Anatomical Landmark Detection for Initializing US and MR Image Registration

Zhijie Fang, Hervé Delingette, Nicholas Ayache

Anatomical Landmark Detection for Initializing US and MR Image Registration

Zhijie Fang, Hervé Delingette, Nicholas Ayache

Centre Inria d'Université Côte d'Azur
2004 Rte des Lucioles, 06902 Valbonne, France
zhijie.fang@inria.fr

Abstract. Targeted MR/ultrasound (US) fusion biopsy is a technology made possible by overlaying ultrasound images of the prostate with MRI sequences for the visualization and the targeting of lesions. However, US and MR image registration requires a good initial alignment based on manual anatomical landmark detection or prostate segmentation, which are time-consuming and often challenging during an intervention. We propose to explicitly and automatically detect anatomical landmarks of prostate in both modalities to achieve initial registration. Firstly, we train a deep neural network to detect three anatomical landmarks for both MR and US images. Instead of relying on heatmap regression or coordinate regression using a fully connected layer, we regress coordinates of landmarks directly by introducing a differentiable layer in U-Net. After being trained and validated on 900 and 152 cases, the proposed method predicts landmarks within a Mean Radial Error (MRE) of 5.55 ± 2.63 mm and 5.77 ± 2.67 mm in 263 test cases for US and MR images, separately. Secondly, least-squares fitting is applied to calculate a rough rigid transformation based on detected anatomical landmarks. Surface registration error (SRE) of 6.62 ± 3.97 mm and Dice score of 0.77 ± 0.11 are achieved, which are both comparable metrics in clinical setting when comparing with previous method.

Keywords: Landmark detection · Image-guided intervention · Convolutional neural network and Prostate cancer.

1 Introduction

Prostate cancer is the 2nd most commonly occurring cancer in men and the 4th most common cancer overall, with around 1.4 million new cases and 370 000 deaths worldwide in 2020 [1]. There are several tests that indicative of a potential prostate cancer, but biopsy analysis is the gold standard. The introduction of multiparametric magnetic resonance imaging (mp-MRI) now allows for imaging-based detection of prostate cancer, which may improve diagnostic accuracy for higher-risk tumors. Targeted MR/ultrasound (US) fusion biopsy is a technology made possible by overlaying ultrasound images of the prostate with MRI sequences for visualization and targeting lesions [2]. However, image fusion is a challenging and time consuming task especially for multi-modal images.

For mono-modal medical image registration, it can be solved as an optimization problem [3] by maximizing image similarity, which indicates how well image intensities correspond. However, it is difficult to engineer a similarity metric for multi-modal image registration. Last but not least, even if many researchers [4–6] worked on the US-MR image registration task, the proposed methods demand an approximate initial alignment for US and MR images, which are usually based on manual anatomical landmark detection or prostate segmentation. However, manually detecting landmarks from both modalities is time-consuming and often challenging during an intervention [7]. Therefore, computer assistance is necessary for anatomical landmarks detection in both modalities in order to achieve a good initialization for US and MR images registration.

Landmark-based Image Registration. Natarajan *et al.* [8] proposed an elastic warping of MR volume to match the US volume acquired for targeted prostate biopsy. The fusion method involves rigid alignment of the two volumes using manually selected anatomical landmarks. Heinrich *et al.* [9] proposed a landmark detection method specifically designed for lung computed tomography (CT) registration, which is not generalizable to other tasks. Grewal *et al.* [10] presented DCNN-Match, that learns to predict landmark correspondences in lower abdominal CT scans and in a self-supervised manner, which significantly improves the performance in deformable image registration. Song *et al.* [11] proposed an affine registration method for US and MR images based on four anatomical landmarks, which requires not only landmark detection network, but also segmentation network.

Landmark Detection. Detecting landmarks in images is a well-studied topic, and this problem has been explored with traditional machine learning techniques [12, 13]. Recently, deep learning methods have been proposed with fully-convolutional architecture such as U-Net [14] to compute a heatmap image as an output that highlights the location of the landmark(s). Thus, landmark localization is turned into an image-to-heatmap regression problem [11, 15, 16], where the ground truth coordinates are used to generate Gaussian blobs (of often arbitrary size) to create training data. Another coordinate regression approach is to add a fully connected layer which produces numerical coordinates [17]. An attractive property of this approach is that it is possible to backpropagate all the way from the predicted numerical coordinates to the input image. However, the weights of the fully-connected layer are highly dependent on the spatial distribution of the inputs during training, hampering the generalization ability of the overall network. Nibali *et al.* [18] proposed differentiable spatial to numerical transform (DSNT) layer for 2D human pose estimation, which is fully differentiable, and exhibits good spatial generalization.

Proposed Method. As shown in Fig. 1, we propose a pipeline to achieve initialization for US and MR image registration automatically. We use three anatomical landmarks, including the apex, the bladder neck, and the posterior median, which are displayed in Fig. 1a. A neural network is adopted to detect three anatomical landmarks in each US and MR image, separately. Least-squares

fitting [19] is applied to calculate a rough rigid transformation based on the detected landmarks from both modalities.

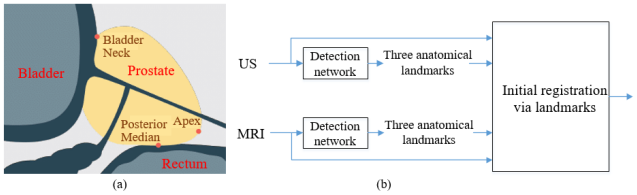

Fig. 1. Overview of the proposed pipeline for initializing US and MR image registration. (a) Three prostate anatomical landmarks: the apex, the bladder neck, the posterior median. (b) Workflow for detecting anatomical landmarks in both modalities and for computing a rough rigid transformation using least-squares fitting.

Contribution. In summary, our work to the state of the art in the following aspects:

1. The proposed pipeline can detect three prostate anatomical landmarks of both US and MR images automatically, and least-squares fitting is applied to calculate a rough rigid transformation based on detected anatomical landmarks, thus achieving initialization for US and MR image registration.
2. Instead of heatmap regression or coordinate regression using fully connected layer, we adopt the differentiable spatial to numerical transform (DSNT) layer [18, 20] and combine it with a 3D U-Net in order to regress coordinates of landmarks.
3. We introduce a novel heatmap regularization term, which penalizes large heatmap values far away from the ground truth location.

2 Methodology

Given a 3D US and MR image pair, F, M, respectively, with corresponding landmarks $\boldsymbol{G_i}^f \in \mathbb{R}^3$ and $\boldsymbol{G_i^m} \in \mathbb{R}^3$, where i is the landmark index, $i = [1, 2 \dots L]$, and L is the number of landmarks. Our goal is to train a neural network that predicts the coordinates of L landmarks, then calculate a rigid transformation based on the detected landmark coordinates from both modalities.

2.1 End-to-end landmark detection network

Inspired by landmark detection in human pose estimation [18], we formulate landmark detection problem as a direct coordinate regression task. The proposed neural network consists in a 3D U-Net and a differentiable spatial to numerical transform (DSNT) layer, which transforms spatial heatmaps from the output of U-Net into numerical coordinates, shown in Fig. 2. We consider an input image I, which is either F or M, of size N, and the network outputs a matrix $\boldsymbol{P}$ of size $L \times 3$. To generate this matrix, each raw heatmap $\tilde{H}^i$ of size N is first

normalized with a softmax activation function into H^i such that $\sum_{j \in I} H^i_j = 1$, $H^i_j \in]0, 1[$ for $i = 1 \dots L$. The DNST layer computes each landmark $\boldsymbol{P_i} \in \mathbb{R}^3$ as the expectation of the voxel position based on each probabilistic maps:

$$\boldsymbol{P}_i = \mathbb{E}_{H^i}(\boldsymbol{V}) = \text{DSNT}(H^i) = \sum_{j \in I} H^i_j \boldsymbol{v}_j \tag{1}$$

where $\boldsymbol{v}_j$ is the 3D position of voxel j in image I and $\boldsymbol{P}_i$ is the estimated i_{th} landmark position.

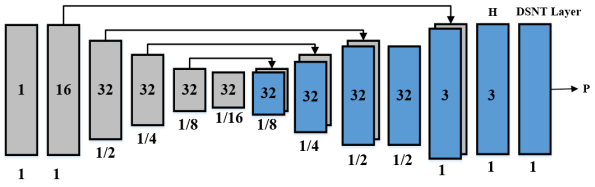

Fig. 2. Overview of the proposed end-to-end landmark detection network with $L = 3$, consisting of 3D U-Net and DSNT layer. The proposed network is trained separately for each modality. It's inspired by Balakrishnan *et al.* [21].

2.2 Regression Loss Function

Since DSNT layer transforms spatial heatmaps into numerical coordinates directly, it's possible to calculate the L1 and L2 norms between the ground truth and prediction coordinate vector (Equation 2), and it is named as vanilla DSNT.

$$\mathcal{L}(\boldsymbol{G}, \boldsymbol{P}) = \|\boldsymbol{P} - \boldsymbol{G}\|_1 + \lambda_1 \|\boldsymbol{P} - \boldsymbol{G}\|_2 \tag{2}$$

where G is the matrix of ground truth landmark position.

The spread of the heatmap has no effect on the output such that heatmaps with small or large variance can produce the same landmark position. Based on [18], we propose to regularize the probabilistic map H variance to achieve better performance than vanilla DSNT. The overall loss function is a combination between coordinate regression loss and heatmap regularization loss.

$$\mathcal{L}(\boldsymbol{G}, \boldsymbol{P}) = \|\boldsymbol{P} - \boldsymbol{G}\|_1 + \lambda_1 \|\boldsymbol{P} - \boldsymbol{G}\|_2 + \lambda_2 \mathcal{L}_{reg}(\boldsymbol{H}) \tag{3}$$

Variance Regularization. As a first option, the variance of each probabilistic map is used to regularize the regression. In [18], the authors proposed a regularization term based on a specific target variance. We instead propose to minimize the overall variance (equivalent to specifying a zero target variance) thus avoiding to pick an additional hyperparameter which may be data dependent. Besides, this choice forces the network to make a bias-variance trade-off in a data driven way. The computation of the variance as the second order moment of the probabilistic maps which extends the approach proposed in [18]:

$$\mathcal{L}_{\text{Var}} = \mathbb{E}_{H^i}(\|\boldsymbol{P}_i - \mathbb{E}_{H^i}(\boldsymbol{V})\|^2) = \sum_{j \in I} H^i_j \|\boldsymbol{v}_j - \boldsymbol{P}_i\|^2 \tag{4}$$

Distance Map Regularization. As a second option, we propose a new regularization term $\mathcal{L}_{\text{Dist}}$, which penalizes high probability values that are far way from the ground truth landmark position $\boldsymbol{G}_i$:

$$\mathcal{L}_{\text{Dist}} = \sum_{i=1}^{L} \sum_{j \in I} H_j^i \|\boldsymbol{v}_j - \boldsymbol{G}_i\| \tag{5}$$

2.3 Multimodal Landmark-based Rigid Registration

Once anatomical landmark coordinates have been predicted in both modalities, the calculation of rigid matrix can be formulated as the least-square optimization problem. $\boldsymbol{P}_i^m$ and $\boldsymbol{P}_i^f$ are the coordinate vectors of the i^{th} landmark in the MR and US images, $\boldsymbol{R}$ is a 3×3 rotation matrix, and $\boldsymbol{t}$ is a 3×1 translation vector. To solve this problem, we use the noniterative SVD-based algorithm proposed in [19], and the equation is $\min_{\boldsymbol{R},\boldsymbol{t}} \sum_{i=1}^{3} \left\| \boldsymbol{P}_i^f - (\boldsymbol{R}\boldsymbol{P}_i^m + \boldsymbol{t}) \right\|^2$.

3 Experiments

3.1 Data and Training

The database contains 1315 patients, each of them consisting of a MRI-US volume pair, prostate segmentations, and landmarks. All cases were scheduled for prostate biopsy. Each MRI volume has 256×256×128 voxels with a voxel size of 0.5 × 0.5× 1.0 mm^3, and each US volume has 256×256×256 voxels with 0.4 mm resolution in all directions. US images are coming from various ultrasound systems, with various probes, both end-fire and side-fire probes. For both MR and US images, three anatomical landmarks (the apex, the bladder neck, the posterior median) are detected by medical experts. We used 900 cases of MRI-US volume pair for training, 152 cases for validation, and 263 cases for testing. For each modality, we train a landmark detection network. As shown in Fig. 2, we used 3D U-Net structure to map a whole 3D image to 3 probability maps ($L = 3$), one for each landmark. We apply 3D convolutions in both the encoder and decoder stages using a kernel size of 3, and a stride of 2. Each convolution is followed by a LeakyReLU layer with parameter 0.2. In the encoder, we use strided convolutions to reduce the spatial dimensions in half at each layer. The softmax activation function is applied to normalize each heatmap. Finally, the DSNT layer transforms spatial heatmaps into numerical coordinates. The loss hyperparameter was empirically chosen as $\lambda_1 = 1$ and $\lambda_2 = 5 \times 10^{-4}$ on MR images (resp. $\lambda_2 = 2 \times 10^{-3}$ on US images) for the variance regularization and $\lambda_2 = 5 \times 10^{-3}$ on MR images (resp. $\lambda_2 = 10^{-2}$ on US images) for the distance map regularization. An Adam optimizer is used with a learning rate initialized to $\text{lr} = 1 \times 10^{-4}$. The neural network was implemented with PyTorch framework and trained on one NVIDIA RTX 8000 GPU with batch size of 4.

3.2 Experimental Results

Landmark Detection Results and Ablation Study. Various landmark detection methods were evaluated qualitatively and quantitatively in Table 1. Following previous works [22], we use Mean Radial Error (MRE) as a metric, which is the average Euclidean distance between the predicted landmarks and the ground-truth landmarks measured in mm. As baseline method, we consider the proposed end-to-end method with DSNT layer without any heatmap regularization. From Table 1, we can see that regularizing heatmap improves significantly the model's performance. The proposed method (with and without regularization) outperforms the heatmap matching or direct regression coordinates methods [15, 17] based on MRE. In Fig. 3, a successful example of posterior median landmark detection is shown for both modalities.

Table 1. Landmark detection results on MR and US images using different methods. We report statistically significant differences based on the Wilcoxon test from the baseline model with a $*$ sign, and from the variance regularization model with a † sign.

Method	US MRE (mm)	MRI MRE (mm)
Direct regression coordinate [17]	11.43 ± 3.82	13.83 ± 7.25
Heatmap matching [15]	6.91 ± 5.12	6.93 ± 7.57
Baseline (no regularization)	5.90 ± 2.91	6.24 ± 2.79
Variance regularization	$5.53 \pm 2.85^{*}$	$6.05 \pm 2.97^{*}$
Distance map regularization	$5.55 \pm 2.63^{*}$	$5.77 \pm 2.67^{*\dagger}$

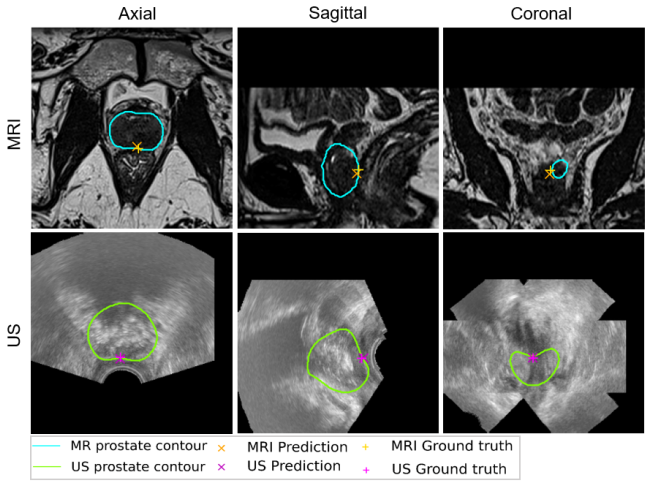

Fig. 3. Visualization results of the posterior median landmark for both modalities. The contour of US and MR prostate segmentation is highlighted in green and blue, separately.

Registration error based on Detected Landmarks. Following [6], we also evaluate the proposed method based on surface registration error (SRE). In the remainder, we consider that the ground truth rigid transformation is best estimated by performing an Iterative Closest Point (ICP) algorithm between the prostate meshes extracted from binary segmentation masks in the US and MR images. We write $\boldsymbol{T}_{\mathbf{gt}}$ as the mesh-based rigid transformation from MR to US images, $\boldsymbol{T}_{\mathbf{man}}$ as the rigid transform estimated from manual landmarks and $\boldsymbol{T}_{\mathbf{pred}}$ as the rigid transform predicted by applying the automatic landmark detection on both MR and US images. The SRE metric measures the displacement error due to a rigid misalignment on the prostate surface. If we write $\boldsymbol{S}_i \in \mathbb{R}^3, i = 1 \dots n$, a surface point of a segmented prostate mesh from the MR T2w image, the SRE for the automatic landmark detection computes as:

$$\mathrm{SRE}(\boldsymbol{T}_{\boldsymbol{gt}}, \boldsymbol{T}_{\boldsymbol{pred}}) = \frac{1}{n} \sum_{i=1}^{n} \left\| \boldsymbol{T}_{\mathbf{gt}} \left(\boldsymbol{S}_i \right) - \boldsymbol{T}_{\mathbf{pred}} \left(\boldsymbol{S}_i \right) \right\|_2 \tag{6}$$

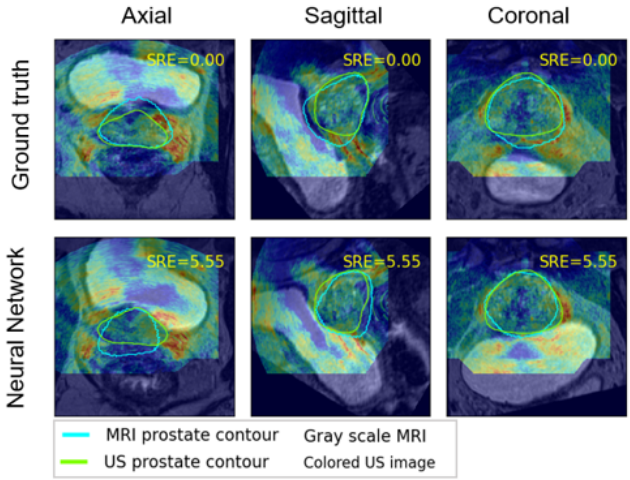

Fig. 4. Visualization of registration result. The top (resp. bottom) row show the rigid registration based on ICP algorithm (resp. neural network landmark detection) aligning MR and US images.

The registration results on all 263 MR-US image pairs in the test set are shown in Table 2, including the surface registration error (SRE) and the Dice score between the transformed MR prostate mask and US prostate mask. Visual comparison between the ground truth transformation $\boldsymbol{T}_{\boldsymbol{gt}}$ and the automatic network prediction is shown in Fig. 4.

It shows that our fully automatic approach does not match the manual landmark accuracy but is comparable to initial SRE and Dice score in clinical setting from Song *et al.* [23] (resp. 7.98 ± 5.01 mm and 0.77 ± 0.14 Dice).

We can see the histogram of SRE values for our fully automatic approach and the manual landmark approach in Table 3. For fully automatic approach,

Table 2. Image registration performance of different methods.

Metric	Neural Network	Manual Landmark
SRE (mm)	6.62 ± 3.97	5.80 ± 3.08
Dice score	0.77 ± 0.11	0.82 ± 0.06

most cases are smaller than 15 mm with few outliers, similarly to the SRE based on manual landmark cases.

Table 3. Frequency distribution (histogram) of SRE on the test set.

Threshold (mm)	5	10	15	20	25	30	45
Neural Network	93	137	28	1	3	0	1
Manual Landmark	122	127	9	4	0	1	0

The residual rotation matrix $\boldsymbol{R_{res}} = \boldsymbol{R_{gt}}\boldsymbol{R_{pred}^{-1}}$ captures the amount of rotation that needs to be compensated by any rigid registration method estimated after the rigid initialization stage based on predicted landmarks. If we convert this residual matrix into a rotation vector, then we produce the histogram of the rotation angle given by the norm of rotation vector, as seen in Table 4. For the fully automatic landmark selection approach, 213 (81%) cases are under 20 degrees, whereas this occurs 229 (87%) for the manually selected landmarks. While both histograms are similar, this suggests that the robustness of the automated landmark detection should be improved. After looking at the data, we found that large registration errors are associated with the following reasons : MR and US image quality, partial view of the prostate in US image, very large deformation of the prostate, ambiguous ground truth position of posterior median landmark (#3). The strategies to deal with it include detecting these US images using intensity, detecting discrepancies between the predictions of two different models (with and without regularization), and visual inspection.

Table 4. Frequency distribution (histogram) of the norm of the residual rotation vector on the test set.

Threshold (degree)	5	10	20	30	40	60	70
Neural Network	18	71	124	37	8	5	0
Manual Landmark	29	88	112	24	3	6	1

4 Conclusion

We have proposed a pipeline to detect prostate anatomical landmarks of both US and MR images automatically, then least-squares fitting is applied to calculate a rough rigid transformation based on detected anatomical landmarks, thus achieving initialization for US and MR image registration. Intense experimental results have demonstrated that our method can detect anatomical landmarks for US and MR images in terms of MRE. A rough rigid transformation is calculated based on detected anatomical landmarks, which achieves comparable results in

terms of SRE and Dice score in clinical setting when comparing with previous method.

Acknowledgements This work has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002.

References

1. Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249, 2021.
2. M Minhaj Siddiqui, Soroush Rais-Bahrami, Baris Turkbey, Arvin K George, Jason Rothwax, Nabeel Shakir, Chinonyerem Okoro, Dima Raskolnikov, Howard L Parnes, W Marston Linehan, et al. Comparison of mr/ultrasound fusion–guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer. *Jama*, 313(4):390–397, 2015.
3. Tony CW Mok and Albert CS Chung. Large deformation diffeomorphic image registration with laplacian pyramid networks. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part III 23*, pages 211–221. Springer, 2020.
4. Yue Sun, Jing Yuan, Wu Qiu, Martin Rajchl, Cesare Romagnoli, and Aaron Fenster. Three-dimensional nonrigid mr-trus registration using dual optimization. *IEEE transactions on medical imaging*, 34(5):1085–1095, 2014.
5. Yipeng Hu, Marc Modat, Eli Gibson, Wenqi Li, Nooshin Ghavami, Ester Bonmati, Guotai Wang, Steven Bandula, Caroline M Moore, Mark Emberton, et al. Weakly-supervised convolutional neural networks for multimodal image registration. *Medical image analysis*, 49:1–13, 2018.
6. Xinrui Song, Hengtao Guo, Xuanang Xu, Hanqing Chao, Sheng Xu, Baris Turkbey, Bradford J Wood, Ge Wang, and Pingkun Yan. Cross-modal attention for mri and ultrasound volume registration. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part IV 24*, pages 66–75. Springer, 2021.
7. Huanye Li, Chau Hung Lee, David Chia, Zhiping Lin, Weimin Huang, and Cher Heng Tan. Machine learning in prostate mri for prostate cancer: current status and future opportunities. *Diagnostics*, 12(2):289, 2022.
8. Shyam Natarajan, Leonard S Marks, Daniel JA Margolis, Jiaoti Huang, Maria Luz Macairan, Patricia Lieu, and Aaron Fenster. Clinical application of a 3d ultrasound-guided prostate biopsy system. In *Urologic oncology: seminars and original investigations*, volume 29, pages 334–342. Elsevier, 2011.
9. Mattias P Heinrich and Lasse Hansen. Voxelmorph++ going beyond the cranial vault with keypoint supervision and multi-channel instance optimisation. In *Biomedical Image Registration: 10th International Workshop, WBIR 2022, Munich, Germany, July 10–12, 2022, Proceedings*, pages 85–95. Springer, 2022.

10. Monika Grewal, Jan Wiersma, Henrike Westerveld, Peter AN Bosman, and Tanja Alderliesten. Automatic landmark correspondence detection in medical images with an application to deformable image registration. *Journal of Medical Imaging*, 10(1):014007–014007, 2023.
11. Xinrui Song, Xuanang Xu, Sheng Xu, Baris Turkbey, Thomas Sanford, Bradford J Wood, and Pingkun Yan. Distance map supervised landmark localization for mr-trus registration. In *Medical Imaging 2023: Image Processing*, volume 12464, pages 708–713. SPIE, 2023.
12. Antonio Criminisi, Duncan Robertson, Ender Konukoglu, Jamie Shotton, Sayan Pathak, Steve White, and Khan Siddiqui. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical image analysis*, 17(8):1293–1303, 2013.
13. Amir Alansary, Ozan Oktay, Yuanwei Li, Loic Le Folgoc, Benjamin Hou, Ghislain Vaillant, Konstantinos Kamnitsas, Athanasios Vlontzos, Ben Glocker, Bernhard Kainz, et al. Evaluating reinforcement learning agents for anatomical landmark detection. *Medical image analysis*, 53:156–164, 2019.
14. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.
15. Christian Payer, Darko Štern, Horst Bischof, and Martin Urschler. Regressing heatmaps for multiple landmark localization using cnns. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 230–238. Springer, 2016.
16. Xu Wang, Xin Yang, Haoran Dou, Shengli Li, Pheng-Ann Heng, and Dong Ni. Joint segmentation and landmark localization of fetal femur in ultrasound volumes. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 1–5. IEEE, 2019.
17. Jiahong Qian, Ming Cheng, Yubo Tao, Jun Lin, and Hai Lin. Cephanet: An improved faster r-cnn for cephalometric landmark detection. In *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*, pages 868–871. IEEE, 2019.
18. Aiden Nibali, Zhen He, Stuart Morgan, and Luke Prendergast. Numerical coordinate regression with convolutional neural networks. *arXiv preprint arXiv:1801.07372*, 2018.
19. K Somani Arun, Thomas S Huang, and Steven D Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on pattern analysis and machine intelligence*, (5):698–700, 1987.
20. Milosz Gajowczyk, Patryk Rygiel, Piotr Grodek, Adrian Korbecki, Michal Sobanski, Przemyslaw Podgorski, and Tomasz Konopczynski. Coronary ostia localization using residual u-net with heatmap matching and 3d dsnt. In *International Workshop on Machine Learning in Medical Imaging*, pages 318–327. Springer, 2022.
21. Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019.
22. James McCouat and Irina Voiculescu. Contour-hugging heatmaps for landmark detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20597–20605, 2022.

23. Xinrui Song, Hanqing Chao, Xuanang Xu, Hengtao Guo, Sheng Xu, Baris Turkbey, Bradford J Wood, Thomas Sanford, Ge Wang, and Pingkun Yan. Cross-modal attention for multi-modal image registration. *Medical Image Analysis*, 82:102612, 2022.