Robust Hough and Spatial-To-Angular Transform Based Rotation Estimation for Orthopedic X-Ray Images

Magdalena Bachmaier[1,2], Maximilian Rohleder[1,2], Benedict Swartman[3], Maxim Privalov[3], Andreas Maier[1], and Holger Kunze[1,2(✉)]

[1] Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
`Holger.hk.kunze@siemens-healthineers.com`
[2] Siemens Healthcare GmbH, Forchheim, Germany
[3] Department for Trauma and Orthopaedic Surgery, BG Trauma Center Ludwigshafen, Ludwigshafen, Germany

Abstract. Standardized image rotation is essential to improve reading performance in interventional X-ray imaging. To minimize user interaction and streamline the 2D imaging workflow, we present a new automated image rotation method. Image rotation can follow two steps: First, an anatomy specific centerline image is predicted which depicts the desired anatomical axis to be aligned vertically after rotation. In a second step, the necessary rotation angle is calculated from the orientation of the predicted line image. We propose an end-to-end trainable model with the Hough transform (HT) and a differentiable spatial-to-angular transform (DSAT) embedded as known operators. This model allows to robustly regress a rotation angle while maintaining an explainable inner structure and allows to be trained with both a centerline segmentation and angle regression loss. The proposed method is compared to a Hu moments-based method on anterior-posterior X-ray images of spine, knee, and wrist. For the wrist images, the HT based method reduces the mean absolute angular error (MAE) from $9.28°$ using the Hu moments-based method to $3.54°$. Similar results for the spinal and knee images can be reported. Furthermore, a large improvement of the 90th percentile of absolute angular error by a factor of 3 indicates a better robustness and reduction of outliers for the proposed method.

Keywords: Machine Learning · Image Rotation · Hough Transform

1 Introduction

Intra-operative X-ray imaging supports the assessment of fracture reduction and implant positioning, which significantly increases surgical precision [7] and the overall outcome [1]. Therefore, standardized image rotation is essential to improve reading performance and interpretation. While in diagnostic imaging, this can be reached by careful alignment of the imaging system with the patient, in intra-operative setups, this alignment is hard to achieve due to mechanical

constraints of the widely used C-arm devices and limited space in the operating room. The rotation of the images must then be corrected in a post-processing step [11]. Therefore, a common task in the radiography workflow is to manually rotate digital X-ray images to a preferred orientation suitable for diagnostic reading. However, during a surgery, user interaction of the surgeon with the system must be minimized due to sterility considerations. Thus, even improved user interaction with the system does not improve the situation for the surgeon much. Supporting staff is often not well trained in operating the vendor-specific systems and changes job positions frequently. Therefore, the speed and quality of the image alignment depends on the experience of the operator and leaves room for mistakes. An automatic X-ray alignment system can help ensure proper alignment resulting in improved reading performance and interpretation also [2]. Previous attempts [3,14,17,18] to automate this rotation typically only allow the identification of a limited set of orientations. In addition, they are often specially designed for certain examinations. More recent publications [4] work on the regression of the correct angular offset. Conventional vision algorithms based on feature extraction with subsequent feature registration to an atlas, typically fail to automate the image rotation of intra-operative images due to the enormous variety of images with different collimation settings, fracture types, and instruments/casts that obscure many image features. Additionally, they hardly generalize to a large number of body regions due to the nature of the hand crafted image features. In this paper, we extend the indirect regression approach of Kunze et al. [11]. It bases on the insight that for many anatomical structures a line can be defined that strongly corresponds with the upright position of the X-ray image. After this line has been determined as a heatmap by a Convolutional neural network (CNN) according to the idea of Kordon et al. [8] this heatmap can be post-processed to calculate the orientation of the line. Kunze et al. [11] have demonstrated that this indirect approach is more robust than the direct regression approaches for angle regression [2,6] based on ResNet or PoseNet. In applications with different crops of anatomic regions, Kordon et al. [10] have shown that the quality of the heatmap can be improved if a companion objective function is integrated into the training that optimizes the regression target directly. Therefore, the derivable calculation of the position and direction based on the Hu moments (HM) was added to the cost function of the segmentation task, enabling end-to-end gradient flow. This approach can be also adapted for the angle regression. However, the HM are an indirect measure leaving much room for the algorithm to shape the heatmap in an uncontrolled manner. Therefore, to limit the space of solutions, we propose to substitue the HM based angle regression with a Hough transform (HT) based one. By doing so, the HM based cost term is replaced by a term, that favors straight, narrow lines. Thereto, we present an algorithm to implement the computation of the line angle based on the HT differentiable to keep end-to-end training based on angular values.

The main contribution of the paper can be summarized as follows:

– proposition of a derivable calculation of the image rotation angle derived from the HT heatmap in the manner of a differentiable spatial to numerical transform (DSNT) [16]
– ablation study which examines the properties of the proposed method and shows that the proposed method increases the robustness of the image rotation compared to state-of-the-art indicated by the 90th percentile of the absolute angular error.
– verification of the approach on a the body regions spine, wrist, and knee.

2   Materials and Methods

2.1   Determination of Rotation Angle

In Kunze et al. [11], two classes of algorithms are compared with which the rotation angle of images can be derived: The direct regression method, uses an encoder structure determining festures of the image followed by a regression network converting them into an angular valued, represented by its sine and cosine value for stability reasons. The indirect method computes a heatmap of a line which represents the upwards direction of the image. Analysing the direction of the line returns the rotation angle of the image. While the first approach consisting of a modified version of ResNet-34 with the fully connected (FC) layer being substituted by two FC layers serves as baseline algorithm [11], in the following the angle determination of the indirect approch is revised.

Angle Regression Exploiting the Hough Transform. As pointed out in the introduction, the training with a cost function based on the HM does not enforce a narrow destinct line, but any structure with a defined main axis contributes to a low cost. A cost function which enforces this behaviour would potentially more distinctive heatmap.

A common approach to detect lines in pattern recognition applications is the HT: Given a line $\mathbb{L}_{\rho,\theta}$, with ρ the angle between line and the x-axis and θ its distance from the origin, its corresponding pixels $\{x_{\rho,\theta,i}, y_{\rho,\theta,i}\}$ vote in the Hough space for the closest bin (ρ, θ). In contrast to the HM based direction calculation, the maximum of the HT is only obtained for a straight line of given extent. The extent can be controlled by the bin width of the HT. This feature of the HT enables the training of a narrow structure with a defined width punishing wide structures. However, the HT does not return an angle itself but returns a new 2-dimensional representation of the heatmap, with the line to be sought represented as its maximum. When this representation shall be integrated in an end-to-end trainable network, a differentiable calculation of the position of this maximum is needed. Thereto, we follow the idea of the DSNT layer introduced by Nibali et al. [16] that calculates the channel-wise centroid of the input. For the current scenario, merely the centroid $\bar{\theta}$ of the θ coordinate indicating the direction of the line is of interest. The $\bar{\theta}$ can be written as

$$\bar{\theta} = \frac{\sum_\theta \sum_\rho \theta g(\theta, \rho)}{\sum_\theta \sum_\rho g(\theta, \rho)} = \frac{\sum_\theta \theta \sum_\rho g(\theta, \rho)}{\sum_\theta \sum_\rho g(\theta, \rho)}. \tag{1}$$

Equation (1) involves a summation over ρ coordinate. Since the zeroth-order Helgason-Ludwig condition of the Radon transform [13] states that the sum over each projection is constant, also the Hough histogram yields a constant first-order moment. This renders the direct implementation incapable for the angle computation.

To overcome this shortcoming, A non-linear function needs to be applied to the Hough histogram to assign high pixel values a greater weight than small ones. Keeping pixel values between 0 and 1 while weighting maximum values and suppressing low values, a steep sigmoid function is employed, which maps input values to a range from 0 to 1. At best, only values from roughly 0.8 to 1 should be used. The steeper the sigmoid function is, the more it approximates the step function, which ideally assigns a weight of 1 to values of 0.8 and higher.

For the end-to-end training, the same issue holds as for the direct regression: the angular value θ has a discontinuity for $\theta = -90°$ and $\theta = 90°$. Both θ values represent the same line. To pass this boundy problem, the cosine- and sine-value are computed instead of the angular value itself. This can be incorporated in Eq. (1) by using a sine- and cosine-weighting instead of linear one. So the total angular value calculation given the HT can be written as

$$(cos(\hat{\theta}), sin(\hat{\theta})) = \left(\frac{\sum_\theta cos\theta \sum_\rho \hat{g}(\theta, \rho)}{\sum_\theta \sum_\rho \hat{g}(\theta, \rho)}, \frac{\sum_\theta sin\theta \sum_\rho \hat{g}(\theta, \rho)}{\sum_\theta \sum_\rho \hat{g}(\theta, \rho)} \right) \qquad (2)$$

with

$$\hat{g}(\theta, \rho) = sigmoid\left(\alpha \cdot g(\theta, \rho) - \beta\right), \qquad (3)$$

where $\alpha = 20$ and $\beta = 17$ were determined heuristically. $\hat{\theta}$ can be calculated by the atan2 function from its sine and cosine value. We call the layer, that implements Eq. (2) in combination with the atan2 method differential spatial to angular transform (DSAT).

Neither the HT layer parameters nor the DSAT layer have parameters, that need to be trained: The HT can be written as a multiplication of a vector representing the values of the image with a sparse matrix mapping the input image into a $[N_\rho \times N_\theta]$ Hough histogram Y, where N_ρ and N_θ are the numbers of discrete offsets and angles [12]. For the DSAT layer, Eq. 1 can be implemented. So both parts of the algorithms can be treated as known operators [15].

2.2   Data

For the experiments, three different datasets containing the spine, the knee and the wrist, were used. The spinal dataset consists of 958 images derived from 148 patients, where 289 images represent the cervical, 150 the lumbar, and 519 the thoracic region. The wrist dataset contains 257 and the knee dataset 113 images. The 16-bit, gray-scale X-ray images were selected randomly retrospectively from anonymized databases which were acquired using the Cios Spin mobile C-arm system of Siemens Healthcare GmbH during orthopedic surgeries. The images are depicted in the AP view and may contain screws, plates, and other surgical tools.
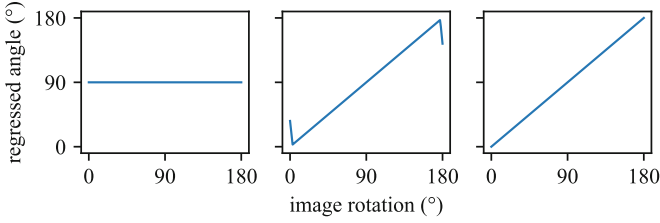
Fig. 1. Regression results for different calculations of the image rotation based on the Hough transform (HT). Left: Linear central moment on HT, Middle: Linear central moment on sigmoid activated HT, Right: Sine-Cosine weighted central moment on sigmoid activated HT

In these images, the centerline was marked using the labelme tool [19] by a trained medical engineer. During annotation, special care has been taken to ensure that the rotation angle of the image can be calculated as the angle between the line and the y-axis by consistently drawing the orientation of the line.

2.3   Training Protocol, Experiments, and Evaluation

To identify the effect of the post processing method on the angle regression error, an ablation study was set up. Following Kunze et al. [11], for the indirect (segmentation-based) method, a D-LinkNet [20] was chosen as underlying network architecture. The output of the D-LinkNet, which is the heatmap of the line, is fed into the HT layer and its result is processed by the DSAT layer. As loss function for the training, the weighted sum of the segmentation loss and the regression loss was chosen. The segmentation loss consists of the sum of Binary Cross Entropy and Mean Squared Error(MSE). For the regression loss, the MSE of the regressed rotation's sine- and cosine-values was selected. Thus, the total loss can be defined as

$$\mathcal{L} = \lambda \cdot \mathcal{L}_{seg} + (1 - \lambda) \cdot \mathcal{L}_{reg}, \tag{4}$$

where $\lambda \in \mathbb{R}$ is a multiplicative weighting term. Using this we examined the following 3 cases:

1. For $\lambda = 1$, the training corresponds to one without the regression loss. In this case, the performance of the HM based and the HT based angle regression methods can be compared on the heatmap.
2. With a constant weighting term $\lambda = 0.5$, the regression loss is considered along with the segmentation loss during the training.
3. When decreasing *lambda* over the epochs from 1 to 0, we obtain a training which is kick-started using the segmentation loss only. But in the end, only the regression loss is used for the training.

For the parameter selection of the HT layer, a grid search was performed on synthetic line heatmaps, resulting in discretization values $d\theta = 0.4°$ and $d\rho = 18px$.

For comparison reasons, the direct regression method based on a ResNet-34 was trained using the MSE based on the sine- and cosine-vales as cost function.

During all trainings, online augmentation for rotation ($\alpha \in [-45°, 45°], p = 0.5$), inverting ($p = 0.5$), shifting ($s \in [-40px, 40px], p = 0.5$) and contrast enhancement was applied. After the augmentation, the images were scaled to the dimension [H:256 × W:256] px, using zero-padding the image to preserve the ratio. Thereafter, sample-wise data normalization using z-scoring was used and combined with batch normalization [9]. The heatmaps for segmentation training were created by placing a Gaussian function of width $7px$ in a neighborhood of $21px$ on the line connecting the augmented start and end point of the center line. The Gaussian was scaled such that the center point obtains a value of 1 [11].

For the evaluation, different matrices based on the angular error were used. The mean error (ME) reveals a directional offset of the algorithm. Further on, the mean absolute error (MSE) and the standard deviation of the absolute error (Std.) are reported. Practically more relevant are the percentiles $\mathbf{P}_{50}$, $\mathbf{P}_{90}$ and $\mathbf{P}_{95}$ based on the absolute error as they reveal to what extent manual rotation corrections are required after the automatic correction was performed. Based on the MAE a Wilcoxon signed-rank test was performed to verify the significance of the observed differences.

To evaluate the performance of the presented methods, a 5-fold cross-validation scheme is employed. The mean performance score across all folds served as final result of the evaluated algorithm. Training was performed using the Adam optimizer for 500 epochs with a learning rate of $l = 0.001$ divided by two every 50 epochs. The batch size was selected to be 8. The weights of the segmentation network as well as of the direct regression model were initialized by He's method [5]. Implementation was done in PyTorch v1.11 (Python v3.9.12, CUDA v11.0), training was performed on a Windows 10 system with 64 GB RAM and 24 GB NVIDIA Titan RTX. Reproducibility was confirmed by repeated trainings.

3   Results

The results of the experiments are shown in Table 1. A comparison of the segmentation results is made in Fig. 1.

The results show for all three anatomies, that the HT based methods are superior to the HM based method. So already as post-processing method, it significantly outperforms the HM based method (spine: $p = 1.0 \times 10^{-95}$, wrist: $p = 5.6 \times 10^{-21}$, knee: $p = 6.7 \times 10^{-8}$). Further improvements of the angle regression results can be realized by incorporating the regression loss in the training (spine: $p = 9.5 \times 10^{-91}$, wrist: $p = 5.7 \times 10^{-21}$, knee: $p = 4.5 \times 10^{-9}$). Best results except for wrist are obtained when the final training is performed on the regression loss only (spine: $p = 1.3 \times 10^{-2}$, wrist: $p = 7.7 \times 10^{-1}$, knee: $p = 2.1 \times 10^{-2}$). For the wrist adding the regression loss to the overall loss does not improve the training. The comparable $\mathbf{P}_{50}$ value for most of the anatomies shows, that HT based and HM based angle regression both work well for standard images.

Table 1. Angular error in degree for the different network topologies and angle calculation methods. The direct regression uses the values obtained by the fully connected layers, while the Hu moments and Hough transform-based methods analyze the detected line at the heatmap output and are integrated as post-processing step as well as additional layers.

Anatomy	Model	ME	MAE	Std.	P_{50}	P_{90}	P_{95}
spine	ResNet-34	0.07	6.95	6.32	5.37	14.54	18.95
	Seg loss, HM post-proc	−0.17	2.14	3.23	1.19	4.87	7.22
	Seg loss, HT post-proc	0.13	1.80	2.36	1.08	3.92	5.69
	Seg loss & HM loss	−0.01	2.05	2.91	1.28	4.33	6.08
	Seg loss & HT loss	0.10	1.79	2.50	1.13	3.61	5.21
	adapt. Seg loss & HM loss	−0.08	1.98	3.50	1.24	4.31	5.97
	adapt. Seg loss & HT loss	0.15	1.61	2.19	1.05	3.40	4.72
wrist	ResNet-34	0.32	7.31	5.12	6.25	14.43	17.28
	Seg loss, HM post-proc	2.26	9.28	16.02	3.13	26.51	43.80
	Seg loss, HT post-proc	0.39	3.73	7.72	1.97	7.63	10.01
	Seg loss & HM loss	4.40	11.24	16.93	3.80	31.13	48.33
	Seg loss & HT loss	0.23	3.54	6.60	2.07	6.75	11.46
	adapt. Seg loss & HM loss	3.24	9.93	16.75	3.66	24.81	49.09
	adapt. Seg loss & HT loss	0.63	4.80	11.67	2.05	6.97	11.47
knee	ResNet-34	0.22	5.63	4.97	4.01	12.48	14.27
	Seg loss, HM post-proc	4.99	8.77	9.18	5.71	19.53	29.97
	Seg loss, HT post-proc	0.15	4.43	11.83	1.68	7.16	10.73
	Seg loss & HM loss	6.00	10.00	15.27	4.77	24.57	40.45
	Seg loss & HT loss	0.44	3.71	4.39	2.23	7.76	11.13
	adapt. Seg loss & HM loss	0.49	8.31	13.14	3.68	16.72	32.62
	adapt. Seg loss & HT loss	1.05	3.48	8.07	1.71	6.29	8.34

For all anatomies, an improvement in the 90th percentile of the absolute angular error can be observed. While for the spine the improvement is about 17%, for the wrist and knee, this value an reduction by a factor of 4.6 and 3.1, respectively, can be observed.

A case-by-case analysis shows that on images for which the centerline is well determined, the HT analysis returns roughly the same result as the HM analysis. However, since the dataset also contains many challenging cases, e.g., images with cropped anatomical structures of interest or images containing metal implants, the centerline cannot always be well determined by the segmentation network. Then occasionally, the heatmap generates short and fragmented estimates of the centerline. For these images, the error of the regressed angle by the HT based method typically is thinner compared to that by the HM computed one (Fig. 2).
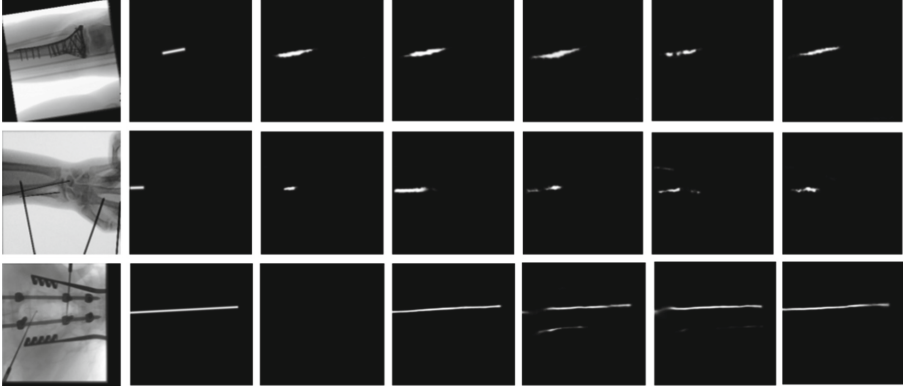
Fig. 2. Comparison of the segmentation results for upper ankle, wrist and spine (top to bottom) trained with segmentation loss only, Hu moments, Hough loss, Hu moments with loss adaption, and Hough loss with loss adaption (column 3–6) with corresponding input images (1st column) and ground truth heatmaps (2nd column).

4   Discussion

We introduced a new HT based method to calculate the rotation of straight heatmaps. For the implementation, special care needs to be taken to cope with constant first-order moments of the HT and boundary problems at $\theta = \pm 90°$.

Doing so, the proposed method has similar performance to HM for X-ray images of good quality. In more difficult cases, when the segmentation algorithm returns multiple and only short segments, like for the wrist or knee, the proposed algorithm is more robust than the HM based one.

The largest benefit of the HT based image rotation calculation can be achieved with its embedding into the cost function of a segmentation network. Then it helps to improve the accuracy compared to a segmentation-only approach. Especially outliers can be reduced resulting in a lower MAE and standard deviation of the error. The presented results confirm the work of Kordon et al. [10] that direct optimization of the target measure in addition to a proxy objective is beneficial to the underlying heatmap-generating problem. That is especially true for body regions like the wrist or knee. For these regions, the lengths of the structures and the annotated lines defining the orientation of the image vary. Then, the algorithm can benefit from a more generic description. For the spine, a line from the top to the bottom of the image can be drawn. Therefore, the regression-related term provides less additional information during the training.

The inspection of the generated heatmaps reveals that the HT based approach can enforce more pronounced, thinner heatmaps compared to the HM - as desired. This observation can be explained by the property of the HT based cost function: it penalizes structures that are not oriented along the desired direction. Only thin straight heatmaps along the direction have a positive effect on

the loss term. Structures that are parallel, like the centerlines of the radius and ulna for the wrist images, are penalized. In contrast to that, the HM derived cost term does not punish broad heatmaps as long as the largest main axis points toward the correct direction. Thus, this regression term does not enforce one single, thin line. Also, clusters are tolerated as the results depict. That explains the reduced number of outliers for the HT bases method compared to the HM based, indicated by P_{90} and P_{95} values. The fact that the adaptation of the parameter λ during the training has only a small effect can be explained by the regression loss being larger compared to the segmentation loss by magnitudes as soon as a certain segmentation level is reached. So the segmentation loss only at the beginning of the training has a distinct influence on the training. Thus, the training process is adapting the influence of the segmentation loss by itself.

Finally, the results confirm the finding of Kunze et al. [11] that the segmentation-based rotation regression is more robust compared to the direct regression by a ResNet-34, at the cost of the uncertainty of an up-down flip. For the wrist, slightly worse results are reported compared to [11]. That can be attributed to the dataset for this study containing only intra-operative X-ray images which are acquired not as standardized as the diagnostic images used in [11].

Data Use Declaration: The data was obtained retrospectively from anonymized databases and not generated intentionally for the study. The acquisition of data from patients had a medical indication. Informed consent was not required.

References

1. Atesok, K.: The use of intraoperative three-dimensional imaging (ISO-C-3D) in fixation of intraarticular fractures. Injury **38**(10), 1163–1169 (2007). https://doi.org/10.1016/j.injury.2007.06.014
2. Baltruschat, I.M., Saalbach, A., Heinrich, M.P., Nickisch, H., Jockel, S.: Orientation regression in hand radiographs: a transfer learning approach. In: Proceedings of SPIE Medical Imaging, vol. 10574, pp. 473–480 (2018). https://doi.org/10.1117/12.2291620
3. Boone, J.M., Seshagiri, S., Steiner, R.M.: Recognition of chest radiograph orientation for picture archiving and communications systems display using neural networks. J. Digit. Imaging **5**(3), 190–193 (1992). https://doi.org/10.1007/BF03167769
4. Fonseca, A., Vieira, G.S., Felix, J., Freire Sobrinho, P., Silva, A.V.P., Soares, F.: Automatic orientation identification of pediatric chest x-rays. In: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1449–1454 (2020). https://doi.org/10.1109/COMPSAC48688.2020.00-51
5. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1026–1034 (2015). https://doi.org/10.1109/ICCV.2015.123
6. Kausch, L., et al.: Toward automatic c-arm positioning for standard projections in orthopedic surgery. Int. J. Comput. Assist. Radiol. Surg. **15**(7), 1095–1105 (2020). https://doi.org/10.1007/s11548-020-02204-0

7. Keil, H., Beisemann, N., Swartman, B., Vetter, S.Y., Grützner, P.A., Franke, J.: Intra-operative imaging in trauma surgery. EFORT Open Rev. **3**(10), 541–549 (2018). https://doi.org/10.1302/2058-5241.3.170074

8. 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

9. Kordon, F., et al.: Improved x-ray bone segmentation by normalization and augmentation strategies. In: Bildverarbeitung für die Medizin 2019. I, pp. 104–109. Springer, Wiesbaden (2019). https://doi.org/10.1007/978-3-658-25326-4_24

10. Kordon, F., Maier, A., Swartman, B., Privalov, M., El Barbari, J.S., Kunze, H.: Deep geometric supervision improves spatial generalization in orthopedic surgery planning. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) Medical Image Computing and Computer Assisted Intervention - MICCAI 2022. LNCS, pp. 615–625. Springer Nature Switzerland, Cham (2022). https://doi.org/10.1007/978-3-031-16449-1_59

11. Kunze, H., Kordon, F., Maier, A., Breininger, K.: Direct and indirect image rotation estimation methods of orthopedic X-ray images. In: Proceedings SPIE Medical Imaging, pp. 640–647 (2022). https://doi.org/10.1117/12.2606045

12. Lin, Y., Pintea, S., Gemert, J.: Semi-supervised lane detection with deep hough transform. pp. 1514–1518 (2021). https://doi.org/10.1109/ICIP42928.2021.9506299

13. Ludwig, D.: The radon transform on Euclidean space. Commun. Pure Appl. Math. **19**(1), 49–81 (1966). https://doi.org/10.1002/cpa.3160190105

14. Luo, H., Luo, J.: Robust online orientation correction for radiographs in PACs environments. IEEE Trans. Med. Imaging **25**(10), 1370–1379 (2006). https://doi.org/10.1109/tmi.2006.880677

15. Maier, A.K., et al.: Learning with known operators reduces maximum error bounds. Nat. Mach. Intell. **1**(8), 373–380 (2019). https://doi.org/10.1038/s42256-019-0077-5

16. Nibali, A., He, Z., Morgan, S., Prendergast, L.: Numerical coordinate regression with convolutional neural networks. arXiv preprint arXiv:1801.07372 (2018). https://doi.org/10.48550/arXiv.1801.07372

17. Nose, H., Unno, Y., Koike, M., Shiraishi, J.: A simple method for identifying image orientation of chest radiographs by use of the center of gravity of the image. Radiol. Phys. Technol. **5**(2), 207–212 (2012). https://doi.org/10.1007/s12194-012-0155-4

18. Pietka, E., Huang, H.: Orientation correction for chest images. J. Digit. Imaging **5**(3), 185–189 (1992). https://doi.org/10.1007/BF03167768

19. Wada, K.: Labelme: image polygonal annotation with python (2016)

20. Zhou, L., Zhang, C., Wu, M.: D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 192–1924 (2018). https://doi.org/10.1109/CVPRW.2018.00034