XA-Sim2Real: Adaptive Representation Learning for Vessel Segmentation in X-ray Angiography

Baochang Zhang[1,2], Zichen Zhang[1], Shuting Liu[1], Shahrooz Faghihroohi[1], Heribert Schunkert[2,3], and Nassir Navab[1]

[1] Computer Aided Medical Procedures,
Technical University of Munich, Munich, Germany
baochang.zhang@tum.de
[2] German Heart Center Munich, Munich, Germany
[3] German Centre for Cardiovascular Research,
Munich Heart Alliance, Munich, Germany

Abstract. Accurate vessel segmentation from X-ray Angiography (XA) is essential for various medical applications, including diagnosis, treatment planning, and image-guided interventions. However, learning-based methods face challenges such as inaccurate or insufficient manual annotations, anatomical variability, and data heterogeneity across different medical institutions. In this paper, we propose XA-Sim2Real, a novel adaptive framework for vessel segmentation in XA image. Our approach leverages Digitally Reconstructed Vascular Radiographs (DRVRs) and a two-stage adaptation process to achieve promising segmentation performance on XA image without the need for manual annotations. The first stage involves an XA simulation module for generating realistic simulated XA images from patients' CT angiography data, providing more accurate vascular shapes and backgrounds than existing curvilinear-structure simulation methods. In the second stage, a novel adaptive representation alignment module addresses data heterogeneity by performing intra-domain adaptation for the complex and diverse nature of XA data in different settings. This module utilizes self-supervised and contrastive learning mechanisms to learn adaptive representations for unlabeled XA image. We extensively evaluate our method on both public and in-house datasets, demonstrating superior performance compared to state-of-the-art self-supervised methods and competitive performance compared to supervised method.

Keywords: Vessel Segmentation · Domain Adaptation · Digitally Reconstructed Vascular Radiographs.

1 Introduction

X-ray Angiography (XA) is a common medical imaging modality for the diagnosis and treatment of cardiovascular diseases. Vessel segmentation from XA images can further highlight vascular positions and structures, which is beneficial

for relevant procedures. While supervised segmentation methods have demonstrated success in 2D/3D vascular segmentation across various anatomical regions [11, 15], their application to XA poses challenges due to the scarcity of extensive datasets and annotations required for training. The inherently noisy visual appearance of XA images, coupled with the complex anatomical background and varying vessel geometries, results in annotation efforts that are labor-intensive and time-consuming. Moreover, existing unsupervised methods encounter difficulties when applied to medical vessel images, primarily due to the intricate nature of vascular structures and the variability in imaging conditions [1, 6].

In recent years, some self-supervised methods have been proposed for XA vessel segmentation. A self-supervised vessel segmentation approach [5] has been presented to learn vessel representations, which requires two different adversarial networks to segment vessels, which leads to increasing training complexity. Then a non-iterative version of diffusion-based generative model [4] was proposed and successfully applied to self-supervised vessel segmentation. However, this model requires significant computational resources and time for training. Both methods utilize mathematically generated fractal shapes [13] to synthesize vessel-like masks for fake angiogram generation, but these fractal shapes exhibit significant morphological discrepancies compared to realistic vessels. And their methods are constrained by the requirement for clean background images as input, thus greatly limiting their applicability. A self-supervised curvilinear object segmentation method [10] was proposed, which presents an enhanced fractal generation system and discards the usage of background images. But it employs Fourier domain adaptation [12] for XA image generation, which imposes limitations on generation capabilities and necessitates additional hyperparameter selection. Recently, a primary but interesting study was conducted for vessel segmentation, which integrates with adversarial learning and self-supervised learning to transfer the knowledge learned from Digitally Reconstructed Vascular Radiographs (DRVRs) to real XA images [16]. However, there is still much room for improvement in performance.

In this paper, we reformulate vessel segmentation for unlabeled XA data into a two-stage domain adaptation procedure to improve segmentation performance. Our contributions are as follows. (1) A novel adaptive representation alignment technique is introduced to address data heterogeneity by performing intra-domain adaptation for the complex and diverse nature of real XA data in different settings. (2) To address the limitations imposed by fractals, we leverage clinical knowledge and XA imaging principles to generate DRVRs and refine their resemblance to real XA images using a generative adversarial network. (3) We build and release the first huge simulated coronary artery XA dataset[4], which contains DRVR images and corresponding vessel masks. (4) Our framework surpasses state-of-the-art self-supervised methods by a large margin and achieves competitive performance to supervised method on a public dataset and two in-house datasets.

[4] DRVR Dataset: https://github.com/BaochangZhang/XSim2Real

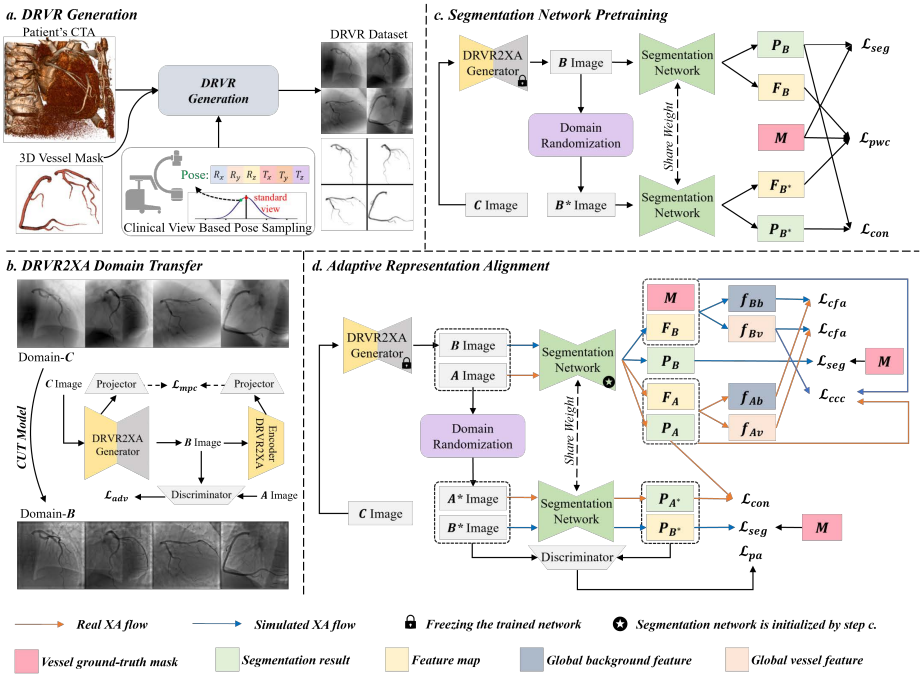

Fig. 1. The overview of our proposed XA-Sim2Real framework. ***A*** image means real XA image, ***B*** image means simulated XA image, ***C*** image means DRVR image, and * means augmented image.

2 Method

An overview of the proposed framework is illustrated in Fig. 1, which can be mainly divided into four parts that are introduced hereafter.

2.1 XA Simulation Module

DRVR Generation. Inspired by Diff-DeepDRR[14], DRVR generation procedure is developed for more realistic XA simulation than fractal-based methods [10], as shown in Fig. 1(a). For the DRVR generation process, given a CTA data $\mathcal{V}$, an available X-ray imaging system Ω and the pose parameters θ, the density at detector point p is calculated as:

$$D(p) = \int P_o(E) exp\Big\{ -\sum_{m} \sum_{\Omega_p(\theta)} \delta(m, G(M(\mathcal{V})|\Omega_p(\theta))) \\ \phi_m(E) G(\rho(\mathcal{V})|\Omega_p(\theta))\Delta s \Big\} dE \tag{1}$$

where $P_o(E)$ means the X-ray beam spectra obtained using SPEKTR3.0 software [8]. $\delta(\cdot,\cdot)$ is the Kronecker delta. $M(\mathcal{V})$ are decomposed materials' mask,

i.e., bone, blood vessel with contrast agent, soft tissue and air. $\phi_m(E)$ is the material and energy dependent linear attenuation coefficient [3]. The attenuation coefficient of blood vessel with contrast agent is the weighted average of the attenuation coefficients of blood and iodine, where the weight of the iodine's attenuation coefficient is 0.2. $\rho(\mathcal{V})$ is the mass density computed from HU values. $\Omega_p(\theta) \in \mathcal{R}^{N\times 3}$ means the world coordinates of N sampling points on the ray from the source to detector point p. Δs is the sampling step along the ray. $G(*|\Omega_p)$ is a 3D gird sampling function based on a bilinear interpolation scheme, which will output the values of N sampling points on the given volume $*$. Based on the public ASOCA dataset [2], A DRVR dataset containing 19384 images and corresponding masks is generated following 11 standard clinical coronary artery view poses, all in a resolution of 512 $\times$ 512 pixels.

DRVR2XA Domain Transfer. To reduce the inter-domain gap between DRVR images and XA images, Contrastive Unpaired Translation (CUT) model [7] is employed to perform a domain transfer, which is trained using our generated DRVR dataset and all XA images from the public XCAD dataset [5]. Fully following CUT, adversarial loss $\mathcal{L}_{adv}$ and multi-layer patch-wise contrastive loss $\mathcal{L}_{mpc}$ are utilized for training. Some DRVR images and corresponding simulated XA images are presented in Fig. 1(b). Note that this is a one-time process and we do not retrain or refine CUT specifically for the new XA dataset.

2.2 Segmentation Network Pre-training

Unet [9] is used as vessel segmentation network, and it is pre-trained using simulated XA images and vessel masks inherited from the DRVR dataset, as shown in Fig. 1(c). Dice loss function $\mathcal{L}_{dice}$ and binary cross entropy loss function $\mathcal{L}_{bce}$ are used as segmentation loss $\mathcal{L}_{seg}$. Meanwhile, a consistency loss $\mathcal{L}_{con}$ based on L1-norm is employed to make the prediction maps (P_{B^*}, P_B) learned from the augmented input and the original input consistent, thereby making the model more robust and achieving better generalization capabilities.

To improve the feature distinctiveness between vessel and background, a pixel-wise contrastive loss $\mathcal{L}_{pwc}$ is applied to the feature maps extracted by the third-to-last convolution layer in Unet. Utilizing ground truth mask M, we sample n vessel features from the feature map F_B extracted from non-augmented input and treat them as a query set $\{q_i\} \in f_v$. A paired positive set $\{k_i^+\} \in f_{v^+}$ is collected from the feature map F_{B^*} extracted from augmented input at the corresponding vessel positions. The negative set $\{k_j^-\} \in f_{v^-}$ is randomly sampled from the feature maps (F_{B^*}, F_B) extracted from both input at the background positions. Then the pixel-wise contrastive loss $\mathcal{L}_{pwc}$ is formulated as,

$$\mathcal{L}_{pwc} = -\frac{1}{n}\sum_{i=1}^{n}\left\{log\frac{exp(q_i \cdot k_i^+/\tau)}{exp(q_i \cdot k_i^+/\tau) + \sum_{j=1}^{w} exp(q_i \cdot k_j^-/\tau)}\right\} \tag{2}$$

where n is set to 1024, w is set to 2048, and the temperature parameter τ is set to 0.1. And the final loss $\mathcal{L}_{pretrain}$ is defined as,

$$\mathcal{L}_{pretrain} = \mathcal{L}_{dice}(M, P_B) + \mathcal{L}_{bce}(M, P_B) + \mathcal{L}_{con}(P_B, P_{B^*}) + \mathcal{L}_{pwc} \tag{3}$$

2.3 Adaptive Representation Alignment

Our objective is to tackle the question: "How can we ensure that a pretrained vessel segmentation model performs effectively on a new real XA dataset without the need for manual annotations?". To this end, we propose adaptive representation alignment shown in Fig. 1(d), as a means to mitigate intra-domain gap between simulated XA data and various sources of real XA data, focusing on aligning their representations. The segmentation network is initialized with pretrained weights. Subsequently, both simulated XA images and their augmented versions, alongside real XA images and their augmented versions, are simultaneously utilized as inputs for the segmentation network. The segmentation loss $\mathcal{L}_{seg}$ employed in Section 2.2 is retained for optimizing the segmentation outcomes of both the simulated XA image and its augmented version. And the consistency loss $\mathcal{L}_{con}$ is also retained, yet this time it is applied between the prediction maps (P_{A^*}, P_A) learned from the augmented XA input and the original XA.

To minimize the average distance between features of simulated XA and real XA, and enhance the feature distinctiveness between vessels and backgrounds on XA, two additional loss functions are proposed and applied to the feature maps (F_A, F_B) extracted by the third-to-last convolution layer in Unet. First, class centroid feature alignment loss $\mathcal{L}_{cfa}$ is introduced to perform representation alignment on the centroids of the vessel and background features between simulated XA and real XA, respectively. For simulated XA flow, the mini-batch vessel class centroid feature $\tilde{f}_{Bv}$ can be calculated by averaging the features of all vessel positions as indicated by the ground truth mask M, similarly, the mini-batch background class centroid feature $\tilde{f}_{Bb}$ follows the same process. For real XA flow, the prediction map P_A is used to distinguish vessel and background positions to calculate mini-batch class centroid feature $\tilde{f}_{Av}$ and $\tilde{f}_{Ab}$, respectively. And the overall class centroid features are updated using an exponentially moving average, which is defined as,

$$f_{Dc} = (1 - \alpha) f_{Dc} + \alpha \tilde{f}_{Dc} \tag{4}$$

where f_{Dc} is the centroid feature of class $c \in \{v, b\}$ from the domain $D \in \{A, B\}$, and α is set to 0.9. Therefore, the class centroid feature alignment loss $\mathcal{L}_{cfa}$ is formulated as,

$$\mathcal{L}_{cfa} = \|f_{Av} - f_{Bv}\|_2 + \|f_{Ab} - f_{Bb}\|_2 \tag{5}$$

Second, class centroid contrastive loss $\mathcal{L}_{ccc}$ is proposed, which is formulated as,

$$\mathcal{L}_{ccc} = -\frac{1}{n} \sum_{i=1}^{n} \Big\{ log \frac{exp(f_{Bv} \cdot h_i / \tau)}{exp(f_{Bv} \cdot h_i / \tau) + \sum_{j=1}^{w} exp(f_{Bv} \cdot g_j / \tau)} \Big\} \tag{6}$$

where the vessel set $\{h_i\}$ is collected from the feature maps F_A and F_B at the corresponding vessel positions. And the background set $\{g_j\}$ is sampled from the feature map F_B at the background positions.

Meanwhile, we utilize adversarial learning to explicitly align the prediction space distribution of real XA images and simulated XA images. A discriminator D is employed to evaluate the quality of the segmentation result. It takes not only the prediction maps as input but also the gradient magnitude maps of XA images. The prediction space adversarial loss $\mathcal{L}_{pa}$ for the discriminator is defined as,

$$\mathcal{L}_{pa}^{D} = \mathbb{E}[log(D(P_{B^*}, \mathcal{G}(B^*)))] + \mathbb{E}[log(1 - D(P_{A^*}, \mathcal{G}(A^*)))] \tag{7}$$

Where $\mathcal{G}(A^*)$ and $\mathcal{G}(B^*)$ are the gradient magnitude maps of the A^* images and the B^* images, respectively. And $\mathcal{L}_{pa}$ for the segmentation network S is defined as,

$$\mathcal{L}_{pa}^{S} = \mathbb{E}[log(D(P_{A^*}, \mathcal{G}(A^*)))] \tag{8}$$

Therefore, the total loss function for segmentation network is summarized as,

$$\mathcal{L}_{total} = 2.5\mathcal{L}_{pa}^{S} + 0.5\mathcal{L}_{ccc} + 5\mathcal{L}_{cfa} + 2\mathcal{L}_{seg} + 100\mathcal{L}_{con} \tag{9}$$

2.4 Implementation Details

The proposed framework is implemented on PyTorch library with one NVIDIA GPU (Quadro RTX A6000). For the pre-training step, the model is trained for 30 epochs on DRVR dataset with a batch size of 16, using the Adam optimizer (initial learning rate is 0.001, decay of 0.75 per 3 epochs). For the adaptive representation alignment training step, the model is trained for 100 epochs on XA dataset with a batch size of 16, using the Adam optimizer (initial learning rate is 0.0001, decay of 0.85 per 10 epochs).

For the domain randomization, each effect is conducted with probability 50%, and only one option is randomly selected for each effect. The employed effects are as follows, (a) random intensity transformation with four options, including random adding offset, random gamma correction, non-linear transformation and contrast-limited adaptive histogram equalization; (b) random noise adding with three options, including Gaussian noise, Poisson noise and speckle noise; and (c) random crop and resize.

3 Experiments and Results

3.1 Dataset and Evaluation Metrics

Our method is evaluated on three XA datasets. (i) The public XCAD dataset [5] contains a training set of 1621 XA images and 1621 background images (i.e., X-ray image without contrast agent), and a test set of 126 XA images with vessel annotations. (ii) An in-house I-XA dataset contains a training set of 90 XA images and 90 background images, and a test set of 70 XA images with vessel annotations. (iii) Another in-house II-XA dataset contains a training set

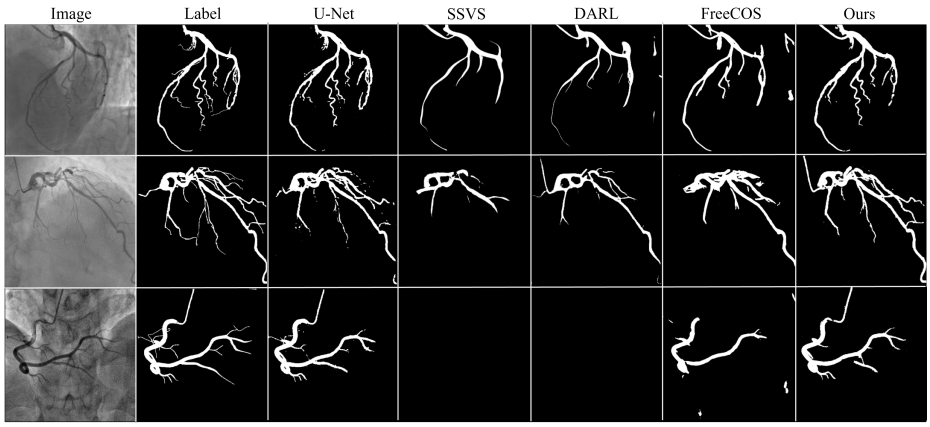

Fig. 2. Visualization results of the XA images, the labels, and the segmentation from the three datasets (from top to bottom): XCAD, I-XA, and II-XA examples.

of 80 XA images and a test set of 70 XA images with vessel annotations. All images in the three datasets have a resolution of 512 × 512 pixels.

Six metrics Jaccard Index (Jac.), Dice Coefficient (Dice), Accuracy (Acc.), Sensitivity (Sen.), Precision (Pre.) and Specificity (Spe.) are used to evaluate the performance of coronary vessel segmentation, following the paper [5].

3.2 Comparison with State-of-the-art

Table. 1 presents a comparison of coronary vessel segmentation performance across three XA datasets between our proposed XA-Sim2Real method and state-of-the-art self-supervised vessel segmentation techniques, including SSVS [5], DARL [4], and FreeCOS [10]. Initially, we benchmark our method against the supervised UNet [9], trained using test sets from the three XA datasets and evaluated through 5-fold cross-validation. Although the supervised UNet achieves slightly higher performance than our method on the XCAD and II-XA datasets, its reliance on manually annotated vessel data for training poses a significant limitation. In contrast, our method achieves competitive performance on these datasets without the need for manual annotations. Notably, it demonstrates superior performance than UNet on the I-XA dataset with a dice score 0.727. When compared to self-supervised methods such as SSVS, DARL, and FreeCOS, our method exhibits significantly superior performance across all metrics on the three datasets. Notably, SSVS and DARL necessitate training with clean background images, limiting their applicability to II-XA data. Moreover, Fig. 2 visually depicts the segmentation results of XA images from the three datasets. As evident from Fig. 2, our method produces segmentations with richer vessel representation and reduced background noise.

Table 1. Comparison results on three XA datasets.

Dataset	*Method*	*Jac.*	*Dice*	*Acc.*	*Sen.*	*Spe.*	*Pre.*
XCAD	Unet [9]	$0.632_{\pm 0.082}$	$0.772_{\pm 0.066}$	$0.977_{\pm 0.006}$	$0.766_{\pm 0.099}$	$0.989_{\pm 0.005}$	$0.789_{\pm 0.077}$
	SSVS [5]	$0.389_{\pm 0.062}$	$0.557_{\pm 0.066}$	$0.948_{\pm 0.008}$	$0.630_{\pm 0.074}$	$0.967_{\pm 0.006}$	$0.510_{\pm 0.095}$
	DARL [4]	$0.481_{\pm 0.095}$	$0.644_{\pm 0.090}$	$0.962_{\pm 0.014}$	$0.657_{\pm 0.095}$	$0.980_{\pm 0.010}$	$0.648_{\pm 0.132}$
	FreeCOS [10]	$0.524_{\pm 0.083}$	$0.684_{\pm 0.073}$	$0.966_{\pm 0.011}$	$0.707_{\pm 0.094}$	$0.980_{\pm 0.010}$	$0.671_{\pm 0.090}$
	Our	$0.631_{\pm 0.077}$	$0.771_{\pm 0.061}$	$0.975_{\pm 0.006}$	$0.810_{\pm 0.092}$	$0.985_{\pm 0.005}$	$0.742_{\pm 0.067}$
I-XA	Unet [9]	$0.561_{\pm 0.101}$	$0.713_{\pm 0.092}$	$0.975_{\pm 0.008}$	$0.726_{\pm 0.118}$	$0.987_{\pm 0.007}$	$0.723_{\pm 0.120}$
	SSVS [5]	$0.329_{\pm 0.073}$	$0.490_{\pm 0.090}$	$0.954_{\pm 0.011}$	$0.513_{\pm 0.101}$	$0.976_{\pm 0.008}$	$0.497_{\pm 0.133}$
	DARL [4]	$0.428_{\pm 0.086}$	$0.594_{\pm 0.091}$	$0.964_{\pm 0.010}$	$0.619_{\pm 0.083}$	$0.981_{\pm 0.008}$	$0.598_{\pm 0.151}$
	FreeCOS [10]	$0.439_{\pm 0.099}$	$0.603_{\pm 0.102}$	$0.963_{\pm 0.017}$	$0.637_{\pm 0.151}$	$0.978_{\pm 0.017}$	$0.606_{\pm 0.138}$
	Our	$0.578_{\pm 0.102}$	$0.727_{\pm 0.087}$	$0.974_{\pm 0.012}$	$0.772_{\pm 0.077}$	$0.983_{\pm 0.010}$	$0.698_{\pm 0.121}$
II-XA	Unet [9]	$0.689_{\pm 0.044}$	$0.815_{\pm 0.031}$	$0.982_{\pm 0.005}$	$0.788_{\pm 0.062}$	$0.993_{\pm 0.003}$	$0.851_{\pm 0.048}$
	FreeCOS [10]	$0.513_{\pm 0.056}$	$0.676_{\pm 0.049}$	$0.967_{\pm 0.008}$	$0.694_{\pm 0.070}$	$0.981_{\pm 0.009}$	$0.672_{\pm 0.091}$
	Our	$0.669_{\pm 0.071}$	$0.800_{\pm 0.054}$	$0.980_{\pm 0.006}$	$0.812_{\pm 0.052}$	$0.989_{\pm 0.005}$	$0.795_{\pm 0.083}$

Table 2. Ablation study results on the I-XA dataset.

XA	$\mathcal{L}_{con}$	$\mathcal{L}_{cfa}$	$\mathcal{L}_{ccc}$	$\mathcal{L}_{pa}$	*Jac.*	*Dice*	*Acc.*	*Sen.*	*Spe.*	*Pre.*
-	-	-	-	-	$0.536_{\pm 0.143}$	$0.685_{\pm 0.137}$	$0.973_{\pm 0.014}$	$0.687_{\pm 0.148}$	$0.986_{\pm 0.008}$	$0.697_{\pm 0.143}$
✓	-	-	-	-	$0.552_{\pm 0.124}$	$0.703_{\pm 0.112}$	$0.973_{\pm 0.013}$	$0.729_{\pm 0.112}$	$0.984_{\pm 0.009}$	$0.689_{\pm 0.129}$
✓	✓	-	-	-	$0.553_{\pm 0.107}$	$0.706_{\pm 0.094}$	$0.974_{\pm 0.012}$	$0.702_{\pm 0.091}$	$0.987_{\pm 0.009}$	$0.723_{\pm 0.129}$
✓	✓	✓	-	-	$0.561_{\pm 0.106}$	$0.712_{\pm 0.093}$	$0.974_{\pm 0.013}$	$0.728_{\pm 0.085}$	$0.985_{\pm 0.010}$	$0.711_{\pm 0.131}$
✓	✓	✓	✓	-	$0.568_{\pm 0.104}$	$0.718_{\pm 0.091}$	$0.974_{\pm 0.013}$	$0.752_{\pm 0.078}$	$0.984_{\pm 0.011}$	$0.700_{\pm 0.128}$
✓	✓	✓	✓	✓	$0.578_{\pm 0.102}$	$0.727_{\pm 0.087}$	$0.974_{\pm 0.012}$	$0.772_{\pm 0.077}$	$0.983_{\pm 0.010}$	$0.698_{\pm 0.121}$

3.3 Ablation Study of Adaptive Representation Alignment

We conduct the ablation study to evaluate the impact on segmentation quality when training with real XA and different loss functions. The results are shown in Table 2. Here, we set the Unet trained by our simulated XA images as the baseline, and its results are shown in the 1st row in Table 2. Notably, our baseline outperforms self-supervised methods, suggesting that our proposed XA simulation module facilitates deep learning-based vessel segmentation without manual annotations. Meanwhile, we observe that simply forwarding the model on target XA images still yields improvements in segmentation performance on the target XA domain. Subsequent rows in Table 2 demonstrate the incremental impact of incorporating different components, represented by various loss functions. Our findings highlight the significant contribution of our method in enhancing vessel segmentation quality, indicating its efficacy in aligning representations between simulated and real XA domains. These results underscore the importance of adaptive representation alignment techniques in improving the generalization of deep learning models across different data distributions.

4 Conclusion

In this paper, we introduce XA-Sim2Real, a novel adaptive representation learning framework tailored for vessel segmentation in XA images. By addressing the challenges posed by limited annotations and data heterogeneity across institutions, our method achieves promising segmentation performance without requiring manual annotations. Leveraging DRVR generation and adaptive representation alignment techniques, XA-Sim2Real effectively bridges the gap between simulated and real XA data, ensuring robust and accurate segmentation results. Extensive experimental evaluations on public and in-house datasets demonstrate the superiority of our approach over state-of-the-art self-supervised methods, reaffirming its applicability and effectiveness in clinical settings.

Acknowledgments. The project was supported by the Bavarian State Ministry of Science and Arts within the framework of the "Digitaler Herz-OP" project under the grant number 1530/891 02 and the China Scholarship Council (File No.202004910390). We also thank BrainLab AG for their partial support.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ahn, E., Feng, D., Kim, J.: A spatial guided self-supervised clustering network for medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. pp. 379–388. Springer (2021)
2. Gharleghi, R., Adikari, D., Ellenberger, K., Webster, M., Ellis, C., Sowmya, A., Ooi, S., Beier, S.: Annotated computed tomography coronary angiogram images and associated data of normal and diseased arteries. Sci Data **10**(1), 128 (2023)
3. Hubbell, J.H., Seltzer, S.M.: Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients 1 kev to 20 mev for elements z= 1 to 92 and 48 additional substances of dosimetric interest. Tech. rep., National Inst. of Standards and Technology-PL, Gaithersburg, MD (United ... (1995)
4. Kim, B., Oh, Y., Ye, J.C.: Diffusion adversarial representation learning for self-supervised vessel segmentation. arXiv preprint arXiv:2209.14566 (2022)
5. Ma, Y., Hua, Y., Deng, H., Song, T., Wang, H., Xue, Z., Cao, H., Ma, R., Guan, H.: Self-supervised vessel segmentation via adversarial learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7536–7545 (2021)
6. Melas-Kyriazi, L., Rupprecht, C., Laina, I., Vedaldi, A.: Deep spectral methods: A surprisingly strong baseline for unsupervised semantic segmentation and localization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8364–8375 (2022)
7. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. pp. 319–345. Springer (2020)

8. Punnoose, J., Xu, J., Sisniega, A., Zbijewski, W., Siewerdsen, J.: spektr 3.0—a computational tool for x-ray spectrum modeling and analysis. Medical physics **43**(8Part1), 4711–4717 (2016)
9. Ronneberger, O., Fischer, P., Brox, T.: 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. pp. 234–241. Springer (2015)
10. Shi, T., Ding, X., Zhang, L., Yang, X.: Freecos: self-supervised learning from fractals and unlabeled images for curvilinear object segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 876–886 (2023)
11. Soares, J.V., Leandro, J.J., Cesar, R.M., Jelinek, H.F., Cree, M.J.: Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. IEEE Transactions on medical Imaging **25**(9), 1214–1222 (2006)
12. Yang, Y., Soatto, S.: Fda: Fourier domain adaptation for semantic segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4085–4095 (2020)
13. Zamir, M.: Arterial branching within the confines of fractal l-system formalism. The Journal of general physiology **118**(3), 267–276 (2001)
14. 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: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 515–524. Springer (2023)
15. Zhang, X., Sun, K., Wu, D., Xiong, X., Liu, J., Yao, L., Li, S., Wang, Y., Feng, J., Shen, D.: An anatomy-and topology-preserving framework for coronary artery segmentation. IEEE Transactions on Medical Imaging (2023)
16. Zhang, Z., Zhang, B., Azampour, M.F., Faghihroohi, S., Tomczak, A., Schunkert, H., Navab, N.: Self-supervised vessel segmentation from x-ray images using digitally reconstructed radiographs. In: BVM Workshop. pp. 220–225. Springer (2024)