step of the proposed plan. This design decision ultimately improves user acceptance of the planning tool by allowing inaccuracies in the automatically proposed plan to be corrected for complex and rare injury patterns. Furthermore, we were able to show empirically that parallel detection of all anatomical features relevant to the plan using a weighted MTL policy benefits positioning accuracy, confirming the long-standing consensus in the MTL literature.

6. Conclusions

In this work, we specifically address the need for accurate X-ray guided 2D planning in the intra-operative environment for common injury patterns in reconstructive orthopedic surgery. The proposed multi-stage algorithm separates automatic feature extraction and the actual planning step, which we consider a pivotal design choice that allows for real-time control of individual planning components by the user. The detection step can be efficiently solved with a lightweight MTL scheme, allowing short execution times and improving

spatial precision over the STL case. We believe that the developed platform will help overcome the limitations of current clinical practice, where planning is done manually on diagnostic 2D data with limited intra-operative value. Generating the surgical plan directly in the OR will elevate surgical accuracy and ultimately motivate the development of novel 2D planning guidelines. Future work should explore the possibilities of end-to-end training of the feature extractor to provide gradient updates in each step of the planning geometry. Additionally, integrating the method with automatic mechanisms for image alignment and surgical tool localization would assist the surgeon not only in planning, but also in image data acquisition and directly in surgical execution.

Author Contributions: Conceptualization, F.K., B.S. and H.K.; methodology, F.K.; software, F.K.; validation, F.K.; formal analysis, F.K.; investigation, F.K.; resources, A.M., B.S., M.P., J.S.E.B. and H.K.; data curation, F.K., B.S. and M.P.; writing—original draft preparation, F.K.; writing—review and editing, A.M., B.S., M.P., J.S.E.B. and H.K.; visualization, F.K.; supervision, A.M. and H.K.; project administration, H.K.; funding acquisition, A.M. and H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to the clinical data being obtained retrospectively from anonymized databases and not generated intentionally for the study.

Informed Consent Statement: The acquisition of data from living patients had a medical indication and informed consent was not required.

Data Availability Statement: The image data used for the evaluation of the software cannot be made available. The methods and information presented here are based on research and are not commercially available.

Acknowledgments: The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the Bavarian State Ministry for Science and Art. The authors acknowledge the support of Siemens Healthcare GmbH, Forchheim, Germany. The authors acknowledge financial support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg within the funding programme “Open Access Publication Funding”.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MPFL	Medial Patellofemoral Ligament
ACL	Anterior Cruciate Ligament
PCL	Posterior Cruciate Ligament
AM	Anteromedial
PL	Posterolateral
OR	Operating room
MTL	Multi-task learning
STL	Single-task learning
GradNorm	Gradient normalization
ASSD	Average symmetric surface distance
ED	Euclidean distance

Appendix A. Pre-Operative Planning Examples

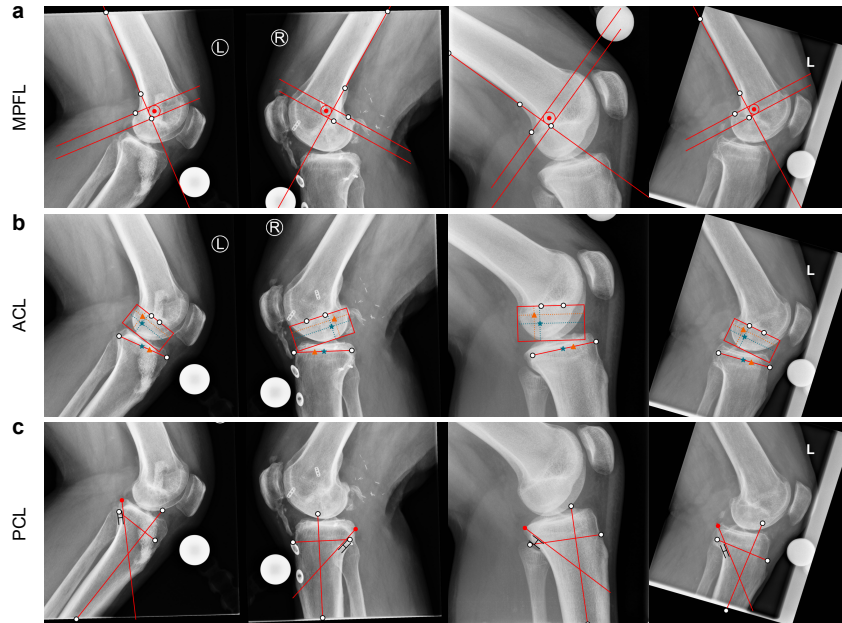


Figure A1. Planning predictions on pre-operative test data with multi-head topology and GradNorm task weighting. (a), MPFL. (b), ACL. (c), PCL.

Appendix B. Planning Geometry

Appendix B.1. MPFL Reconstruction

In the following, we outline the geometric calculation of the Schoettle Point p_{sp} for MPFL reconstruction planning (Figure A2a) [21]. First, the heatmap is thresholded at 1σ likelihood, and the bone contour C is calculated using morphological operations on the binary segmentation $S1$ [50]. The thresholded heatmap is used to select the bone shaft's relevant contour points. A line is fitted to this 2D point cloud using major axis regression, and the cloud's two extreme points are used as start and end points l_1 and l_2 of the cortex vector \vec{v} .

Next, the detected keypoints k_1 and k_2 are projected onto the unit vector \hat{v} .

$$\begin{aligned} k'_1 &= l_2 + ((k_1 - l_2) \cdot \hat{v})\hat{v} \\ k'_2 &= l_2 + ((k_2 - l_2) \cdot \hat{v})\hat{v} \end{aligned} \quad (A1)$$

This allows to calculate the radius r of the Schoettle confidence circle around the Schoettle Point as $r = \|(k'_1 - k'_2)/2\|_2$. To ensure that p_{sp} is always positioned anterior to \vec{v} , k_2 is assumed to always be posterior to \hat{v} . This assumption lets us calculate the knee configuration τ using the z-component magnitude of the 3D cross-product's between \hat{v} and $(k_2 - l_2)$.

$$\tau = \text{sgn}\left(\det\begin{bmatrix} \hat{v}_x & (k_2 - l_2)_x \\ \hat{v}_y & (k_2 - l_2)_y \end{bmatrix}\right) \quad (A2)$$

$\tau > 0$ refers to a left knee, $\tau < 0$ denotes to a right knee, and $\tau = 0$ denotes placement of k_2 on the line through \hat{v} . Finally, the target Schoettle Point p_{sp} is computed as follows.

$$\begin{aligned} \hat{v}^\perp &= R(90^\circ) \hat{v} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \hat{v} \\ p_{sp} &= k'_2 + \frac{(k'_1 - k'_2)}{2} - (\tau r)\hat{v}^\perp \end{aligned} \quad (A3)$$

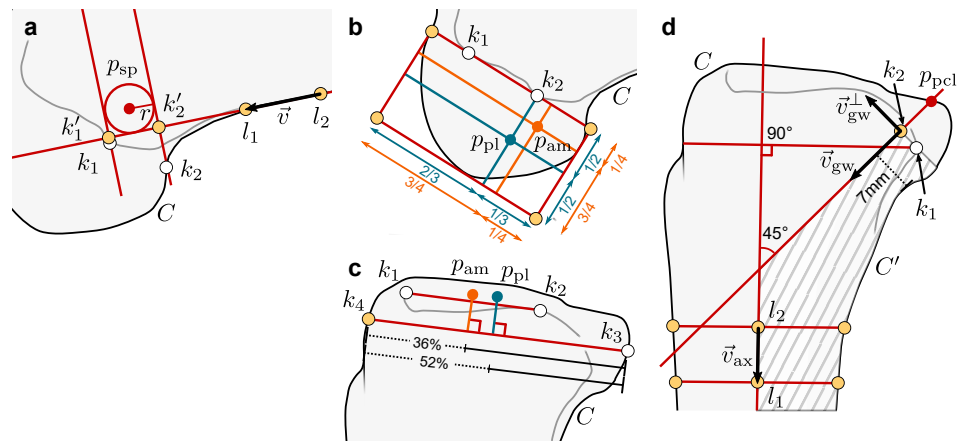


Figure A2. Geometric relations of the considered surgery planning types. (a), MPFL. (b,c), femoral and tibial ACL. (d), PCL.

Appendix B.2. ACL Reconstruction

Standard ACL reconstruction planning can be divided into femoral and tibial planning of the anteromedial (AM) and patellolateral (PL) bundles. Geometric planning on the femoral side follows the quadrant method by Bernard et al. [32]. Tibial attachment points are related to fractions on the Amis/Jakob line [33] originally described by Stäubli/Rauschnig [34].

Appendix B.2.1. Femoral Bundle Attachments

The basic geometric relations are visualized in Figure A2b. After extracting the femoral bone contour C , a line is defined that passes through the anterior and posterior Blumensaat points k_1, k_2 . The bounds of the rectangular grid are calculated as follows. First, the anterior and posterior intersection of the Blumensaat line with C are obtained by iterative search on small contour sections with Bézier clipping. Overlapping groups of four contour points are interpolated with a Bézier spline, generating a set of candidate intersections. The actual intersection points are selected by evaluating the minimum orthogonal distance to the Blumensaat line. This process is repeated for multiple distal shifts of the detected Blumensaat line to extract to maximum head width. The grid's distal bound are similarly identified by selecting the largest possible parallel shift of the Blumensaat line that still intersects with C .

The projection of the AM bundle center p_{am} onto the Blumensaat line is at $\sim 25\%$ from the most anterior aspect, and in orthogonal projection $\sim 25\%$ from the most distal aspect of the femoral condyle. The center of the projected PL bundle p_{pl} lays at $\sim 33\%$ from the most anterior aspect, whereas the orthogonal projection is located at $\sim 50\%$ from the most distal aspect of the femoral condyle [61].

Appendix B.2.2. Tibial Bundle Attachments

The bundle positions on the tibial head are visualized in Figure A2c. After extracting the tibial bone contour C , a parallel line to the tibial plateau ($k_1 - k_2$) passing through k_3 is computed, i.e., the Amis/Jakob line. The intersection is again calculated using iterative search along C with Bézier clipping. Afterwards, the optimal positions of the AM and PL bundles along the Amis/Jakob line are at 36% and 52% from the posterior tibial epiphysis, respectively.

Appendix B.3. PCL Reconstruction

Planning of the transtibial drilling tunnel orientation and PCL insertion site follows the radiographic construction described by Johannsen et al. (Figure A2d) [23]. We refer the reader to [55] for a comprehensive explanation of this construction and [50] for details on the utilized anatomical axis detection.

In brief, after extracting the tibial contour C and estimating the anatomical axis \vec{v}_{ax} , the drilling path \vec{v}_{gw} is initialized at a 45° acute angle oriented towards the anterior aspect of the tibia. Similar to planning of the MPFL Schoettle Point, this step necessitates knowledge of the knee orientation. The orientation τ can be calculated using the vectors $\vec{v}_{ax} = (l_1 - l_2)$ and $\vec{v}_{cgd} = (k_1 - l_2)$:

$$\tau = \text{sgn} \left(\det \begin{bmatrix} (\vec{v}_{ax})_x & (\vec{v}_{cgd})_x \\ (\vec{v}_{ax})_y & (\vec{v}_{cgd})_y \end{bmatrix} \right). \tag{A4}$$

The initial position of the PCL insertion point p_{pcl} is constructed such that the drilling path \vec{v}_{gw} intersects C at the anatomic insertion site on a parallel line with distance 5.5 mm to the champagne glass drop-off line (line orthogonal to the bone axis passing through k_1). The search for smooth intersection p_{pcl} is done iteratively by Bézier spline fitting to overlapping groups of four contour points and subsequent Bezier clipping. This procedure reduces the effect of segmentation inaccuracy.

After initialization, a safety margin between the drilling path and the bone cortex is established by the smallest possible parallel shift of \vec{v}_{gw} such that $\|k_1 - k_2\|_2 \geq 7$ mm.

Appendix C. Neural Network Topology

Table A1. Detailed topology of the used multi-task network in the multi-head configuration. The architecture is adapted from the Stacked Hourglass Network by Newell et al. [76]. The bottleneck and head topology are illustrated in Figure A3. (s: stride, p: padding, \oplus : feature fusion by element-wise addition, NN: nearest-neighbor).

	Block	Input	Operation	C_{in}	C_{out}	Input Size	Output Size
Pre	P1	–	3x3 conv, p=1	1	64	256×256	256×256
	P2	P1	Bottleneck	64	128	256×256	256×256
	P3	P2	Bottleneck	128	128	256×256	256×256
Encoder	E1	P3	Bottleneck 2×2 max pool, s=2	128 128	128 128	256×256 256×256	256×256 128×128
	E2	E1	Bottleneck 2×2 max pool, s=2	128 128	128 128	128×128 128×128	128×128 64×64
	E3	E2	Bottleneck 2×2 max pool, s=2	128 128	128 128	64×64 64×64	64×64 32×32
	E4	E3	Bottleneck 2×2 max pool, s=2	128 128	128 128	32×32 32×32	32×32 16×16
	Skip Con.	S1	P3	Bottleneck	128	128	256×256
S2		E1	Bottleneck	128	128	128×128	128×128
S3		E2	Bottleneck	128	128	64×64	64×64
S4		E3	Bottleneck	128	128	32×32	32×32
Decoder	D1	E4	Bottleneck x2 NN up-sampling	128 128	128 128	16×16 16×16	16×16 32×32
	D2	$D1 \oplus S1$	Bottleneck x2 NN up-sampling	128 128	128 128	32×32 32×32	32×32 64×64
	D3	$D2 \oplus S2$	Bottleneck x2 NN up-sampling	128 128	128 128	64×64 64×64	64×64 128×128
	D4	$D3 \oplus S3$	Bottleneck x2 NN up-sampling	128 128	128 128	128×128 128×128	128×128 256×256
	Heads	H1 (t = 1)	$D4 \oplus S4$	Head	128	C_1	256×256
H2 (t = 2)		$D4 \oplus S4$	Head	128	C_2	256×256	256×256
H3 (t = 3)		$D4 \oplus S4$	Head	128	C_3	256×256	256×256

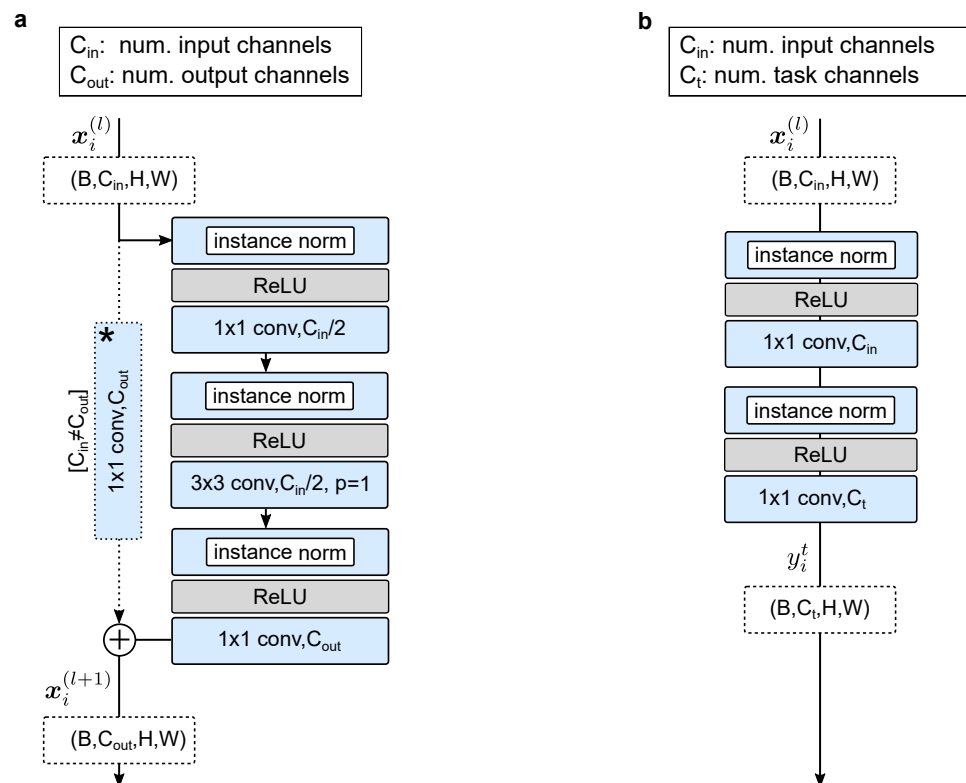


Figure A3. Topology of utilized neural network building blocks based on the work from Newell et al. [76] and He et al. [77]. (a), layout of 3-layer pre-activation bottleneck with identity mapping [77]. If the number of input and output channels is different, the identity function is replaced by a 1×1 convolution. (b), task heads which output C_t predictions. (p: padding, \oplus : feature fusion by element-wise addition).

References

- Hernandez, A.J.; Favaro, E.; Almeida, A.; Bonavides, A.; Demange, M.K.; Camanho, G.L. Reconstruction of the Medial Patellofemoral Ligament in Skeletally Immature Patients. *Tech. Knee Surg.* **2009**, *8*, 42–46. [CrossRef]
- Deie, M.; Ochi, M.; Sumen, Y.; Yasumoto, M.; Kobayashi, K.; Kimura, H. Reconstruction of the medial patellofemoral ligament for the treatment of habitual or recurrent dislocation of the patella in children. *J. Bone Joint Surg. Br.* **2003**, *85*, 887–890. [CrossRef]
- Diederichs, G.; Issever, A.S.; Scheffler, S. MR imaging of patellar instability: Injury patterns and assessment of risk factors. *Radiographics* **2010**, *30*, 961–981. [CrossRef] [PubMed]
- Longo, U.G.; others. Epidemiology of Posterior Cruciate Ligament Reconstructions in Italy: A 15-Year Study. *J. Clin. Med.* **2021**, *10*, 499. [CrossRef]
- Rünnow, A. The Dislocating Patella. *Acta Orthop. Scand.* **1983**, *54*, 1–53. [CrossRef]
- Stefancin, J.J.; Parker, R.D. First-time traumatic patellar dislocation: A systematic review. *Clin. Orthop. Relat. Res.* **2007**, *455*, 93–101. [CrossRef]
- Dornacher, D.; Lippacher, S.; Nelitz, M.; Reichel, H.; Ignatius, A.; Dürselen, L.; Seitz, A.M. Impact of five different medial patellofemoral ligament-reconstruction strategies and three different graft pre-tensioning states on the mean patellofemoral contact pressure: A biomechanical study on human cadaver knees. *J. Exp. Orthop.* **2018**, *5*, 25. [CrossRef]
- Arendt, E.A.; Fithian, D.C.; Cohen, E. Current concepts of lateral patella dislocation. *Clin. Sports Med.* **2002**, *21*, 499–519. [CrossRef]
- Askenberger, M.; Ekström, W.; Finnbogason, T.; Janarv, P.M. Occult Intra-articular Knee Injuries in Children With Hemarthrosis. *Am. J. Sports Med.* **2014**, *42*, 1600–1606. [CrossRef]
- Fithian, D.C.; Paxton, E.W.; Stone, M.L.; Silva, P.; Davis, D.K.; Elias, D.A.; White, L.M. Epidemiology and natural history of acute patellar dislocation. *Am. J. Sports Med.* **2004**, *32*, 1114–1121. [CrossRef]
- Mäenpää, H.; Lehto, M.U.K. Patellar dislocation. The long-term results of nonoperative management in 100 patients. *Am. J. Sports Med.* **1997**, *25*, 213–217. [CrossRef] [PubMed]
- Panni, A.S.; Cerciello, S.; Vasso, M. Patellofemoral instability: Surgical treatment of soft tissues. *Joints* **2013**, *1*, 34–39. [PubMed]
- Petri, M.; Ettinger, M.; Stuebig, T.; Brand, S.; Krettek, C.; Jagodzinski, M.; Omar, M. Current Concepts for Patellar Dislocation. *Arch. Trauma Res.* **2015**, *4*, e29301. [CrossRef] [PubMed]

14. Amis, A.A.; Firer, P.; Mountney, J.; Senavongse, W.; Thomas, N.P. Anatomy and biomechanics of the medial patellofemoral ligament. *Knee* **2003**, *10*, 215–220. [[CrossRef](#)]
15. Dragoo, J.L.; Nguyen, M.; Gatewood, C.T.; Taunton, J.D.; Young, S. Medial Patellofemoral Ligament Repair Versus Reconstruction for Recurrent Patellar Instability: Two-Year Results of an Algorithm-Based Approach. *Orthop. J. Sports Med.* **2017**, *5*, 2325967116689465. [[CrossRef](#)] [[PubMed](#)]
16. Camp, C.L.; Krych, A.J.; Dahm, D.L.; Levy, B.A.; Stuart, M.J. Medial patellofemoral ligament repair for recurrent patellar dislocation. *Am. J. Sports Med.* **2010**, *38*, 2248–2254. [[CrossRef](#)]
17. Buckens, C.F.M.; Saris, D.B.F. Reconstruction of the medial patellofemoral ligament for treatment of patellofemoral instability: A systematic review. *Am. J. Sports Med.* **2010**, *38*, 181–188. [[CrossRef](#)]
18. Nelitz, M.; Williams, S.R.M. Kombinierte Trochleoplastik und Rekonstruktion des medialen patellofemorales Ligaments zur Behandlung der patellofemorales Instabilität. *Oper. Orthop. Traumatol.* **2015**, *27*, 495–504. [[CrossRef](#)]
19. Shah, J.N.; Howard, J.S.; Flanigan, D.C.; Brophy, R.H.; Carey, J.L.; Lattermann, C. A systematic review of complications and failures associated with medial patellofemoral ligament reconstruction for recurrent patellar dislocation. *Am. J. Sports Med.* **2012**, *40*, 1916–1923. [[CrossRef](#)]
20. Vainionpää, S.; Laasonen, E.; Silvennoinen, T.; Vasenius, J.; Rokkanen, P. Acute dislocation of the patella. A prospective review of operative treatment. *J. Bone Joint Surg. Br.* **1990**, *72*, 366–369. [[CrossRef](#)]
21. Schöttle, P.B.; Schmeling, A.; Rosenstiel, N.; Weiler, A. Radiographic landmarks for femoral tunnel placement in medial patellofemoral ligament reconstruction. *Am. J. Sports Med.* **2007**, *35*, 801–804. [[CrossRef](#)]
22. Chahla, J.; Nitri, M.; Civitarese, D.; Dean, C.S.; Moulton, S.G.; LaPrade, R.F. Anatomic Double-Bundle Posterior Cruciate Ligament Reconstruction. *Arthrosc. Tech.* **2016**, *5*, e149–e156. [[CrossRef](#)] [[PubMed](#)]
23. Johannsen, A.M.; Anderson, C.J.; Wijedicks, C.A.; Engebretsen, L.; LaPrade, R.F. Radiographic landmarks for tunnel positioning in posterior cruciate ligament reconstructions. *Am. J. Sports Med.* **2013**, *41*, 35–42. [[CrossRef](#)] [[PubMed](#)]
24. LaPrade, R.F.; others. Double-Bundle Posterior Cruciate Ligament Reconstruction in 100 Patients at a Mean 3 Years' Follow-up: Outcomes Were Comparable to Anterior Cruciate Ligament Reconstructions. *Am. J. Sports Med.* **2018**, *46*, 1809–1818. [[CrossRef](#)] [[PubMed](#)]
25. Spiridonov, S.I.; Slinkard, N.J.; LaPrade, R.F. Isolated and combined grade-III posterior cruciate ligament tears treated with double-bundle reconstruction with use of endoscopically placed femoral tunnels and grafts: Operative technique and clinical outcomes. *J. Bone Joint Surg. Am.* **2011**, *93*, 1773–1780. [[CrossRef](#)] [[PubMed](#)]
26. Weimann, A.; Wolfert, A.; Zantop, T.; Eggers, A.K.; Raschke, M.; Petersen, W. Reducing the “killer turn” in posterior cruciate ligament reconstruction by fixation level and smoothing the tibial aperture. *Arthroscopy* **2007**, *23*, 1104–1111. [[CrossRef](#)]
27. Jackson, D.W.; Proctor, C.S.; Simon, T.M. Arthroscopic assisted PCL reconstruction: A technical note on potential neurovascular injury related to drill bit configuration. *Arthroscopy* **1993**, *9*, 224–227. [[CrossRef](#)]
28. Montgomery, S.R.; Johnson, J.S.; McAllister, D.R.; Petrigliano, F.A. Surgical management of PCL injuries: Indications, techniques, and outcomes. *Curr. Rev. Musculoskelet. Med.* **2013**, *6*, 115–123. [[CrossRef](#)]
29. Nicodeme, J.D.; Löcherbach, C.; Jolles, B.M. Tibial tunnel placement in posterior cruciate ligament reconstruction: A systematic review. *Knee Surg. Sports Traumatol. Arthrosc.* **2014**, *22*, 1556–1562. [[CrossRef](#)]
30. Yao, J.; others. Effect of tibial drill-guide angle on the mechanical environment at bone tunnel aperture after anatomic single-bundle anterior cruciate ligament reconstruction. *Int. Orthop.* **2014**, *38*, 973–981. [[CrossRef](#)]
31. Bertollo, N.; Walsh, W.R. Drilling of Bone: Practicality, Limitations and Complications Associated with Surgical Drill-Bits. In *Biomechanics in Applications*; Klika, V., Ed.; IntechOpen: Rijeka, Croatia, 2011; Chapter 3. [[CrossRef](#)]
32. Bernard, M.; Hertel, P.; Hornung, H.; Cierpinski, T. Femoral insertion of the ACL. Radiographic quadrant method. *J. Knee Surg.* **1997**, *10*, 14–21; discussion 21–22.
33. Amis, A.A.; Jakob, R.P. Anterior cruciate ligament graft positioning, tensioning and twisting. *Knee Surg. Sports Traumatol. Arthrosc.* **1998**, *6* (Suppl. 1), S2–S12. [[CrossRef](#)]
34. Stäubli, H.U.; Rauschnig, W. Tibial attachment area of the anterior cruciate ligament in the extended knee position. Anatomy and cryosections in vitro complemented by magnetic resonance arthrography in vivo. *Knee Surg. Sports Traumatol. Arthrosc.* **1994**, *2*, 138–146. [[CrossRef](#)] [[PubMed](#)]
35. Linte, C.A.; Moore, J.T.; Chen, E.C.; Peters, T.M. Image-Guided Procedures. In *Bioengineering for Surgery*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 59–90. [[CrossRef](#)]
36. Kubicek, J.; Tomanec, F.; Cerny, M.; Vilimek, D.; Kalova, M.; Oczka, D. Recent Trends, Technical Concepts and Components of Computer-Assisted Orthopedic Surgery Systems: A Comprehensive Review. *Sensors* **2019**, *19*, 5199. [[CrossRef](#)] [[PubMed](#)]
37. Joskowicz, L.; Hazan, E.J. Computer Aided Orthopaedic Surgery: Incremental shift or paradigm change? *Med. Image Anal.* **2016**, *33*, 84–90. [[CrossRef](#)] [[PubMed](#)]
38. Bai, L.; Yang, J.; Chen, X.; Sun, Y.; Li, X. Medical Robotics in Bone Fracture Reduction Surgery: A Review. *Sensors* **2019**, *19*, 3593. [[CrossRef](#)] [[PubMed](#)]
39. Innocenti, B.; Bori, E. Robotics in orthopaedic surgery: Why, what and how? *Arch. Orthop. Trauma Surg.* **2021**, *141*, 2035–2042. [[CrossRef](#)]
40. Jinnah, A.H.; Luo, T.D.; Plate, J.F.; Jinnah, R.H. General Concepts in Robotics in Orthopedics. In *Robotics in Knee and Hip Arthroplasty*; Lonner, J.H., Ed.; Springer International Publishing: Cham, Switzerland, 2019; pp. 27–35. [[CrossRef](#)]

41. Karthik, K.; Colegate-Stone, T.; Dasgupta, P.; Tavakkolizadeh, A.; Sinha, J. Robotic surgery in trauma and orthopaedics. *Bone Joint J.* **2015**, *97*, 292–299. [[CrossRef](#)]
42. Stübig, T.; Windhagen, H.; Krettek, C.; Ettinger, M. Computer-Assisted Orthopedic and Trauma Surgery. *Dtsch. Arztebl. Int.* **2020**, *117*, 793–800. [[CrossRef](#)]
43. Zhao, J.X.; Li, C.; Ren, H.; Hao, M.; Zhang, L.C.; Tang, P.F. Evolution and Current Applications of Robot-Assisted Fracture Reduction: A Comprehensive Review. *Ann. Biomed. Eng.* **2020**, *48*, 203–224. [[CrossRef](#)]
44. Ewurum, Chidozie H. and Guo, Yingying and Pagnha, Seang and Feng, Zhao and Luo, Xiongbiao. Surgical Navigation in Orthopedics: Workflow and System Review. In *Intelligent Orthopaedics: Artificial Intelligence and Smart Image-Guided Technology for Orthopaedics*; Zheng, G., Tian, W., Zhuang, X., Eds.; Springer Singapore: Singapore, 2018; pp. 47–63. [[CrossRef](#)]
45. Székely, G.; Nolte, L.P. Image guidance in orthopaedics and traumatology: A historical perspective. *Med. Image Anal.* **2016**, *33*, 79–83. [[CrossRef](#)]
46. Zheng, G.; Nolte, L.P., Computer-Aided Orthopaedic Surgery: State-of-the-Art and Future Perspectives. In *Intelligent Orthopaedics: Artificial Intelligence and Smart Image-guided Technology for Orthopaedics*; Springer Singapore: Singapore, 2018; pp. 1–20. [[CrossRef](#)]
47. Wang, M.; Li, D.; Shang, X.; Wang, J. A review of computer-assisted orthopaedic surgery systems. *Int. J. Med. Robot.* **2020**, *16*, 1–28. [[CrossRef](#)] [[PubMed](#)]
48. Casari, F.A.; Navab, N.; Hrubby, L.A.; Kriechling, P.; Nakamura, R.; Tori, R.; de Lourdes Dos Santos Nunes, F.; Queiroz, M.C.; Fürnstahl, P.; Farshad, M. Augmented Reality in Orthopedic Surgery Is Emerging from Proof of Concept Towards Clinical Studies: A Literature Review Explaining the Technology and Current State of the Art. *Curr. Rev. Musculoskelet. Med.* **2021**, *14*, 192–203. [[CrossRef](#)] [[PubMed](#)]
49. Essert, C.; Joskowicz, L. Chapter 32 - Image-based surgery planning. In *Handbook of Medical Image Computing and Computer Assisted Intervention*; Zhou, S.K., Rueckert, D., Fichtinger, G., Eds.; Academic Press, 2020; pp. 795–815. [[CrossRef](#)]
50. Kordon, F.; Maier, A.; Swartman, B.; Privalov, M.; El Barbari, J.S.; Kunze, H. Contour-Based Bone Axis Detection for X-ray Guided Surgery on the Knee. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Martel, A.L.; Abolmaesumi, P.; Stoyanov, D.; Mateus D.; Zuluaga, M.A.; Zhou, S.K.; Racoceanu, D.; Joskowicz, L. Eds.; Springer: Cham, Switzerland, 2020; Volume 12266, pp. 671–680. [[CrossRef](#)]
51. Bier, B.; Unberath, M.; Zaech, J.N.; Fotouhi, J.; Armand, M.; Osgood, G.; Navab, N.; Maier, A. X-ray-transform Invariant Anatomical Landmark Detection for Pelvic Trauma Surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; Volume 11073, pp. 55–63. [[CrossRef](#)]
52. Lee, S.M.; Kim, H.P.; Jeon, K.; Lee, S.H.; Seo, J.K. Automatic 3D cephalometric annotation system using shadowed 2D image-based machine learning. *Phys. Med. Biol.* **2019**, *64*, 055002. [[CrossRef](#)] [[PubMed](#)]
53. Hsu, W.E.; Yu, C.H.; Chang, C.J.; Wu, H.K.; Yu, T.H.; Tseng, C.S. C-Arm Image-Based Surgical Path Planning Method for Distal Locking of Intramedullary Nails. *Appl. Bionics Biomech.* **2018**, *2018*, 4530386. [[CrossRef](#)] [[PubMed](#)]
54. Kordon, F.; others. Multi-task Localization and Segmentation for X-Ray Guided Planning in Knee Surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Shen, D.; Liu, T., Peters, T.M., Staib, L.H., Essert, C., Zhou, S., Yap, P.-T., Khan, A., Eds.; Springer: Cham, Switzerland, 2019; Volume 11769, pp. 622–630. [[CrossRef](#)]
55. Kordon, F.; Maier, A.; Swartman, B.; Privalov, M.; El Barbari, J.S.; Kunze, H. Automatic Path Planning for Safe Guide Pin Insertion in PCL Reconstruction Surgery. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; de Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C., Eds.; Springer International Publishing: Cham, Switzerland, 2021; Volume 12904, pp. 560–570. [[CrossRef](#)]
56. Swartman, B.; Frere, D.; Wei, W.; Schnetzke, M.; Beisemann, N.; Keil, H.; Franke, J.; Grützner, P.; Vetter, S. 2D projection-based software application for mobile C-arms optimises wire placement in the proximal femur—An experimental study. *Injury* **2017**, *48*, 2068–2073. [[CrossRef](#)]
57. Lisa Kausch.; Moritz Scherer.; Sarina Thomas.; André Klein.; Fabian Isensee.; Klaus Maier-Hein. Automatic image-based pedicle screw planning. In *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling*; Linte, C.A., Siewerdsen, J.H., Eds.; SPIE: Bellingham, WA, USA, 2021, Volume 11598, pp. 406–414. [[CrossRef](#)]
58. Joskowicz, L., Future Perspectives on Statistical Shape Models in Computer-Aided Orthopedic Surgery: Beyond Statistical Shape Models and on to Big Data. In *Computer Assisted Orthopaedic Surgery for Hip and Knee: Current State of the Art in Clinical Application and Basic Research*; Springer Singapore: Singapore, 2018; pp. 199–206. [[CrossRef](#)]
59. Kamiya, N., Deep Learning Technique for Musculoskeletal Analysis. In *Deep Learning in Medical Image Analysis: Challenges and Applications*; Springer International Publishing: Cham, Switzerland, 2020; pp. 165–176. [[CrossRef](#)]
60. Benos, L.; Stanev, D.; Spyrou, L.; Moustakas, K.; Tsaopoulos, D.E. A Review on Finite Element Modeling and Simulation of the Anterior Cruciate Ligament Reconstruction. *Front. Bioeng. Biotechnol.* **2020**, *8*, 967. [[CrossRef](#)]
61. Colombet, P.; Robinson, J.; Christel, P.; Franceschi, J.P.; Djian, P.; Bellier, G.; Sbihi, A. Morphology of anterior cruciate ligament attachments for anatomic reconstruction: A cadaveric dissection and radiographic study. *Arthroscopy* **2006**, *22*, 984–992. [[CrossRef](#)]
62. Caruana, R., Multitask Learning. In *Learning to Learn*; Springer US: Boston, MA, USA, 1998; pp. 95–133. [[CrossRef](#)]
63. Ruder, S. An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv* **2017**, arXiv:1706.05098.
64. Thung, K.H.; Wee, C.Y. A brief review on multi-task learning. *Multimed. Tools Appl.* **2018**, *77*, 29705–29725. [[CrossRef](#)]
65. Zhang, Y.; Yang, Q. An overview of multi-task learning. *Natl. Sci. Rev.* **2018**, *5*, 30–43. [[CrossRef](#)]

66. Baxter, J. A Model of Inductive Bias Learning. *J. Artif. Int. Res.* **2000**, *12*, 149–198. [CrossRef]
67. Lin, B.; Ye, F.; Zhang, Y. A Closer Look at Loss Weighting in Multi-Task Learning. *arXiv* **2021**, arXiv:2111.10603.
68. Chen, Z.; Badrinarayanan, V.; Lee, C.Y.; Rabinovich, A. GradNorm: Gradient Normalization for Adaptive Loss Balancing in Deep Multitask Networks. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Dy, J., Krause, A., Eds.; PMLR: Stockholm, Sweden, 10–15 July 2018; Volume 80, pp. 794–803.
69. Liu, S.; Johns, E.; Davison, A.J. End-To-End Multi-Task Learning With Attention. In Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 1871–1880. [CrossRef]
70. Kendall, A.; Gal, Y.; Cipolla, R. Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics. In Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018.
71. Sener, O.; Koltun, V. Multi-Task Learning as Multi-Objective Optimization. In *Advances in Neural Information Processing Systems*; Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2018; Volume 31.
72. Yu, T.; Kumar, S.; Gupta, A.; Levine, S.; Hausman, K.; Finn, C. Gradient Surgery for Multi-Task Learning. In *Advances in Neural Information Processing Systems*; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H., Eds.; Curran Associates, Inc.: Red Hook, NY, USA 2020; Volume 33, pp. 5824–5836.
73. Wang, Z.; Tsvetkov, Y.; Firat, O.; Cao, Y. Gradient Vaccine: Investigating and Improving Multi-task Optimization in Massively Multilingual Models. *arXiv* **2021**, arXiv:2010.05874.
74. Chen, Z.; Ngiam, J.; Huang, Y.; Luong, T.; Kretzschmar, H.; Chai, Y.; Anguelov, D. Just Pick a Sign: Optimizing Deep Multitask Models with Gradient Sign Dropout. In *Advances in Neural Information Processing Systems*; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H., Eds.; Curran Associates, Inc.: Red Hook, NY, USA 2020; Volume 33, pp. 2039–2050.
75. Liu, L.; Li, Y.; Kuang, Z.; Xue, J.H.; Chen, Y.; Yang, W.; Liao, Q.; Zhang, W. Towards Impartial Multi-task Learning. In Proceedings of the International Conference on Learning Representations, Virtual Event, 3–7 May 2021.
76. Newell, A.; Yang, K.; Deng, J. Stacked Hourglass Networks for Human Pose Estimation. In *Computer Vision—ECCV 2016*; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer: Cham, Switzerland, 2016; pp. 483–499.
77. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity Mappings in Deep Residual Networks. In *Computer Vision—ECCV 2016*; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 630–645.
78. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
79. Greff, K.; Srivastava, R.K.; Schmidhuber, J. Highway and Residual Networks learn Unrolled Iterative Estimation. In Proceedings of the 5th International Conference on Learning Representations, ICLR, Toulon, France, 24–26 April 2017.
80. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9351, pp. 234–241.
81. Wada, K. Labelme: Image Polygonal Annotation with Python; 2016. Available online: <https://github.com/wkentaro/labelme> (accessed on 28 February 2022).
82. Kordon, F.; Lasowski, R.; Swartman, B.; Franke, J.; Fischer, P.; Kunze, H. Improved X-Ray Bone Segmentation by Normalization and Augmentation Strategies. In *Bildverarbeitung für Die Medizin 2019*; Handels, H., Deserno, T.M., Maier, A., Maier-Hein, K.H., Palm, C., Tolxdorff, T., Eds.; Informatik Aktuell, Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2019; pp. 104–109. [CrossRef]
83. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the Importance of Initialization and Momentum in Deep Learning. In Proceedings of the 30th International Conference on International Conference on Machine Learning, Atlanta, GA, USA, 17–19 June 2013.
84. Smith, L.N. Cyclical learning rates for training neural networks. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 24–31 March 2017; pp. 464–472.
85. Akkus, Z.; Galimzianova, A.; Hoogi, A.; Rubin, D.L.; Erickson, B.J. Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *J. Digit. Imaging* **2017**, *30*, 449–459. [CrossRef]
86. Hu, M.K. Visual pattern recognition by moment invariants. *IEEE Trans. Inf. Theory* **1962**, *8*, 179–187. [CrossRef]
87. Ziegler, C.G.; Fulkerson, J.P.; Edgar, C. Radiographic Reference Points Are Inaccurate with and without a True Lateral Radiograph: The Importance of Anatomy in Medial Patellofemoral Ligament Reconstruction. *Am. J. Sports Med.* **2016**, *44*, 133–142. [CrossRef]
88. Sanchis-Alfonso, V.; Ramírez-Fuentes, C.; Montesinos-Berry, E.; Elía, I.; Martí-Bonmatí, L. Radiographic Location Does Not Ensure a Precise Anatomic Location of the Femoral Fixation Site in Medial Patellofemoral Ligament Reconstruction. *Orthop. J. Sports Med.* **2017**, *5*, 2325967117739252. [CrossRef]
89. Kausch, L.; Thomas, S.; Kunze, H.; Privalov, M.; Vetter, S.; Franke, J.; Mahnken, A.H.; Maier-Hein, L.; Maier-Hein, K. Toward automatic C-arm positioning for standard projections in orthopedic surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2020**, *15*, 1095–1105. [CrossRef]
90. Kausch, L.; Thomas, S.; Kunze, H.; Norajitra, T.; Klein, A.; El Barbari, J.S.; Privalov, M.; Vetter, S.; Mahnken, A.; Maier-Hein, L.; Maier-Hein, K.H. C-Arm Positioning for Spinal Standard Projections in Different Intra-operative Settings. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; de Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 352–362.

91. Cho, W.J.; Kim, J.M.; Kim, D.E.; Lee, J.G.; Park, J.W.; Han, Y.H.; Seo, H.G. Accuracy of the femoral tunnel position in robot-assisted anterior cruciate ligament reconstruction using a magnetic resonance imaging-based navigation system: A preliminary report. *Int. J. Med. Robot. Comput. Assist. Surg.* **2018**, *14*, e1933. [[CrossRef](#)]
92. Morita, K.; Kobashi, S.; Kashiwa, K.; Nakayama, H.; Kambara, S.; Morimoto, M.; Yoshiya, S.; Aikawa, S. Computer-aided Surgical Planning of Anterior Cruciate Ligament Reconstruction in MR Images. *Procedia Comput. Sci.* **2015**, *60*, 1659–1667. [[CrossRef](#)]
93. Morita, K. Computer-aided Diagnosis Systems Based on Medical Image Registration. Ph.D. Thesis, University of Hyogo, Kobe, Japan, 2019.
94. Raposo, C.; Barreto, J.P.; Sousa, C.; Ribeiro, L.; Melo, R.; Oliveira, J.P.; Marques, P.; Fonseca, F.; Barrett, D. Video-based computer navigation in knee arthroscopy for patient-specific ACL reconstruction. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 1529–1539. [[CrossRef](#)] [[PubMed](#)]
95. Choi, J.H.; Lim, S.; Kim, Y.; Lee, D.; Park, S.; Park, S.; Wang, J.H. 3D preoperative surgical planning software for anterior cruciate ligament reconstruction. In Proceedings of the International Conference on Control, Automation and Systems (ICCAS), Gwangju, Korea, 20–23 October 2013; pp. 344–346. [[CrossRef](#)]
96. Le Huynh, M.; Kim, Y.H. A computer-aided and robot-assisted surgical system for reconstruction of anterior cruciate ligament. *Int. J. Precis. Eng. Manuf.* **2013**, *14*, 49–54. [[CrossRef](#)]
97. Kitamura, G.; Albers, M.B.V.; Lesniak, B.P.; Rabuck, S.J.; Musahl, V.; Andrews, C.L.; Ghodadra, A.; Fu, F. 3-Dimensional Printed Models May Be a Useful Tool When Planning Revision Anterior Cruciate Ligament Reconstruction. *Arthrosc. Sports Med. Rehab.* **2019**, *1*, e41–e46. [[CrossRef](#)]