



A Deep Learning Approach for Single Shot C-Arm Pose Estimation

Hooman Esfandiari¹, Sebastian Andreß², Maternus Herold³, Wolfgang Böcker²,
Simon Weidert², and Antony J. Hodgson¹

¹ University of British Columbia, Biomedical Engineering, Vancouver, Canada
hooman.esfandiari@ubc.ca

² University Hospital, LMU Munich, Department of General, Trauma and Reconstructive Surgery,
Munich, Germany

³ LMU Munich, Faculty of Mathematics, Informatics and Statistics, Munich, Germany

Abstract

During a typical fluoroscopic guided surgery, it is common to acquire multiple x-ray images to correctly position the C-arm. This can be a long process resulting in an increase in operation time and ionizing radiation exposure. Our purpose in this study is to implement a machine learning system for predicting the position of the C-arm based on the intraoperative radiographs. The prediction is achieved by training a Deep Learning Network based on Digitally Reconstructed Radiographs. We first showed a high prediction accuracy (4.5 mm and 1.1°) when patient-specific training was implemented. Additionally, we demonstrated a similar range of accuracy by applying transfer-learning on the last layers of the network while reducing the processing time by 83%. In conclusion, in this study, we propose a C-arm position prediction system based on machine learning that can potentially reduce the number of intraoperatively acquired X-rays in a common orthopaedic surgical procedure.

1 Introduction

Fluoroscopic images are frequently used in orthopedic surgery. In a typical fluoroscopic-guided surgery often the surgeon has to iterate between multiple views, as important anatomical structures must be viewed from different points of view to better estimate the three-dimensional structures and surgical tool positions based on the two-dimensional radiographs.

Previous studies have shown that predicting the fluoroscopy image that can be acquired at a given location can lead to a reduction in radiation dose to the patient [1]. It was shown in [5] that it is possible to generate an accurate, scaled 3D model of a patient's pelvis based on a single anterior-posterior radiograph; however, because only a single shot was used and no object of known size was present in the image, this technique was unable to resolve the scale. More recently, it was shown that it is possible to identify anatomical landmarks on intraoperative X-rays using a Convolutional Neural Network [3], which can assist the surgeons in better estimating the 3D position of the desired anatomy based on the 2D radiographs.

In this study, we propose to train a network on a pre-operatively available Computed Tomographic (CT) volumeset and to use it to recognize the position of the C-Arm based on a single X-ray image. Such a network could be used in combination with a C-arm tracking system to predict in real time the location of the C-arm relative to the CT volume.

2 Methods

We collected a dataset consisting of 22 pelvis CTs that were separated into training (12 males, 2 females), cross-validation (2 males, 1 female), and test sets (4 males, 1 female). All CT volumes were manually aligned to a neutral position corresponding to a nominally lateral view of the pelvis as might typically be used in an S1 iliosacral screw placement procedure. Based on this target application, we defined a range of clinically relevant simulated C-arm views:

- Anterior-posterior (AP) and cranio-caudal (CC) translations (movement): ± 50 mm
- Mediolateral (ML) translation: ± 200 mm
- AP and CC axis rotation (pitch and yaw): ± 22.5 degree
- ML axis rotation (roll): ± 45 degree

We used VGG19 as the core architecture of our model. Accuracy assessment was done using the testset CT scans and creating 333 new DRR images per volume. We then conducted the following analyses:

- Patient-Specific Model (PSM): we trained one model per CT volume in the testset based on 30,000 DRR images specific to that CT
- Non-Patient-Specific Model (NPSM): a model was trained based on DRR images generated from all volumes in the training set ($30,000 \times 14$).
- Patient-Tuned Model (PTM): the NPSM was used as the basis for transfer learning. For this, we re-trained the last convolutional layer based on only 3,000 DRRs created from each of the test CTs (representing the target patients).

Image generation and training times were recorded (GPU 10 GB, CPU 8 cores x 2.1 Ghz).

3 Results

The results from the three models trained in this study are reported in Figure 1 and Table 1. The PSM achieved an average prediction error of 4.5 mm and 1.1° . The NPSM had considerably higher errors (on average 15.3 mm and 4.5°), but the PTM (transfer learning-based) performed similarly to the PSM, with an average error of 4.9 mm and 1.3° , despite only requiring a short training time (up to 83.4% less than the PSM).

4 Discussion and Conclusion

The PS and PT models presented in this study were able to predict the C-arm position with sufficient accuracy (with errors < 5 mm and 1.3°). A pre-trained model (NPSM) applied to

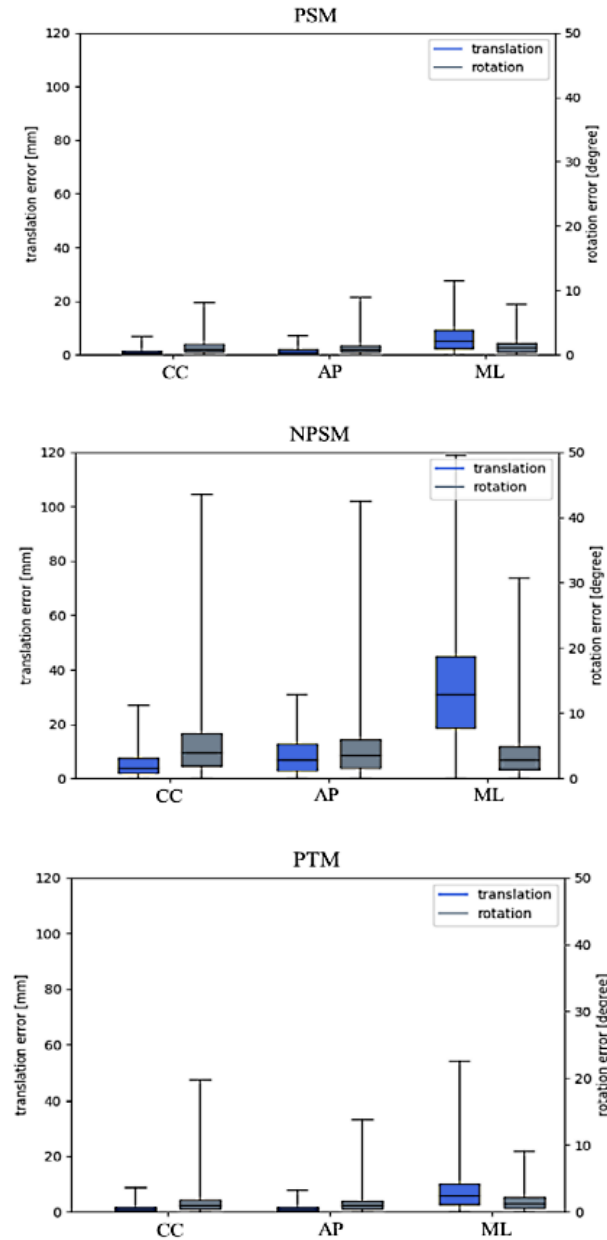


Figure 1: Errors of the position predictions for the three models.

images from a previously-unseen patient had markedly lower spatial prediction accuracy (especially along the imaging axis, likely because there was little scale information available in the image).

However, we demonstrated that applying transfer learning with a small number of images drawn from a patient-specific CT allows us to markedly reduce the preoperative training time with

Table 1: Comparison of the processing times and mean errors of all three experiments.

	PSM	NPSM	PTM	
Image Generation Time [hh:mm:ss]	00:50:46	NA	00:05:09	
Training Time [hh:mm:ss]	03:51:34	NA	00:11:54	
Epochs	42	NA	41	
Time per Epoch [hh:mm:ss]	00:05:27	NA	00:00:16	
Translation Error [mm]	CC	0.99	5.05	1.25
	AP	1.37	8.13	1.26
	ML	6.43	32.77	7.45
Rotation Error [°]	CC	1.14	5.50	1.33
	AP	1.04	4.76	1.28
	ML	1.31	3.49	1.57

minimal loss in position prediction accuracy.

Compared to C-arm pose estimation methods that require external calibration objects such as [2], our model achieved lower, though likely still acceptable, accuracy (~ 5 mm in our study vs 0.4 mm in [2], though the metrics used in the two studies are somewhat different in nature) but does not require use of a calibration object or additional processing steps during the surgical procedure.

One limitation of the current study is that the validation has only been conducted on DRR images; therefore a natural future step would be to try training the model with more realistic X-ray images based on the deep DRR approach introduced in [4] or other more refined DRR simulators and testing based on real clinical-grade X-rays.

References

- [1] Peter Belei, Adrian Skwara, Matías De La Fuente, Erik Schkommodau, Susanne Fuchs, Dieter-C. Wirtz, Christian Kämper, and Klaus Radermacher. Fluoroscopic navigation system for hip surface replacement. *Computer Aided Surgery*, 12(3):160–167, January 2007.
- [2] Hooman Esfandiari, Jose F Martinez, Alba G Alvarez, Pierre Guy, John Street, Carolyn Anglin, and Antony J Hodgson. An automated, robust and closed form mini-RSA system for intraoperative C-Arm calibration. In *CARS 2017 - Computer Assisted Radiology and Surgery*, volume 12 (Suppl 1), pages S37–S38, Barcelona, Spain, 2017.
- [3] Javier Esteban, Matthias Grimm, Mathias Unberath, Guillaume Zahnd, and Nassir Navab. Towards Fully Automatic X-Ray to CT Registration. In Dinggang Shen, Tianming Liu, Terry M. Peters, Lawrence H. Staib, Caroline Essert, Sean Zhou, Pew-Thian Yap, and Ali Khan, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, volume 11769, pages 631–639. Springer International Publishing, Cham, 2019. Series Title: Lecture Notes in Computer Science.
- [4] Mathias Unberath, Jan-Nico Zaech, Sing Chun Lee, Bastian Bier, Javad Fotouhi, Mehran Armand, and Nassir Navab. DeepDRR – A catalyst for machine learning in fluoroscopy-guided procedures. In Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-Lopez, and Gabor Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, volume 11073, pages 98–106. Springer International Publishing, Cham, 2018.

- [5] Guoyan Zheng. Statistical shape model-based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard AP x-ray radiograph: SSM-based surface model reconstruction from a single image. *Medical Physics*, 37(4):1424–1439, March 2010.