



Deep Geometric Supervision Improves Spatial Generalization in Orthopedic Surgery Planning

Florian Kordon^{1,2,3} , Andreas Maier^{1,2} , Benedict Swartman⁴, Maxim Privalov⁴, Jan S. El Barbari⁴, and Holger Kunze³ 

¹ Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
florian.kordon@fau.de

² Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

³ Siemens Healthcare GmbH, Forchheim, Germany

⁴ Department for Trauma and Orthopaedic Surgery, BG Trauma Center Ludwigshafen, Ludwigshafen, Germany

Abstract. Careful surgical planning facilitates the precise and safe placement of implants and grafts in reconstructive orthopedics. Current attempts to (semi-)automatic planning separate the extraction of relevant anatomical structures on X-ray images and perform the actual positioning step using geometric post-processing. Such separation requires optimization of a proxy objective different from the actual planning target, limiting generalization to complex image impressions and the positioning accuracy that can be achieved. We address this problem by translating the geometric steps to a continuously differentiable function, enabling end-to-end gradient flow. Combining this companion objective function with the original proxy formulation improves target positioning directly while preserving the geometric relation of the underlying anatomical structures. We name this concept Deep Geometric Supervision. The developed method is evaluated for graft fixation site identification in medial patellofemoral ligament (MPFL) reconstruction surgery on (1) 221 diagnostic and (2) 89 intra-operative knee radiographs. Using the companion objective reduces the median Euclidean Distance error for MPFL insertion site localization from (1) 2.29 mm to 1.58 mm and (2) 8.70 px to 3.44 px, respectively. Furthermore, we empirically show that our method improves spatial generalization for strongly truncated anatomy.

The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the Bavarian State Ministry for Science and Art.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-16449-1_59.

Keywords: Surgical planning · Orthopedics · Landmark detection

1 Introduction

Careful planning of the individual surgical steps is an indispensable tool for the orthopedic surgeon, elevating the procedure’s safety and ensuring high levels of surgical precision [4, 6, 14, 15]. A surgical plan for routine interventions like ligament reconstruction describes several salient landmarks on a 2D X-ray image and relates them in a geometric construction [5, 8, 22]. Previous attempts to automate this planning type typically separate automatic feature localization with a learning algorithm and geometric post-processing [11–13]. The separation allows to mimic the manual step-wise workflow and enables granular control over each planning step. However, this approach comes with the drawbacks of optimizing a proxy criterion. While this surrogate has shown to be a low-error approximation of the actual planning target for well-aligned anatomy, the strength of correlation depends on the level of image truncation and the visibility of the contained radiographic landmarks [2, 24, 27]. In manual planning, the user compensates for these effects by extrapolating visual cues and using prior anatomical knowledge. As the learning algorithm has no direct access to this knowledge, the variance in correlation limits spatial generalization to unseen data with a broad range of image characteristics. In this work, we develop and analyze a companion objective function that optimizes the planning target directly. We exploit that the planning geometry can be formulated as a continuously differentiable function, enabling end-to-end gradient flow. Through the combination with the original optimization of anatomical feature localization, the relations of the planning geometry can be retained. We name this concept Deep Geometric Supervision (DGS). We test its effectiveness by studying the following research questions.

- RQ 1.** How does DGS affect the overall positioning accuracy?
- RQ 2.** Does DGS improve spatial generalisation on truncated images?
- RQ 3.** Can the potential improvements be applied to more complex imaging data in an intra-operative setting?

The developed method is evaluated for medial patellofemoral ligament (MPFL) reconstruction planning on diagnostic and intra-operative knee radiographs. This planning involves calculating the Schoettle Point (SP) [22], which determines the physiologically correct insertion point of the replacement graft and ensures long-term joint stability. We demonstrate that DGS significantly improves localization accuracy, increases the success rate of plannings to be within the required precision range of 2.5 mm, and enables generalization to severely truncated images.

2 Materials and Methods

2.1 Automatic Approach to Orthopedic Surgical Planning

We build on the two-stage planning method proposed by Kordon et al. [12, 13]. First, the positions of salient anatomical landmarks are automatically extracted

using a multitask learning (MTL) approach. Then, the landmarks are interrelated through geometric post-processing to locate the actual planning target.

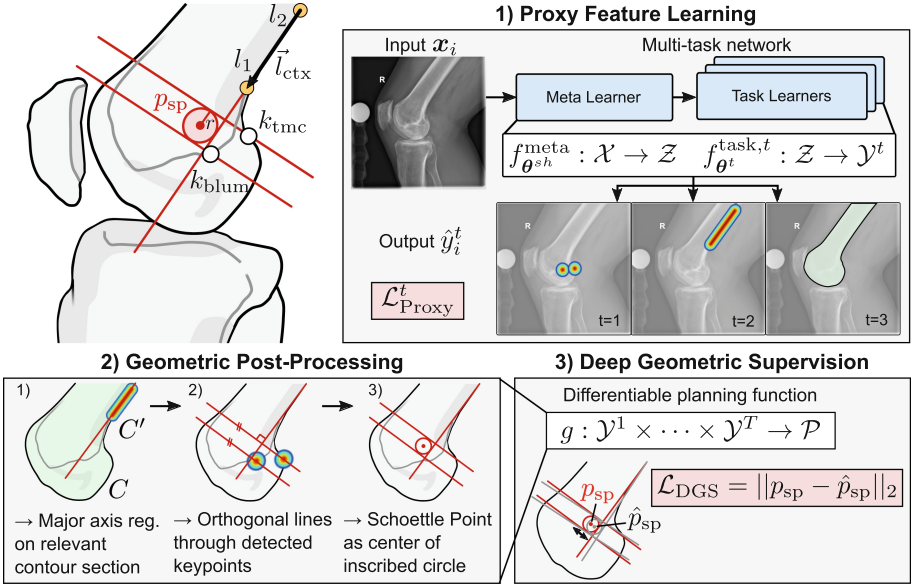


Fig. 1. Overview of the proposed method to automatic surgical planning with DGS.

In the first stage (Fig. 1-1), we want to optimize a mapping $f_{\theta}^t : \mathcal{X} \rightarrow \mathcal{Y}^t$ from the input domain \mathcal{X} to several task solution spaces $\{\mathcal{Y}^t\}_{t \in [T]}$. θ marks a set of trainable function parameters, and T is the number of parallel tasks to solve. The function f_{θ}^t is optimized in a supervised manner using M data-points $\{x_i, y_i^1, \dots, y_i^T\}_{i \in [M]}$ with ground truth y_i^t . To exploit similarities between tasks and maintain task-specific complexity at the same time, we employ hard parameter-sharing [3, 21]. Therefore, the model capacity θ is separated into disjoint sets of shared parameters θ^{sh} and task-specific parameters $\{\theta^t\}_{t \in [T]}$. According to Baxter [1], this separation can be interpreted as a subdivision of function f_{θ}^t into a meta learner $f_{\theta^{sh}}^{meta} : \mathcal{X} \rightarrow \mathcal{Z}$ and task-specific learners $f_{\theta^t}^{task,t} : \mathcal{Z} \rightarrow \mathcal{Y}^t$, such that the composition $(f_{\theta^t}^{task,t} \circ f_{\theta^{sh}}^{meta}) = f_{\theta^{sh}, \theta^t}^t : \mathcal{X} \rightarrow \mathcal{Y}^t$. The task-specific parameters can be trained by minimizing the loss function $\mathcal{L}_{Proxy}^t(\cdot, \cdot) : \mathcal{Y}^t \times \mathcal{Y}^t \rightarrow \mathbb{R}_0^+$. Following common practice in MTL literature, the shared parameters of the meta learner are optimized using a linear combination of all task losses [23]. Using this rationale, we arrive at the empirical risk minimization (ERM) objective

$$\min_{\theta^{sh}; \theta^1, \dots, \theta^T} \sum_{t=1}^T \hat{\mathcal{L}}_{Proxy}^t(\theta^{sh}, \theta^t), \quad (1)$$

where $\hat{\mathcal{L}}_{\text{Proxy}}^t(\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^t) \hat{=} \frac{1}{M} \sum_{i=1}^M \mathcal{L}_{\text{Proxy}}^t \left(f_{\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^t}^t(\mathbf{x}_i), y_i^t \right)$ is the task-specific empirical risk estimated on the training data. For the specific example of SP construction [22], we have to consider three distinct tasks (Fig. 1).

1. ($t=1$): *Keypoint detection* of the posterior Blumensaat line point k_{blum} and turning point on the medial femur condyle k_{tmc} . Both points are encoded as heatmaps sampled from a bivariate Gaussian function with mean at the point coordinates and standard deviation of $\sigma_{\text{hm}} = 6$ px. The correspondence between the predicted heatmap \hat{y} and ground truth heatmap y is optimized using pixel-wise mean squared error (MSE) loss function given by $\text{MSE}(\hat{y}, y) = \mathbb{E} \|\hat{y} - y\|_2^2$. Consequently, $\mathcal{L}_{\text{Proxy}}^1 := \text{MSE}(\hat{y}_i^1, y_i^1)$.
2. ($t=2$): *Line regression* of the tangent to the posterior femur shaft cortex \mathcal{I}_{ctx} . The line is encoded as a heatmap, where the intensities are computed by evaluating the point-to-line distances with a Gaussian function with $\sigma_{\text{hm}} = 6$ px [11, 13]. Similarly, $\mathcal{L}_{\text{Proxy}}^2 := \text{MSE}(\hat{y}_i^2, y_i^2)$.
3. ($t=3$): *Semantic segmentation* of the femur region S . As described in [11], S can be combined with the line heatmap to mask the relevant section $C' \subseteq C$ of segmentation contour $C \subseteq S$ for more precise positioning and angulation in the subsequent major axis regression [26] of relevant line points. For the loss function, we use a pixel-wise binary cross entropy (BCE) given by $\text{BCE}(\hat{y}, y) := -[y \log(\sigma(\hat{y})) + (1 - y) \log(1 - \sigma(\hat{y}))]$ with sigmoid nonlinearity σ . Consequently, $\mathcal{L}_{\text{Proxy}}^3 := \text{BCE}(\hat{y}_i^3, y_i^3)$.

After extracting the relevant features, they are converted to geometric primitives and interrelated according to the planning geometry (Fig. 1-2). This geometry describes consecutive calculations to localize the planning target relevant to the surgeon. For MPFL planning, the cortex tangent line is determined by major axis regression [26] on relevant contour points. The SP can be approximated by the center of the inscribed circle bounded by the tangent and two orthogonal lines intersecting both detected keypoints (Fig. 1).

2.2 Deep Geometric Supervision (DGS)

The disconnect between the proxy function and the actual planning target limits generalization to unfavorable but common image characteristics. We approach this issue by introducing the concept of Deep Geometric Supervision. To this end, we add a companion objective function to the original ERM term (Eq. 1) that directly minimizes positioning errors of the planning target while retaining the relations of the planning geometry (Fig. 1-3).

Mathematically, we start by combining all geometric steps in a single non-parametric function $g : \mathcal{Y}^1 \times \dots \times \mathcal{Y}^T \rightarrow \mathcal{P}$ that operates on the outputs of the anatomical feature extractor $f_{\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^t}^t$. The output of g is the desired planning target, e.g., a keypoint. Next, we calculate the positioning error of the planning target with the loss function $\mathcal{L}_{\text{DGS}}(\cdot, \cdot) : \mathcal{P} \times \mathcal{P} \rightarrow \mathbb{R}_0^+$. The empirical risk is given by $\hat{\mathcal{L}}_{\text{DGS}}(\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^1, \dots, \boldsymbol{\theta}^T) \hat{=} \frac{1}{M} \sum_{i=1}^M \mathcal{L}_{\text{DGS}} \left(g \left(f_{\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^1}^1(\mathbf{x}_i), \dots, f_{\boldsymbol{\theta}^{sh}, \boldsymbol{\theta}^T}^T(\mathbf{x}_i) \right), p_i \right)$.

Finally, adding this term to the original ERM formulation yields

$$\min_{\theta^{sh}; \theta^1, \dots, \theta^T} \lambda \sum_{t=1}^T \underbrace{\hat{\mathcal{L}}_{\text{Proxy}}^t(\theta^{sh}, \theta^t)}_{\text{Anatomical Features}} + (1 - \lambda) \underbrace{\hat{\mathcal{L}}_{\text{DGS}}(\theta^{sh}, \theta^1, \dots, \theta^T)}_{\text{Surgical Target}}, \quad (2)$$

where $\lambda \in \mathbb{R}$ is a multiplicative risk-weighting term. Here, $\mathcal{L}_{\text{DGS}} := \|\hat{p}_i - p_i\|_2$. Since the planning function g is not subject to trainable parameters, the minimization of the additional risk term $\hat{\mathcal{L}}_{\text{DGS}}$ should directly contribute to updates of θ^{sh} and θ^t . For that purpose, $(g \circ f_{\theta^{sh}, \theta^t}^t)$ must be a continuously differentiable function, such that $(g \circ f_{\theta^{sh}, \theta^t}^t) \in C^1(\mathcal{X}, \mathcal{P})$. To fulfill this constraint, the representations and objective functions for keypoint and line regression need to be changed from their original formulation in [13]. Matching the keypoint heatmaps with MSE is not feasible as it is typically followed by a subsequent non-differentiable argmax operation to extract the intensity peak’s x and y coordinates. We instead use the regularized spatial-to-numerical transform (DSNT) [18]. For that purpose, the predicted heatmap \hat{y} is rectified applying a spatial softmax and normalized with the L1 norm. The result of this standardization \hat{y}' is transformed to the numerical coordinate $\hat{c} = \text{DSNT}(\hat{y}')$ in the range of $[-1, 1]$ exploiting a probabilistic interpretation of \hat{y}' . This allows the cost function to operate directly on numerical coordinates and optimize the heatmaps implicitly. Finally, the keypoint loss function is updated to $\mathcal{L}_{\text{Proxy}}^1 := \|\hat{c}_i^1 - c_i^1\|_2 + D_{\text{JS}}(p(\hat{y}_i^1) \parallel \mathcal{N}(y_i^1, \sigma_{\text{hm}}^2 I_2))$. $p(\cdot)$ is a probability mass function under the interpretation of the predicted coordinates as discrete bivariate random vectors, and $D_{\text{JS}}(\cdot \parallel \cdot)$ is the Jensen-Shannon divergence which encourages similarity of the heatmaps to a Gaussian prior [18].

A differentiable representation of the line’s position and orientation is obtained by calculating raw image moments M_{pq} and second order central moments μ'_{pq} on the line heatmap [7, 20]. The centroid $c_{x,y}$ and orientation angle γ are given by $c_{x,y} = (M_{10}/M_{00}, M_{01}/M_{00})$ and $\gamma = \frac{1}{2} \arctan(2\mu'_{11}/(\mu'_{20} - \mu'_{02})) + \frac{\pi}{2} [\mu'_{20} < \mu'_{02}]$. $[\cdot]$ marks the Iverson bracket.

2.3 Model Variants

We define three model variants to evaluate the effect of DGS on planning accuracy and spatial generalization. *A) Proxy*, which uses all three anatomical detection tasks and uses the original ERM term Eq. 1 without DGS. The planning target is calculated with geometric post-processing. *B) Proxy-Seg+DGS*, which utilizes the updated ERM term Eq. 2. Here, g is used to calculate the planning target, and the segmentation task is omitted. *C) Proxy+DGS*, which uses ERM term Eq. 2 but keeps the segmentation task to allow direct comparison without the external factor of different task and parameter counts.

2.4 Datasets and Training Protocol

Method evaluation was done on two radiographic image cohorts of the lateral knee joint. *Cohort 1)* contains 221 diagnostic radiographs collected retrospectively

tively from anonymized databases. For each image, the SP geometry was annotated by an expert trauma surgeon with a proprietary tool by Siemens Healthcare GmbH. The femur polygon was labeled by a medical engineer using labelme software [25]. The images were split into three sets for training (167), validation (16), and testing (38) with stratified sampling [10]. There, all data showing a steel sphere of 30 mm diameter was assigned to the test split, allowing conversion from pixel to mm space. *Cohort 2*) contains 89 intra-operative X-ray images from 43 patients acquired with mobile C-arm systems. Most images show severely truncated bone shafts and instrumented anatomy. The images were annotated by a medical engineer using a custom extension of labelme [25]. The data was divided into training (61), validation (9), and test (19) with no patient overlap. During optimization, the training data was augmented using horizontal flipping ($p = 0.5$), rotation ($\alpha \in [-45^\circ, 45^\circ]$, $p = 1$), and scaling ($s \in [0.8, 1.2]$, $p = 1$) [13]. After min-max normalization of the images to the intensity range of $[0, 1]$, the data were standardized to the dimensions $[H:256 \times W:256]$ px by bicubic sampling and zero-padding to preserve the original aspect ratios. For each cohort and variant, an MTL hourglass network [13, 17] (128 feature root) was trained for 450 epochs using an Adam optimizer, a learning rate of 0.001/0.0006 (cohort 1/2), a batch size of 2, and multiplicative learning rate decay of 0.1 after 350 epochs. The risk-weighting with $\lambda = 0.99$ was re-balanced [9] by decreasing $\lambda \in [0.01, 0.99]$ by 0.01 every fourth epoch. Implementation was done in PyTorch v1.8 [19] (Python v3.8.12, CUDA v11.0) and reproducibility was confirmed.

3 Results

Effects on Feature Extraction and Positioning Accuracy (RQ 1). The results of model evaluation on Cohort 1) are summarized in Fig. 2. DGS reduces the median SP Euclidean Distance (ED) error from 2.29 $[1.84, 2.82]_{CI95}$ mm (A) to 1.68 $[1.19, 2.17]_{CI95}$ mm (B) and 1.58 $[1.152, 0.9]_{CI95}$ mm (C), respectively (Fig. 2-a). Between the two DGS variants, we observe no drop in performance when additionally solving the segmentation task. This understanding lets us reject insufficient model capacity or conflicting task configurations as a reason for inferior performance of the proxy optimization. Furthermore, DGS increases the number of predictions that fall within the clinically relevant precision range of 2.5 mm [22] from 63.2% (Model A) to 76.3% (Model B&C) (Fig. 2-b). Interestingly, DGS slightly changes the spatial appearance of the line heatmaps, increasing activations in the posterior aspect of the Blumensaat line (Fig. 2-c). Since this area resides on the tangential extension of the shaft cortex, we argue that activation in this region allows to compensate for small errors in line alignment.

Effects on Spatial Generalization (RQ2). To evaluate potential effects of DGS on spatial generalization, we constructed a secondary test set with different levels of shaft truncation. For this purpose, multiple crops per image were created such that the visible bone shaft corresponds to a fixed ratio $t \in [0.8, 2.7]$ between bone axis length and femur head width. The results are summarized in

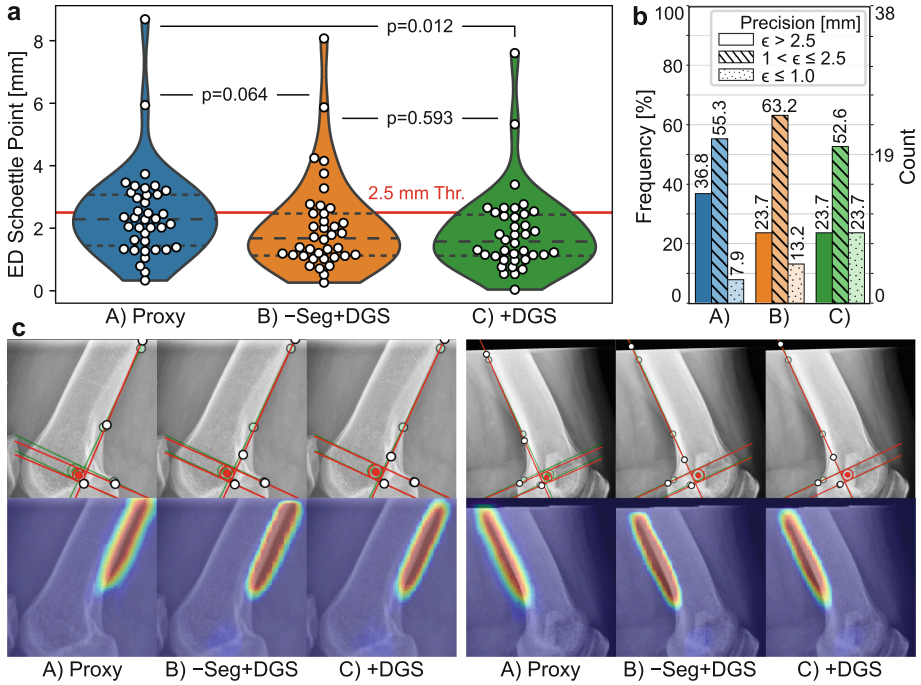


Fig. 2. Summary of model evaluation on Cohort 1). **a**, violin plots of the positioning errors on the test set (ED: Euclidean Distance; white dots mark individual planning samples). Statistical significance was evaluated with a two-sided Mann-Whitney U rank test. **b**, error distribution w.r.t. clinically relevant precision ranges. **c**, planning geometry (red: prediction, green: ground truth) and spatial appearance of line heatmaps. (Color figure online)

Fig. 3. We see considerable improvements for severely truncated shafts for both variants with DGS. This observation is underlined by a strong Spearman’s rank correlation of $r_s = -0.79$ for Model A compared to moderate correlations of $r_s = -0.48$ and $r_s = -0.53$ for the DGS Models B and C, respectively ($p \ll 0.01$ for all correlations). Visual inspection shows that the tangent determined via the bone contour experiences a systematic angular offset in the postero-distal direction for very short shaft lengths. Optimization with DGS can recover the optimal tangent direction in most of these cases despite insufficient image information.

Application to Complex Intra-Operative Data (RQ 3). The evaluation on the intra-operative Cohort 2) is summarized in Fig. 4. Similar to Cohort 1), DGS reduces the positioning error significantly, yielding median ED scores of 3.50 [2.47, 6.92]_{CI95} px and 3.44 [2.09, 6.67]_{CI95} px for Model A and B, respectively. The improvements over the original proxy formulation with a median error of 8.70 [5.06, 16.60]_{CI95} px can be explained by generally shorter shaft lengths caused by a smaller field of view of the mobile C-arm imaging device and less

standardized acquisition. As seen before, this characteristic leads to misaligned tangent predictions during proxy optimization. The compensation effect previously identified in the DGS variants, which is characterized by additional activation peaks in the distal region of the femur, is clearly enhanced (Fig. 4-c).

4 Discussion

Directly optimizing the planning target position while preserving the geometric relation of the anatomical structures promises a more precise, better generalizing, and clinically motivated planning automation in orthopedics. To accommodate this rationale, we developed and analyzed the concept of Deep Geometric Supervision. By interpreting the planning geometry as a differentiable function, the

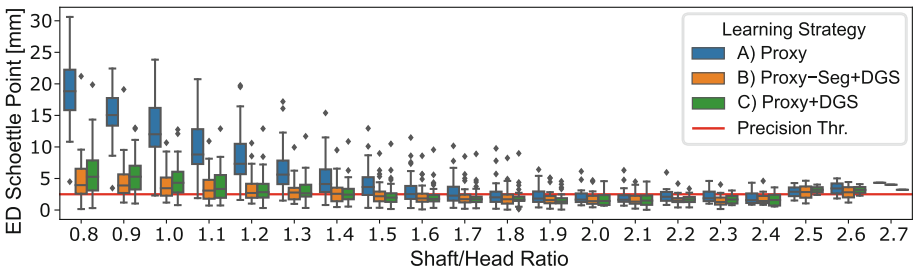


Fig. 3. Dependencies of the model variants to different levels of bone shaft truncation.

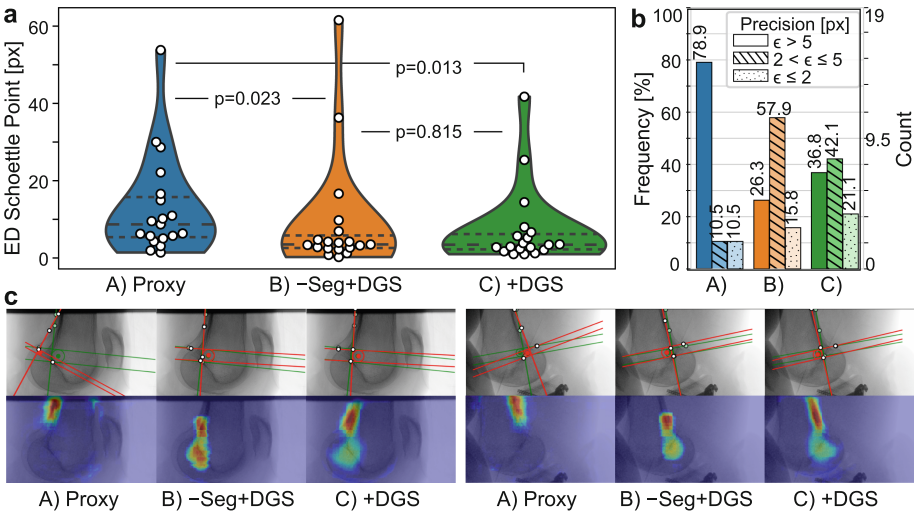


Fig. 4. Evaluation on Cohort 2). **a**, test set positioning errors. **b**, relevant precision classes. **c**, planning geometry (red: prediction, green: ground truth) and line heatmaps. (Color figure online)

planning target and the anatomical feature extractor can be optimized jointly. Improving target positioning accuracy while maintaining the core idea of step-wise geometric planning is a critical design decision that fosters clinical acceptance. Intriguingly, integrating the planning function into the computation graph can be interpreted as learning with a known operator [16], allowing end-to-end training and effectively reducing the upper error bound. In this context, it should be noted that minimizing only the DGS term yields a trivial solution where the extracted landmarks collapse to a single point at the planning target position. While these solutions offer competitive precision, they are undesirable because they do not mimic the clinically established planning workflow and cannot be easily verified for anatomic fidelity. An important trait of DGS is the improvement in spatial generalization. Especially in the intra-operative environment with constrained patient and device positioning, we cannot always expect standard acquisitions with sufficiently large bone shafts. There, DGS successfully bridges the semantic gaps present in the current proxy optimization strategy, reducing malpositioning when landmark visibility is limited. Besides these advantages, our current implementation of the planning function imposes little geometric constraints and, in theory, allows for different anatomical feature configurations that arrive at the same planning target. Reducing the space of possible solutions could ensure planning fidelity and smooth the optimization landscape.

Despite this limitation, our method effectively improves positioning accuracy and spatial generalization in orthopedic surgical planning. At the same time, it allows maintaining the clinically established planning geometry. We believe that these aspects facilitate the translation of planning automation concepts to the field and will ultimately motivate the development of new planning guidelines.

Disclaimer. The methods and information presented here are based on research and are not commercially available.

References

1. Baxter, J.: A model of inductive bias learning. *J. Artif. Int. Res.* **12**(1), 149–198 (2000)
2. Burrus, M.T., Werner, B.C., Cancienne, J.M., Diduch, D.R.: Correct positioning of the medial patellofemoral ligament: troubleshooting in the operating room. *Am. J. Orthop.* (Belle Mead N.J.) **46**(2), 76–81 (2017)
3. Caruana, R.: Multitask learning. In: Thrun, S., Pratt, L. (eds.) *Learning to Learn*, pp. 95–133. Springer, Boston (1998). https://doi.org/10.1007/978-1-4615-5529-2_5
4. Casari, F.A., et al.: 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.* **14**(2), 192–203 (2021). <https://doi.org/10.1007/s12178-021-09699-3>
5. Colombet, P., Robinson, J., Christel, P., Franceschi, J.P., Djian, P., Bellier, G., et al.: Morphology of anterior cruciate ligament attachments for anatomic reconstruction: a cadaveric dissection and radiographic study. *Arthroscopy* **22**(9), 984–992 (2006). <https://doi.org/10.1016/j.arthro.2006.04.102>

6. Essert, C., Joskowicz, L.: Image-based surgery planning. In: Zhou, S.K., Rueckert, D., Fichtinger, G. (eds.) *Handbook of Medical Image Computing and Computer Assisted Intervention*, pp. 795–815. The Elsevier and MICCAI Society Book Series. Academic Press (2020). <https://doi.org/10.1016/B978-0-12-816176-0.00037-5>
7. Hu, M.K.: Visual pattern recognition by moment invariants. *IEEE Trans. Inf. Theory* **8**(2), 179–187 (1962). <https://doi.org/10.1109/TIT.1962.1057692>
8. Johannsen, A.M., Anderson, C.J., Wijdicks, C.A., Engebretsen, L., LaPrade, R.F.: Radiographic landmarks for tunnel positioning in posterior cruciate ligament reconstructions. *Am. J. Sports Med.* **41**(1), 35–42 (2013). <https://doi.org/10.1177/0363546512465072>
9. Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., Ben Ayed, I.: Boundary loss for highly unbalanced segmentation. In: Cardoso, M.J., et al. (eds.) *International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research*, vol. 102, pp. 285–296. PMLR (2019)
10. Kordon, F., Lasowski, R., Swartman, B., Franke, J., Fischer, P., Kunze, H.: Improved X-ray bone segmentation by normalization and augmentation strategies. In: Handels, H., Deserno, T., Maier, A., Maier-Hein, K., Palm, C., Tolxdorff, T. (eds.) *Bildverarbeitung für die Medizin 2019. I*, pp. 104–109. Springer, Wiesbaden (2019). https://doi.org/10.1007/978-3-658-25326-4_24
11. 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: Martel, A.L., et al. (eds.) *MICCAI 2020. LNCS*, vol. 12266, pp. 671–680. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59725-2_65
12. 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* **8**(4), 108 (2022). <https://doi.org/10.3390/jimaging8040108>
13. Kordon, F., et al.: Multi-task localization and segmentation for X-Ray guided planning in knee surgery. In: Shen, D., et al. (eds.) *MICCAI 2019. LNCS*, vol. 11769, pp. 622–630. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32226-7_69
14. 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* **19**(23), 5199 (2019). <https://doi.org/10.3390/s19235199>
15. Linte, C.A., Moore, J.T., Chen, E.C., Peters, T.M.: Image-guided procedures. *Bioeng. Surg.* 59–90 (2016). <https://doi.org/10.1016/b978-0-08-100123-3.00004-x>
16. Maier, A., Syben, C., Stimpel, B., Würfl, T., Hoffmann, M., Schebesch, F., et al.: Learning with known operators reduces maximum error bounds. *Nat. Mach. Intell.* **2019**(1), 373–380 (2019). <https://doi.org/10.1038/s42256-019-0077-5>
17. Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016. LNCS*, vol. 9912, pp. 483–499. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46484-8_29
18. Nibali, A., He, Z., Morgan, S., Prendergast, L.: Numerical coordinate regression with convolutional neural networks. arXiv e-prints [arXiv:1801.07372](https://arxiv.org/abs/1801.07372), January 2018
19. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F.D., Fox, E., Garnett, R. (eds.) *Advances in Neural Information Processing Systems*, pp. 8026–8037. Curran Associates, Inc. (2019)

20. Rocha, L., Velho, L., Carvalho, P.: Image moments-based structuring and tracking of objects. In: XV Brazilian Symposium on Computer Graphics and Image Processing, pp. 99–105 (2002). <https://doi.org/10.1109/SIBGRA.2002.1167130>
21. Ruder, S.: An overview of multi-task learning in deep neural networks. arXiv e-prints [arXiv:1706.05098](https://arxiv.org/abs/1706.05098), June 2017
22. Schöttle, P.B., Schmeling, A., Rosenstiel, N., Weiler, A.: Radiographic landmarks for femoral tunnel placement in medial patellofemoral ligament reconstruction. *Am. J. Sports Med.* **35**(5), 801–804 (2007). <https://doi.org/10.1177/0363546506296415>
23. Sener, O., Koltun, V.: Multi-task learning as multi-objective optimization. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (eds.) *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc. (2018)
24. Tanaka, M.J., Chahla, J., Farr, J., LaPrade, R.F., Arendt, E.A., Sanchis-Alfonso, V., et al.: Recognition of evolving medial patellofemoral anatomy provides insight for reconstruction. *Knee Surg. Sports Traumatol. Arthrosc.* (2018). <https://doi.org/10.1007/s00167-018-5266-y>
25. Wada, K.: Labelme: image polygonal annotation with Python (2016)
26. Warton, D.I., Wright, I.J., Falster, D.S., Westoby, M.: Bivariate line-fitting methods for allometry. *Biol. Rev. Biol. Proc. Camb. Philos. Soc.* **81**(2), 259–291 (2006). <https://doi.org/10.1017/S1464793106007007>
27. 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.* **44**(1), 133–142 (2016). <https://doi.org/10.1177/0363546515611652>