ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

AI-based Collimation Optimization for X-Ray Imaging using Time-of-Flight Cameras

Dominik Mairhöfer[1*] Manuel Laufer[1*] Lennart Berkel[2]
Arpad Bischof [2,3] Erhardt Barth[1] Jörg Barkhausen[2]
Thomas Martinetz[1 †]

1- University of Lübeck - Institute for Neuro- and Bioinformatics
Ratzeburger Allee 160, 23562 Lübeck - Germany

2- University Medical Center Schleswig-Holstein
Ratzeburger Allee 160, 23562 Lübeck - Germany

3- IMAGE Information Systems Europe GmbH
Lange Str. 16, 18055 Rostock - Germany

Abstract. Collimation during radiography, which is the process of defining the area to be radiated, is a crucial factor for the protection of the patient and for the diagnostic quality of a radiograph. Moreover, incorrect collimation is one of the main causes for a retake and the associated costs. In this paper we propose a novel collimation optimization approach using Time-of-Flight cameras and deep Neural Networks trained end-to-end to increase the diagnostic quality of a radiograph. For this we acquired a new dataset in a clinical environment consisting of depth images of the lower leg and the abdomen. Using this dataset we are able to segment depth images for the optimal collimation with an average IoU of 83%.

1 Introduction

Radiographs are still one of the most frequently used imaging modalities in clinical practice. The most essential aspect for the patient's health in such a procedure is to reduce the radiation exposure to a minimum. This goal is the basis for the *ALARA* (As Low As Reasonably Achievable) principle, regarding the effective radiation dose. The collimation of the X-ray beam is an essential tool for following the requirements of this principle as it directly influences the effective radiation dose. While a too large collimation area of the X-ray beam would expose the patient unnecessarily to radiation, a too small collimation area would result in diagnostic relevant anatomies not being completely visible in the radiograph, leading to a poor diagnostic quality and thus a retake. Therefore, optimal collimation is a crucial factor for the protection of the patient and diagnostic quality of the radiograph.

In clinical practice collimation is done by radiographers. However, time pressure, inexperience, patients with irregular body shapes or patient movement after collimation are reasons why collimation is often not ideal. Various studies

*Contributed equally. The order of author names was randomly determined.

†We thank Celina Schubbe for her help and support in collecting the dataset. This work was funded by the Bundesministerium für Wirtschaft und Klimaschutz (BMWK) through the KI-SIGS project.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

show that in many cases the radiation dose is unnecessarily high due to incorrect collimation [1, 2, 3, 4]. The authors of [3] show that, after incorrect patient positioning, a too small collimation area is the main reason for the need of a retake, since diagnostic relevant anatomies may not appear in the radiograph.

In this paper, we propose an approach that uses Time-of-Flight (ToF) cameras to capture depth images of the patient and train deep Neural Networks to directly predict the optimal collimation. To evaluate our approach, we acquired a total of 1020 depth images of the anterior posterior (a.p.) view of the abdomen and the upper ankle joint during a recreation of the clinical process of radiographing these anatomies. Due to regulatory and legal challenges, it was not readily possible to capture the corresponding radiograph in addition to the depth images, nor was it possible to mount the cameras in the clinical process, to obtain both. These radiographs could have been used to obtain a label for correct collimation. Instead, the correct collimation was directly labeled based on the depth information in close exchange with experienced radiographers. The application of this approach in clinical practice could support the collimation decisions of radiographers, reducing the radiation dose for the patient and improving the overall workflow and diagnostic quality.

2 Related Work

While there is already research regarding the determination of the optimal collimation on radiographs [5, 6], at this point the possibly bad radiograph has already been made. In [7] an approach is presented to learn the optimal collimation based on depth images. Since the authors' dataset consists of 177 depth images and radiographs of the chest in a.p. and lateral view, they were able to label on the radiograph. However, for this, they used landmarks as handpicked features instead of directly labelling the collimation. This leads to dividing the problem into a landmark detection, implemented as boosted tree classifiers, and a multivariate regression problem, for the actual collimation area prediction. With such an approach an end-to-end training becomes impossible. In contrast, we present an end-to-end trainable deep learning approach that is able to leverage depth information from multiple cameras and evaluate its functionality for two anatomies. Furthermore, by labelling the actual collimation area, our approach requires less label effort. To our knowledge, there are no other studies that automatically optimize the collimation before exposure to radiation.

3 Dataset

Since there was no public dataset in which subjects were captured under an X-ray device using depth cameras, a novel dataset was acquired for this work. For this purpose, two Microsoft Azure Kinect cameras were attached directly to the X-ray tube in an X-ray room used in everyday clinical practice. The two cameras were mounted on both ends of the X-ray tube to avoid possible occlusions of the

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

patient. Using this setup, four subjects (three male, one female) were captured in various typical poses of the upper ankle or abdomen in each case in the a.p. view on the X-ray table. A notable feature of our dataset is the inclusion of positions in which subjects had not yet moved to the correct position for radiography. After sorting out unusable images a total of 510 depth image pairs (216 upper ankle, 294 abdomen) were selected for the dataset. To aggregate information from both cameras, these pairs of depth images were converted into one registered point cloud. Optimal collimation was then labeled on these merged point clouds, in close exchange with experienced radiographers. Based on these data, 10 augmented point clouds and 10 augmented depth images together with the augmented label were created from each merged point cloud for training. Including the augmented data the dataset contains 2376 labeled point clouds and depth images of the upper ankle a.p. view and 3289 of the abdomen a.p. view. Figure 1a visualizes the described dataset creation process.

3.1 Point Cloud Generation

Since the original point clouds contain a lot of unnecessary information, all points further than 1.5m from both ToF cameras were removed. Before removing the points, the positions of the camera are translated randomly by up to 25cm, which corresponds to a translation of the X-ray device before the exposure. From the remaining points, 75,000 points were then sampled using furthest-point-sampling to create a uniformly distributed point cloud.

3.2 Depth Image Generation

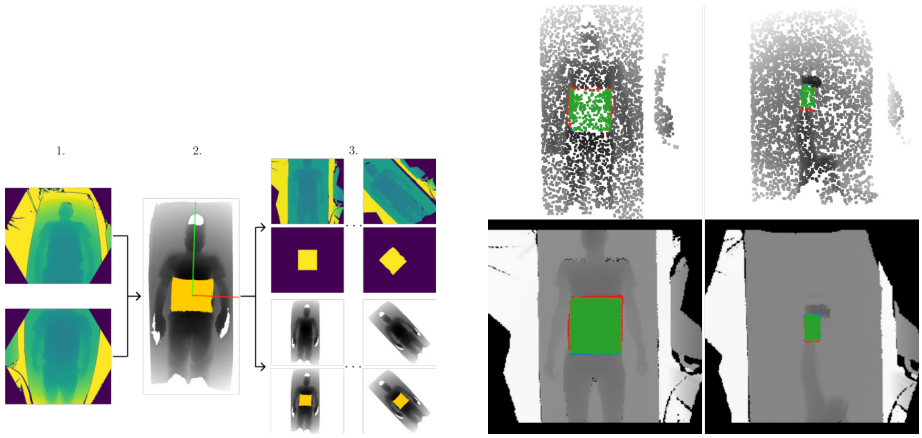
To obtain a single depth image containing information from both cameras and avoiding a multiview-problem, a synthetic depth image was rendered from the merged point cloud. For this the origin of the central beam was chosen as the camera position, since the optimal collimation can be determined from this perspective. In order to create the synthetic depth image, the labeled optimal position of the X-ray device is randomly translated by up to 25cm and then the depth image is rendered from the new position. While the Azure Kinect captures depth images at a resolution of 640 x 576, we rendered the images at half that resolution, which helps to reduce invalid points in the rendered images.

4 Experiments and Training

We implemented and tested different architectures suitable for our task. In all cases the prediction was modeled as a segmentation task using the pixel-wise/point-wise labels. For the point clouds as input data we used the PVCNN++ [8] with a PointNet++ [9] as backbone. For the depth images as input data, we used a UNet++ [10] as a state-of-the-art segmentation model. The training process was identical for all models of one input modality, but differed slightly across image and point cloud based networks. Both networks were trained independently for both anatomies using the Cross Entropy Loss as loss function

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

for a point-wise/pixel-wise classification in the classes *collimation* and *no collimation*. For the point cloud models the Adam [11] optimizer with an initial learn rate of $1 \cdot 10^{-3}$ and a batch size of 16 was used for training for 400,000 iterations. The learning rate was decreased 3 times by a factor of 10, at 250,000, 300,000 and 350,000 iterations. For the depth images, we used the stochastic gradient descent with $1 \cdot 10^{-3}$ as learning rate, a momentum of 0.9 and a weight decay of $1 \cdot 10^{-6}$. The network was trained for 350,000 iterations using a batch size of 32 with the learning rate decreased 3 times by a factor of 10 at 200,000, 275,000 and 325,000 iterations. From the four subjects of the dataset, one was randomly chosen before the training process and specified as test set and one as validation set. For the final training process we trained on all three subjects and tested on the initially selected test subject 5 times with random initializations.



(a) (b)

Fig. 1: (*a*) Dataset creation: 1. acquire depth images, 2. merge them into a registered point cloud and label them and 3. create augmented depth images and point clouds. (*b*) Collimation prediction examples. True positive pixel/points are colored in green, false positives in red and false negatives in blue.

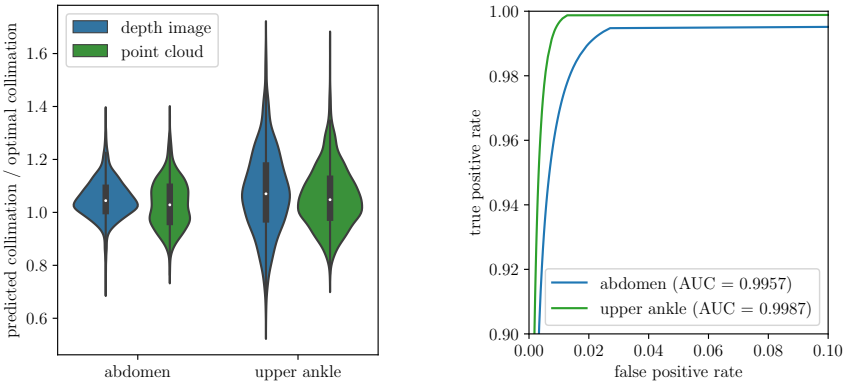
5 Results and Discussion

We modeled the problem as segmentation task and measured the accuracy of the prediction using the Intersection over Union (IoU) metric. In order to enable a more precise evaluation with regard to a too small/large collimation, the specificity and sensitivity are also measured in addition to the pixel-wise accuracy. A sensitivity close to 100% indicates that the area of labeled collimation is almost completely correctly covered. A high specificity, on the other hand, indicates that not much area was exposed unnecessarily. Since it is preferable to choose a collimation slightly larger than absolutely necessary than to have to repeat the acquisition due to a collimation being too small, a higher sensitivity

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Table 1: Results for both anatomies and both input modalities respectively. It is visible that high IoU scores can be achieved for both anatomies.

Metric	upper ankle		abdomen	
	point cloud	depth image	point cloud	depth image
IoU [%]	81.63±0.70	78.76±0.31	83.81±0.61	88.47±0.48
Accuracy [%]	99.39±0.03	99.68±0.01	98.12±0.08	98.85±0.05
Sensitivity [%]	92.37±0.54	91.21±0.41	92.50±0.26	96.11±0.38
Specificity [%]	99.61±0.03	99.79±0.01	98.80±0.08	99.15±0.05
FN/FP Ratio	0.60±0.07	0.57±0.05	0.75±0.05	0.48±0.06



(a) (b)

Fig. 2: (a) shows the distribution of the ratio of the predicted collimation to the optimal collimation. The majority of the results are slightly above the ideal value of 1 with the mean value of 104.97% and 107.69% for abdomen and upper ankle on depth images as well as 103.22% and 106.07% on point clouds. (b) shows the cropped ROC curve for depth images and different discrimination thresholds.

is more important than a higher specificity. Table 1 shows that for both point clouds as well as depth images high IoU scores are achieved for both anatomies. As it can be seen in Figure 2a the collimation prediction is on average only 4.09% too large for abdomen and 6.88% for the upper ankle regarding the optimal collimation. While not directly compareable due to the different target anatomies, this is significantly less than the 26% from [4]. The results also show that the absolute number of false negatives is already significantly below the false positives (see FN/FP Ratio in Table 1). In Figure 2b, the relationship between specificity and sensitivity with respect to different thresholds can be seen for depth images. The mean AUC value of 99.72% supports that a correct classification can be achieved with a high probability. In Figure 1b random

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

chosen example collimation predictions on point clouds and depth images from the same pose are shown. Our results show that for both anatomies, precise segmentation of the optimal collimation can be achieved with an IoU of up to 88% being on average only 5.49% larger than the optimal collimation. It is even possible to train our approach without any adjustment if a radiograph is available for more accurate labeling. In addition, the presented approach can also be transferred to any other body parts. This could furthermore reduce the patient's exposure to radiation and improve the overall workflow in radiography.

References

[1] Zahra Farzanegan, Marziyeh Tahmasbi, Mohsen Cheki, Fatameh Yousefvand, and Mohammad Rajabi. Evaluating the principles of radiation protection in diagnostic radiologic examinations: Collimation, exposure factors and use of protective equipment for the patients and their companions. *Journal of Medical Radiation Sciences*, 67(2):119–127, 2020.

[2] Vahid Karami, Mansour Zabihzadeh, Abdolreza Gilavand, and Nasim Shams. Survey of the Use of X-ray Beam Collimator and Shielding Tools during Infant Chest Radiography. *International Journal of Pediatrics*, 4(4):1637–1642, April 2016.

[3] Kevin J. Little, Ingrid Reiser, Lili Liu, Tiffany Kinsey, Adrian A. Sánchez, Kateland Haas, Florence Mallory, Carmen Froman, and Zheng Feng Lu. Unified Database for Rejected Image Analysis Across Multiple Vendors in Radiography. *Journal of the American College of Radiology*, 14(2):208–216, February 2017.

[4] V. Karami and M. Zabihzadeh. Beam Collimation during Lumbar Spine Radiography: A Retrospective Study. *Journal of Biomedical Physics & Engineering*, 7(2):101–106, June 2017.

[5] Pedersen Alex Elgaard, Lysdahlgaard Simon, and Martin Weber Kusk. Collimation border with U-net segmentation on chest radiographs compared to radiologists. *Journal of Medical Imaging and Radiation Sciences*, 53(4, Supplement 1):S43–S44, December 2022.

[6] Jens von Berg, Sven Krönke, André Gooßen, Daniel Bystrov, Matthias Brück, Tim Harder, Nataly Wieberneit, and Stewart Young. Robust chest x-ray quality assessment using convolutional neural networks and atlas regularization. In *Medical Imaging 2020: Image Processing*, volume 11313, pages 391–398. SPIE, March 2020.

[7] Julien Sénégas, Axel Saalbach, Martin Bergtholdt, Sascha Jockel, Detlef Mentrup, and Roman Fischbach. Evaluation of Collimation Prediction Based on Depth Images and Automated Landmark Detection for Routine Clinical Chest X-Ray Exams. In Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Lecture Notes in Computer Science, pages 571–579, Cham, 2018. Springer International Publishing.

[8] Zhijian Liu, Haotian Tang, Yujun Lin, and Song Han. Point-Voxel CNN for Efficient 3D Deep Learning. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[9] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[10] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. UNet++: A Nested U-Net Architecture for Medical Image Segmentation, July 2018.

[11] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017.