

# Veriserum: A dual-plane fluoroscopic dataset with knee implant phantoms for deep learning in medical imaging

Jinhao Wang, Florian Vogl, Pascal Schütz, Saša Ćuković, and William R. Taylor\*

ETH Zürich, Laboratory for Movement Biomechanics, Institute for Biomechanics,  
Switzerland  
[jinhao.wang@hest.ethz.ch](mailto:jinhao.wang@hest.ethz.ch)

**Abstract.** Veriserum is an open-source dataset designed to support the training of deep learning registration for dual-plane fluoroscopic analysis. It comprises approximately 110,000 X-ray images of 10 knee implant pair combinations (2 femur and 5 tibia implants) captured during 1,600 trials, incorporating poses associated with daily activities such as level gait and ramp descent. Each image is annotated with an automatically registered ground-truth pose, while 200 images include manually registered poses for benchmarking.

Key features of Veriserum include dual-plane images and calibration tools. The dataset aims to support the development of applications such as 2D/3D image registration, image segmentation, X-ray distortion correction, and 3D reconstruction. Freely accessible, Veriserum aims to advance computer vision and medical imaging research by providing a reproducible benchmark for algorithm development and evaluation. The Veriserum dataset used in this study is publicly available via <https://movement.ethz.ch/data-repository/veriserum.html>, with the data stored at ETH Zürich Research Collections: <https://doi.org/10.3929/ethz-b-000701146>.

**Keywords:** Dual-plane fluoroscope · Medical imaging · Deep learning · 2D/3D image registration.

## 1 Introduction

2D/3D image registration based on X-ray imaging is a key requirement for various biomechanical and medical applications, ranging from joint research to surgery guidance systems [1,7,14]. Traditional registration algorithms typically rely on digitally reconstructed radiographs (DRRs), or rendering, which matches a rendered image of the 3D model to the target image by maximizing their similarity via optimization or manual vision [3,12,19,16]. While this approach has

---

\* Corresponding author: [bt@ethz.ch](mailto:bt@ethz.ch)

been effective, it faces computational complexity and robustness challenges, especially in clinical settings. Recently, deep learning approaches have achieved considerable success in image registration tasks by leveraging large datasets [9,17]. However, their application in X-ray image registration is constrained by the scarcity of high-quality medical data and stringent legal restrictions on data usage.

X-ray imaging presents practical challenges for registration purposes, including geometric distortion, limited resolution, and the precise measurement of source-intensifier distances [13]. Addressing these factors requires specialized calibration and distortion correction techniques, often lacking standardized datasets for benchmarking their performance. Despite these challenges, combining the strengths of deep learning and traditional registration methods plausibly holds strong potential to establish a robust, automatic 2D/3D registration pipeline [4,8,20,22]. This pipeline requires large-scale, high-quality datasets that incorporate realistic poses and provide ground-truth data for model evaluation.

Existing single-plane fluoroscopic datasets provide accurate pose labels for hip anatomy (DeepFluoro) [6] and for knee implants in total knee arthroplasty (CAMS-Knee, Stan) [18,2] where precise kinematics are available. However, dual-plane fluoroscopy has gained increasing attention over single-plane imaging due to its superior out-of-plane accuracy [7]. Despite its advantages, an open-source dual-plane fluoroscopic dataset for knee joint image registration that not only provides accurate kinematics but also addresses data privacy issues remains an unmet need. While many studies evaluate their algorithms on existing datasets, most research pipelines still rely on proprietary, unpublished code and resources, limiting reproducibility and comparability in the field.

In this work, we introduce Veriserum, a unique dataset of 110,990 dual-plane X-ray images of tibial and femoral knee implants in various poses associated with activities of daily living. Using a precision robot, we simulated realistic movement patterns by positioning modified implants to predefined positions based on motion data from patients with knee implants. Alongside the X-ray images, the dataset includes raw calibration phantom measurements and calibration functions to enable distortion correction and the calculation of source-to-intensifier distances.

By including dual-plane fluoroscopic images with tools for calibration correction, our dataset aims to a) provide the biomechanical research community with a high-quality dataset and ground-truth data to benchmark 2D/3D image registration techniques, b) support the computer vision community in developing novel automated methods for 2D/3D registration [21], image segmentation [20], or 3D reconstruction [11] specific to knee implants, and c) facilitate the development of clinically relevant applications, such as kinematics assessment and surgical guidance systems, by enabling a practical and reproducible image registration pipeline. Through these contributions, Veriserum provides a robust dataset for advancing research at the intersection of biomechanics, computer vision, and medical imaging.

## 2 Method

The experimental setup consists of a dual-plane fluoroscopic system and a precision robotic system with an implant model attached to the end-effector (Fig.1). This section describes the necessary elements involved in the data acquisition process.



Fig. 1: Schematic of the dual-plane X-ray imaging setup.

### 2.1 Dual-plane fluoroscope and X-ray system

The imaging system comprised two X-ray tubes, two collimators, and two X-ray image intensifiers connected to high-speed cameras, all constructed in a bi-planar configuration. The two X-ray planes were mounted at an angle of  $110^\circ \pm 0.5^\circ$ , with source-intensifier distances of 1850 and 1855 mm, and intensifier sizes of 360 mm. Image acquisition was performed at a tube voltage of 40 kV and a tube current of 25 mA on a large focus. Each fluoroscopic image had a resolution of  $1664 \times 1600$  pixels.

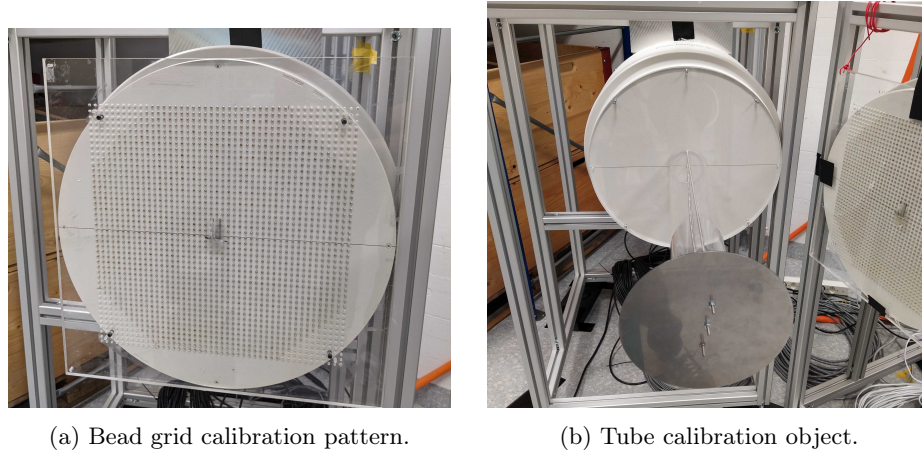


Fig. 2: Calibration objects used in the dual-plane X-ray setup.

## 2.2 Data acquisition and implant design

The study included two femoral and five tibial radiopaque implant designs manufactured from aluminium to mimic widely used geometries (Fig. 3). The implant geometries were adapted from CAD models provided by Zimmer Biomet (Winterthur, Switzerland) by smoothing all intricate geometries and adjusting non-load support wings while retaining the overall topology. A total of ten femur-tibia implant combinations were tested, with each combination following kinematic trajectories derived from the CAMS-Knee datasets [18,2]. The implant poses for each time instant replicated the complete trials of various activities.

Each component was securely mounted to the end-effector of an industrial-grade six-axis precision robot (HORST600, Fruitcore Robotics GmbH, Konstanz, Germany), operating with a reported precision of  $\pm 0.05$  mm. An adapter ensured proper positioning and prevented interference between the implant and the X-ray intensifier. The robotic system sequentially moved the implant to the predefined poses, holding each position for 500 ms before acquiring an X-ray image. This delay ensured that the end-effector remained stationary during imaging.

Data acquisition spanned ten measurement days, with one implant combination measured per day. Femur and tibia components were imaged separately to enable independent analysis as well as together to simulate complete knee implant configurations.

## 2.3 Calibration measurements

Prior to measurements, calibration procedures were performed to correct imaging distortions and establish accurate geometric relationships between the X-ray sources and intensifiers. Two key calibration processes were performed: distortion calibration (DISCAL) and source-intensifier calibration (SICAL).



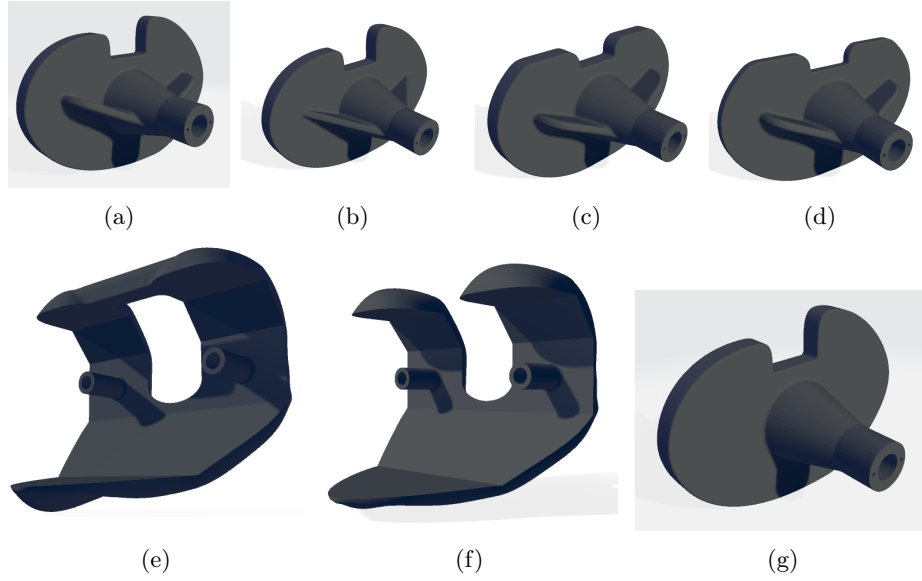


Fig. 3: Comparison of different tibial and femoral implant designs: (a) Asymmetric round wings with fixed bearing, (b) Asymmetric straight wings with fixed bearing, (c) Symmetric round wings with fixed bearing, (d) Symmetric round wings with mobile bearing, (e) Femur crossbar design, (f) Femur normal design, (g) Asymmetric no-wings with fixed bearing.

For distortion correction, a bead grid phantom ( $45 \times 45$  beads with 7 mm spacing) was affixed to the intensifier surface (Fig. 2a). The known geometry of the bead grid enabled precise distortion correction using a coherent point drift method to establish correspondences, followed by third-order polynomial fitting with Powell optimisation to compute the transformation.

For source-intensifier calibration (SICAL), a tube phantom was positioned 500 mm in front of the intensifier. This phantom featured a 200 mm diameter metal plate, concentrically aligned with the intensifier (Fig. 2b), allowing accurate geometric calibration of the imaging system. The relative position of the X-ray source was determined by analyzing the projected shadows of the circular metal plate.

All resulting DISCAL and SICAL values, as well as the raw distortion grid images (Fig. 4a&4b), phantom measurements, and precomputed correction functions, were stored and are available within the Veriserum database. The database also includes code for computing and applying DISCAL and SICAL functions. While precomputed calibration values are available in the database for immediate application, users are welcome to reprocess the calibration parameters if necessary.

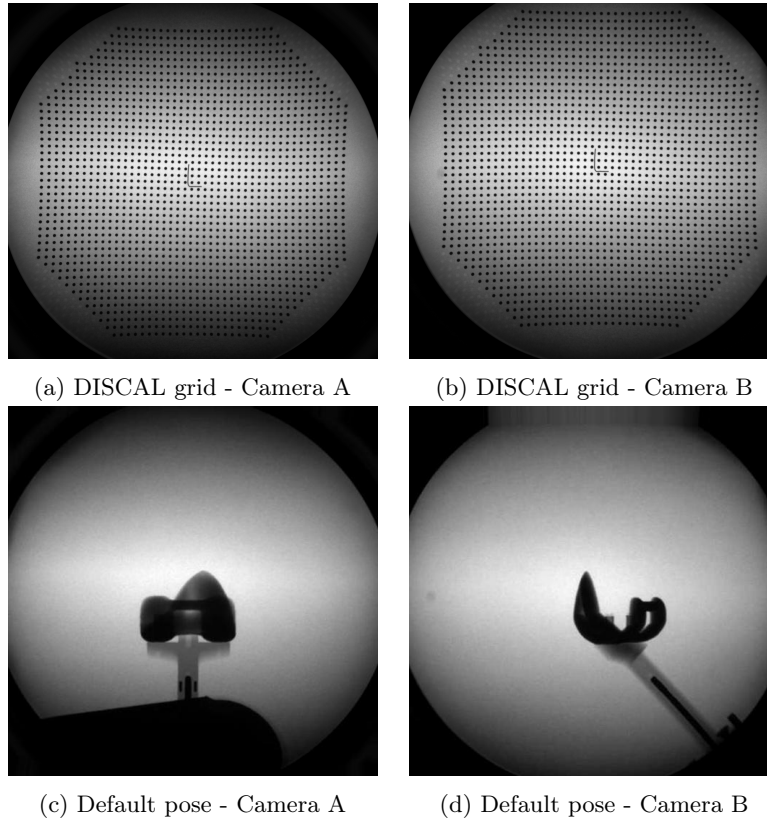


Fig. 4: Calibration images used in the experimental setup.

## 2.4 Poses

A total of 110,990 implant poses were recorded, covering movements from level walking (28k), stair descent (25k), ramp descent (14k), and chair sitting (42k) activities. Poses were derived from the CAMS-Knee kinematics database. Each pose corresponded to a predefined end-effector position provided to the robotic system.

Despite high robotic precision, minor deviations in the accuracy of reproducing the target pose occurred due to inherent robotic control system limitations. An automated image registration pipeline (Section 3.2) estimated the true implant pose by finding the rigid transformation between the rendered implant projection and the acquired fluoroscopic image. Additionally, a subset of 200 images was manually registered using a self-designed image registration software, ensuring correct calibrations. These manually registered poses served as a baseline for evaluating the automated registration performance.

### 3 Experiments and Results

The published dataset includes X-ray images with corresponding target robot poses (110k), corresponding automatically registered poses for all images, and manually registered poses for a subset of 200 randomly selected images.

#### 3.1 Data Collection and Repeatability

Data was collected over ten days, each day dedicated to a different femur-tibia implant combination. As part of quality assurance to check precision, the robotic system repositioned each implant to its default pose between each trial (Fig.4c&4d). The average 2D correlations between the default pose images were computed using Normalized Cross Correlation (NCC) to ensure the repeatability of the data [5]. Across all measurement days, the correlation scores consistently exceeded 0.985, demonstrating a high inter-day positional repeatability of the robotic system. However, despite this high precision, discrepancies remained between the programmed target robot pose and the true implant pose in the fluoroscopic images due to minor inaccuracies inherent in the robotic control system (Fig. 5).

#### 3.2 Automated Pose Registration

We developed a novel differentiable surface renderer using PyTorch3D for automated implant registration [10,15]. This module is compatible with PyTorch, supports gradient-based optimization, and is designed for future integration with AI-based networks. The open-source code is publicly available within the Veriserum dataset to support the research community.

Since the rendered projections of the initial robot pose did not fully align with the fluoroscopic images, estimation of the true implant pose was necessary for accurate dataset annotation. Our differentiable renderer was used together with our automated registration method to refine the initial robot-based positioning and obtain the matched pose.

Each robot pose was optimised using the ADAM optimiser with NCC as the similarity metric, over 200 steps and a learning rate of 0.25. To assess registration accuracy, meticulously manually registered poses served as the gold standard reference. Translation was evaluated using in-plane L1 loss function, and rotation via the geodesic angle that measures the shortest path between two transformations on the rotation manifold.

In approximately 65% of the datasets, the error difference between the target robot pose and the manually registered ground truth exceeded 2.5 mm in L1 translational error and 1° in geodesic rotational error (Fig.5). In contrast, the automatically registered poses achieved submillimeter accuracy compared to ground truth, with over 80% images showing errors below 1 mm and 1°. This result demonstrates that the automated registration pipeline significantly improves pose accuracy ( $MAE < 0.8$  mm,  $0.9^\circ$ ), producing results comparable to

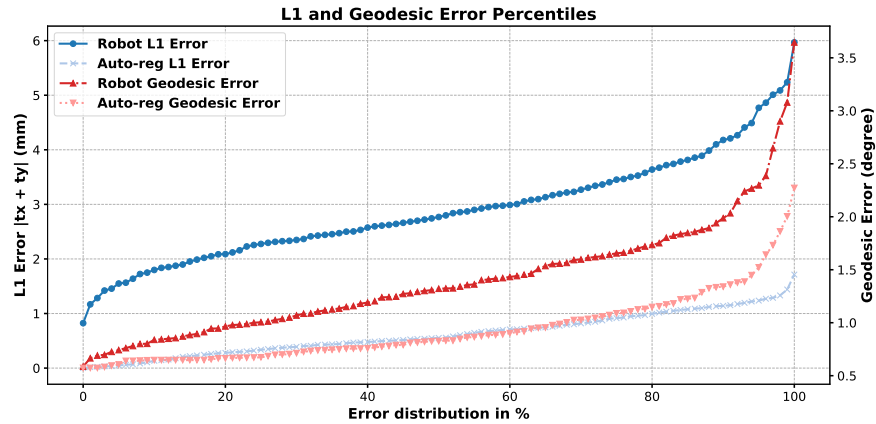


Fig. 5: Error percentile plot of target robot poses and automatically registered poses (Auto-reg), compared to manual ground truth.

time-intensive manual registration while offering a scalable and reproducible solution. Once this dataset is openly available to the community, both manually registered poses and automated registered poses can serve as ground truth for deep learning training.

## 4 Conclusion and Discussion

This study presents Veriserum, the first open-source, high-quality dual-plane fluoroscopic dataset for knee implant analysis. It includes fluoroscopic images of diverse, clinically relevant poses, with both automated and manually registered pose annotations obtained through systematic calibration, making it a strong benchmark for evaluating 2D/3D image registration. Its pose diversity promotes robustness for enhancing data-driven methods [21,9,20], and segmentation masks generated via rendering can assist implant segmentation and pose initialization [8]. Veriserum, therefore, supports downstream tasks such as image registration, implant segmentation, and 3D reconstruction in post-operative patients [11].

While Veriserum provides a valuable resource for medical imaging research, the absence of soft-tissue structures and real patient anatomy may limit the generalisability of deep learning models trained using these datasets. Additionally, since Veriserum time-series were reconstructed from datasets originally captured at only 25 or 30 Hz, its use in supporting time-series analysis maybe limited in some cases. Finally, the original fluoroscopic images were captured using analogue source-intensifier technology rather than digital flat panels, and hence spatial pixel errors were relatively large. Future work could expand the dataset to include natural bones with soft tissues, upgrade the X-ray system with a flat

panel technology to avoid DISCAL-related errors, and increase the frequency of dynamic fluoroscopic sequences.

With its open availability, Veriserum is intended to advance biomechanics and AI-assisted fluoroscopic image analysis. As a comprehensive and reproducible benchmark, it enables development of automated 2D/3D registration, segmentation, and 3D reconstruction methods. As deep learning evolves, Veriserum represents a step towards bridging traditional model-based registration and AI-driven solutions.

**Acknowledgments.** The authors acknowledge the developers of the open-source tools used in this study, including PyTorch3D. Additionally, we would like to thank our collaborator HorstCosmos for providing the Fruitcore Robotics system used in the experiments.

**Disclosure of Interests.** The authors declare no competing interests.

## References

1. Chen, M., Li, T., Zhang, Z., Kong, Y.: An optimization-based baseline for rigid 2d/3d registration applied to spine surgical navigation using cma-es (2 2024), <http://arxiv.org/abs/2402.05642>
2. Dreyer, M.J., Trepczynski, A., Nasab, S.H.H., Kutzner, I., Schütz, P., Weisse, B., Dymke, J., Postolka, B., Moewis, P., Bergmann, G., Duda, G.N., Taylor, W.R., Damm, P., Smith, C.R.: European society of biomechanics s.m. perren award 2022: Standardized tibio-femoral implant loads and kinematics. *Journal of Biomechanics* **141** (8 2022). <https://doi.org/10.1016/j.jbiomech.2022.111171>
3. Flood, P.D., Banks, S.A.: Automated registration of 3-d knee implant models to fluoroscopic images using lipschitzian optimization. *IEEE Transactions on Medical Imaging* **37**, 326–335 (1 2018). <https://doi.org/10.1109/TMI.2017.2773398>
4. Gopalakrishnan, V., Dey, N., Golland, P.: Intraoperative 2d/3d image registration via differentiable x-ray rendering (12 2023), <http://arxiv.org/abs/2312.06358>
5. Grupp, R.B., Armand, M., Taylor, R.H.: Patch-based image similarity for intra-operative 2d/3d pelvis registration during periacetabular osteotomy. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. vol. 11041 LNCS, pp. 153–163. Springer Verlag (2018). [https://doi.org/10.1007/978-3-030-01201-4\\_17](https://doi.org/10.1007/978-3-030-01201-4_17)
6. Grupp, R.B., Unberath, M., Gao, C., Hegeman, R.A., Murphy, R.J., Alexander, C.P., Otake, Y., McArthur, B.A., Armand, M., Taylor, R.H.: Automatic annotation of hip anatomy in fluoroscopy for robust and efficient 2d/3d registration. *International Journal of Computer Assisted Radiology and Surgery* **15**, 759–769 (5 2020). <https://doi.org/10.1007/s11548-020-02162-7>
7. Guan, S., Gray, H.A., Keynejad, F., Pandey, M.G.: Mobile biplane x-ray imaging system for measuring 3d dynamic joint motion during overground gait. *IEEE Transactions on Medical Imaging* **35**, 326–336 (1 2016). <https://doi.org/10.1109/TMI.2015.2473168>
8. Jensen, A.J., Flood, P.D., Palm-Vlasak, L.S., Burton, W.S., Chevalier, A., Rulkoetter, P.J., Banks, S.A.: Joint track machine learning: An autonomous method of measuring total knee arthroplasty kinematics from single-plane x-ray images. *Journal of Arthroplasty* (10 2023). <https://doi.org/10.1016/j.arth.2023.05.029>

9. Li, Y., Wang, G., Ji, X., Xiang, Y., Fox, D.: Deepim: Deep iterative matching for 6d pose estimation (3 2018). <https://doi.org/10.1007/s11263-019-01250-9>, <http://arxiv.org/abs/1804.00175><http://dx.doi.org/10.1007/s11263-019-01250-9>
10. Liu, S., Li, T., Chen, W., Li, H.: Soft rasterizer: A differentiable renderer for image-based 3d reasoning (4 2019), <http://arxiv.org/abs/1904.01786>
11. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space (2019), <https://arxiv.org/abs/1812.03828>
12. Petersen, E.T., Vind, T.D., Jürgens-Lahnstein, J.H., Christensen, R., de Raedt, S., Brüel, A., Rytter, S., Andersen, M.S., Stilling, M.: Evaluation of automated radiostereometric image registration in total knee arthroplasty utilizing a synthetic-based and a ct-based volumetric model. *Journal of Orthopaedic Research* **41**, 436–446 (2 2023). <https://doi.org/10.1002/jor.25359>
13. Postolka, B., Schütz, P., Fucntese, S.F., Freeman, M.A., Pinskerova, V., List, R., Taylor, W.R.: Tibio-femoral kinematics of the healthy knee joint throughout complete cycles of gait activities. *Journal of Biomechanics* **110** (9 2020). <https://doi.org/10.1016/j.jbiomech.2020.109915>
14. Postolka, B., Taylor, W.R., List, R., Fucntese, S.F., Koch, P.P., Schütz, P.: Isb clinical biomechanics award winner 2021: Tibio-femoral kinematics of natural versus replaced knees – a comparison using dynamic videofluoroscopy. *Clinical Biomechanics* **96** (6 2022). <https://doi.org/10.1016/j.clinbiomech.2022.105667>
15. Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, W.Y., Johnson, J., Gkioxari, G.: Accelerating 3d deep learning with pytorch3d (7 2020), <http://arxiv.org/abs/2007.08501>
16. Stentz-Olesen, K., Nielsen, E.T., de Raedt, S., Jørgensen, P.B., Sørensen, O.G., Kaptein, B., Søballe, K., Stilling, M.: Reconstructing the anterolateral ligament does not decrease rotational knee laxity in acl-reconstructed knees. *Knee Surgery, Sports Traumatology, Arthroscopy* **25**, 1125–1131 (4 2017). <https://doi.org/10.1007/s00167-017-4500-3>
17. Stevsic, S., Hilliges, O.: Spatial attention improves iterative 6d object pose estimation (1 2021), <http://arxiv.org/abs/2101.01659>
18. Taylor, W.R., Schütz, P., Bergmann, G., List, R., Postolka, B., Hitz, M., Dymke, J., Damm, P., Duda, G., Gerber, H., Schwachmeyer, V., Nasab, S.H.H., Trepczynski, A., Kutzner, I.: A comprehensive assessment of the musculoskeletal system: The cams-knee data set. *Journal of Biomechanics* **65**, 32–39 (12 2017). <https://doi.org/10.1016/j.jbiomech.2017.09.022>
19. Tsai, T.Y., Lu, T.W., Kuo, M.Y., Lin, C.C.: Effects of soft tissue artifacts on the calculated kinematics and kinetics of the knee during stair-ascent. *Journal of Biomechanics* **44**, 1182–1188 (4 2011). <https://doi.org/10.1016/j.jbiomech.2011.01.009>
20. Vogl, F., Schutz, P., Postolka, B., List, R., Taylor, W.: Personalised pose estimation from singleplane moving fluoroscope images using deep convolutional neural networks. *PLoS ONE* **17** (6 2022). <https://doi.org/10.1371/journal.pone.0270596>
21. Wang, C., Xie, S., Li, K., Wang, C., Liu, X., Zhao, L., Tsai, T.Y.: Multi-view point-based registration for native knee kinematics measurement with feature transfer learning. *Engineering* **7**, 881–888 (6 2021). <https://doi.org/10.1016/j.eng.2020.03.016>



22. Zhang, B., Faghihroohi, S., Azampour, M.F., Liu, S., Ghotbi, R., Schunkert, H., Navab, N.: A patient-specific self-supervised model for automatic x-ray/ct registration. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 14228 LNCS, pp. 515–524. Springer Science and Business Media Deutschland GmbH (2023). [https://doi.org/10.1007/978-3-031-43996-4\\_49](https://doi.org/10.1007/978-3-031-43996-4_49)