VeCAS: Vessel-Focused Contrast-Free Angiogram Synthesis for Vascular Interventions

- Supplementary Material -

De-Xing Huang, Chen-Yu Wang, Hao Liang, Xiao-Hu Zhou, Mei-Jiang Gui, Tian-Yu Xiang, Qin-Yi Zhang, Chen Wang, Xiao-Liang Xie, Shi-Qi Liu, Ming-Yuan Liu, Zhen-Chang Wang, and Zeng-Guang Hou

S1. Detailed Scoring Rules for Physician Assessment

To evaluate the visual realism and potential procedural utility of the synthesized X-ray angiograms, experienced physicians were asked to compare each synthesized image with its corresponding real angiogram. The assessment used three criteria: vascular morphological realism (VMR), contrast opacification realism (COR), and procedural usability (PU). Each criterion was scored using a five-point Likert scale. The detailed scoring rules are summarized in Table S1.

S2. Additional Details of the Phantom Experiment

Phantom Fabrication. Two patient-specific vascular phantoms were fabricated for the phantom experiment. We first selected one case from each of two patients who underwent lower-limb vascular interventions. For each case, the vascular structures were annotated on the corresponding X-ray angiogram by an experienced physician. The annotated vascular structures were then converted into physical phantom models using 3D printing. Each phantom had an in-plane size of $25\ \mathrm{cm} \times 25\ \mathrm{cm}$ and a height of $1\ \mathrm{cm}$. The in-plane dimensions were chosen to preserve the spatial scale and vascular morphology of the original X-ray angiograms as closely as possible, ensuring that the printed vascular paths were consistent with the patient-specific vessel trajectories. Both phantoms were printed using polylactic acid (PLA), which provided sufficient structural rigidity for repeated guidewire navigation experiments.

Real-Time Image Display and Guidewire Overlay. The workflow for real-time image display and guidewire overlay in the phantom experiment is illustrated in Fig. S1. The workflow consisted of four steps. First, a live video stream was acquired from an overhead camera to capture the phantom and guidewire motion during navigation. Second, the phantom was automatically detected as the region of interest (ROI) based on the rectangular frame surrounding the phantom. The detected ROI was then perspective-corrected to obtain a square top-view image. After ROI correction, a guidewire-free background frame was captured and used as the reference for subsequent guidewire extraction. Third, the selected non-contrast X-ray image was loaded into the local interface as the guidance background. During navigation, each incoming video frame was transformed using the same ROI correction, and the guidewire was segmented by comparing the current frame with the guidewire-free background frame. The extracted guidewire mask was then overlaid on the non-contrast X-ray image in real time. Fourth, for VeCAS-guided navigation, the same non-contrast X-ray image was sent to a remote server running VeCAS for angiogram synthesis. After the synthesized angiogram was returned to the local workstation, the live guidewire mask was overlaid on it.

Guidance Conditions. Three guidance conditions were evaluated in the phantom navigation experiment: non-contrast guidance, VeCAS guidance, and contrast guidance, as shown in Fig. S2. In the non-contrast guidance condition, the selected non-contrast X-ray image was displayed as the background image, and the live guidewire mask extracted from the camera stream was overlaid on this image in real time. This condition simulated guidewire navigation without vascular visualization. In the VeCAS guidance condition, the synthesized angiogram was displayed as the background image, and the same live guidewire mask was overlaid on the angiogram in real time.

This work was supported in part by the National Natural Science Foundation of China under Grant 62373351, Grant 62503474, Grant 82327801, Grant 62303463, Grant 52475038, in part by the Chinese Academy of Sciences Project for Research Instrument Development under Grant No. PTYQ2026YZ0072, Young Scientists in Basic Research under Grant No. YSBR-104, in part by the Beijing Natural Science Foundation under Grant F252068, Grant 4254107, in part by Beijing Nova Program under Grant 20250484813, in part by China Postdoctoral Science Foundation under Grant 2024M763535, in part by the Postdoctoral Fellowship Program of CPSF under Grant GZC20251170, and in part by Xiaomi Young Talents Program/Xiaomi Foundation.

†: D.-X. Huang, C.-Y. Wang, and H. Liang contributed equally to this work (email: huangdexing2022@ia.ac.cn, chenyuwang.cs@gmail.com, lh250312@163.com).

∗ Corresponding authors: Xiao-Hu Zhou, Ming-Yuan Liu, and Zeng-Guang Hou (email: {xiaohu.zhou, zengguang.hou}@ia.ac.cn, dr.mingyuanliu@pku.edu.cn).

D.-X. Huang, C.-Y. Wang, X.-H. Zhou, M.-J. Gui, T.-Y. Xiang, Q.-Y. Zhang, C. Wang, X.-L. Xie, S.-Q. Liu, and Z.-G. Hou are with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.

D.-X. Huang, X.-H. Zhou, T.-Y. Xiang, Q.-Y. Zhang, C. Wang, and Z.-G. Hou are also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China.

C.-Y. Wang is also with the School of Advanced Interdisciplinary Sciences, University of Chinese Academy of Sciences, Beijing 101408, China.

H. Liang, M.-Y. Liu, and Z.-C. Wang are with the Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China.

TABLE S1
DETAILED SCORING RULES USED IN THE PHYSICIAN ASSESSMENT OF SYNTHESIZED X-RAY ANGIOGRAMS.

Score	Description
Criterion I: Vascular Morphological Realism (VMR)	
5	The synthesized angiogram closely matches the reference angiogram in terms of vessel locations and trajectories, vessel continuity, boundary definition, and depiction of major branches. The vascular morphology appears highly realistic.
4	The synthesized angiogram largely matches the reference angiogram in terms of vessel locations and trajectories, vessel continuity, boundary definition, and depiction of major branches. Only minor localized differences are observed, and these do not substantially affect the overall interpretation.
3	The major vessels are generally visible, and their overall trajectories are largely reasonable. However, some inconsistencies remain in vessel continuity, boundary accuracy, or depiction of major branches.
2	Some vascular structures are visible, but evident deviations in vessel locations or trajectories, vessel continuity, boundary definition, or depiction of major branches result in substantial differences from the reference angiogram.
1	The vascular morphology is clearly incorrect. Vessel locations or trajectories, vessel continuity, boundary definition, or depiction of major branches differ severely from the reference angiogram, and obvious artifacts or hallucinated structures are present.
Criterion II: Contrast Opacification Realism (COR)	
5	The synthesized contrast appearance is highly natural. The intravascular intensity distribution, contrast variation, and edge transitions closely resemble those of the reference angiogram.
4	The synthesized contrast appearance is relatively natural. The intravascular intensity distribution, contrast variation, and edge transitions are generally realistic, with only minor unnatural features.
3	The synthesized contrast appearance is generally acceptable, but localized unnatural intensity variations, uneven intravascular brightness, or unrealistic edge transitions are still observed.
2	The synthesized contrast appearance is artificial, with obvious over-enhancement, under-enhancement, smearing, block-like artifacts, or unnatural contrast variations.
1	The synthesized contrast appearance is clearly unrealistic. The intravascular intensity distribution, contrast variation, or edge transitions are severely abnormal, making the image difficult to regard as a realistic angiogram.
Criterion III: Procedural Usability (PU)	
5	The synthesized angiogram has high procedural usability. It clearly provides useful information about vessel trajectories and major branches for potential procedural guidance or planning.
4	The synthesized angiogram has good procedural usability. It provides useful information about vessel trajectories and major branches and may support procedural guidance or planning, although key clinical decisions still require confirmation with real angiography.
3	The synthesized angiogram has some procedural utility and can aid the interpretation of part of the vascular anatomy, but real angiography is still required for reliable confirmation.
2	The synthesized angiogram has limited procedural utility and provides only limited information about vessel trajectories and major branches for potential procedural guidance, with a relatively high risk of misinterpretation.
1	The synthesized angiogram has no procedural utility and may lead to misinterpretation of vessel trajectories, major branches, or lesion-related regions. It is unsuitable for procedural guidance or planning.

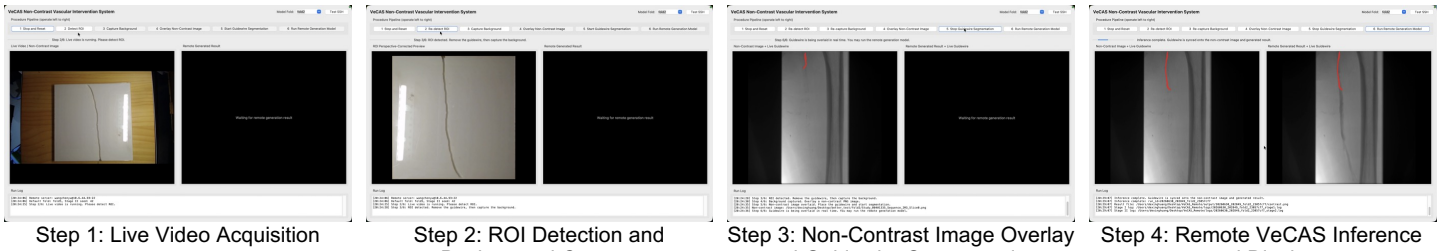

Fig. S1. **Workflow of the Real-Time Image Display and Guidewire Overlay Interface Used in the Phantom Experiment.**

This condition was used to evaluate whether the synthesized angiogram could provide useful guidance during navigation. In the contrast guidance condition, the real X-ray angiogram corresponding to the selected case was displayed as the background image, and the same live guidewire mask was overlaid on the real angiogram in real time. This condition served as the conventional guidance reference.

Protocol. Three operators were invited to perform guidewire navigation on two patient-specific vascular phantoms. For each phantom, each operator was asked to navigate the guidewire tip from the inlet region to a predefined distal target point, as illustrated in Fig. S4. For each operator, phantom, and guidance condition, the navigation task was repeated five times. Before each trial, the guidewire was reinitialized at the predefined inlet region of the phantom. A trial started when the operator issued the first control command and ended when the guidewire tip reached the predefined target region. During each trial, the operator manipulated the guidewire using an Xbox game controller while observing the corresponding guidance image with the live guidewire overlay.

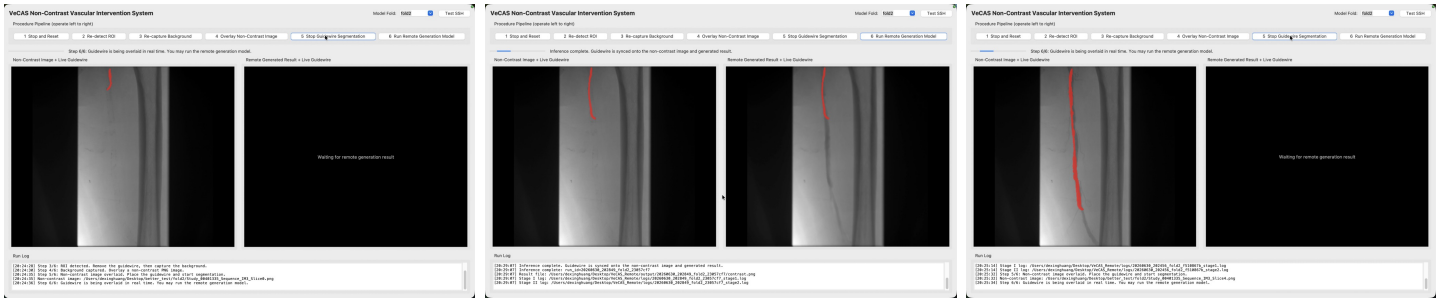

Fig. S2. **Interface Views under the Three Guidance Conditions in the Phantom Navigation Experiment.**

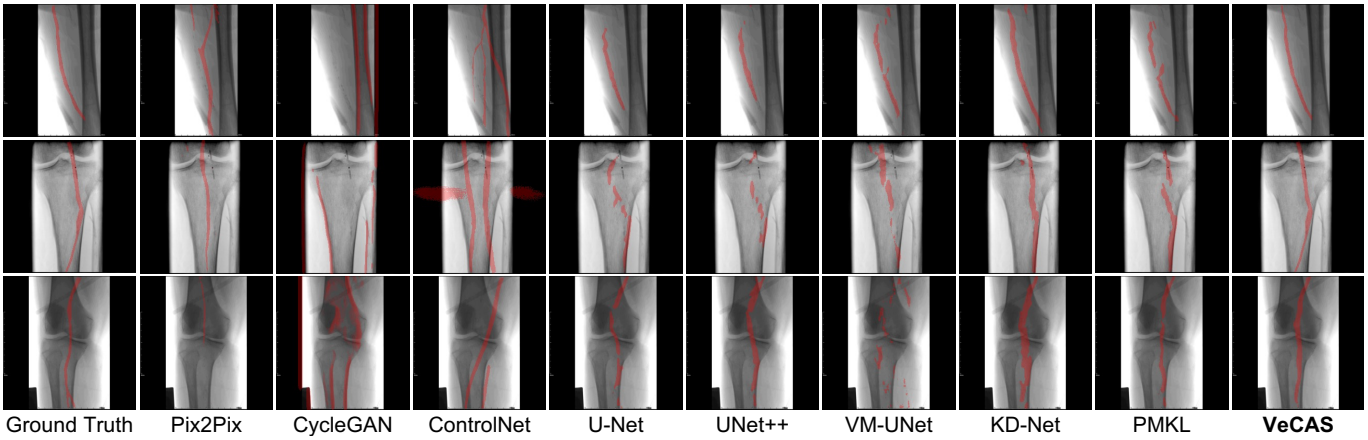

Fig. S3. **Qualitative Results of Stage I.** Compared with the baseline methods, VeCAS achieves more accurate vascular localization, with fewer false positives and better preservation of vascular continuity.

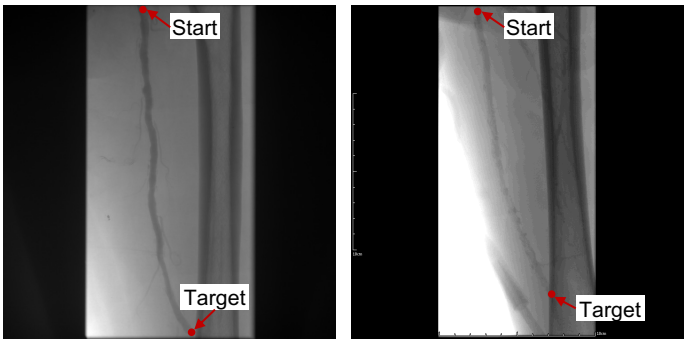

Fig. S4. **Navigation Task Definition for Two Vascular Phantoms.** The guidewire was initialized at the starting point and advanced toward the predefined distal target point. The same task setup was used under non-contrast, VeCAS, and contrast guidance.

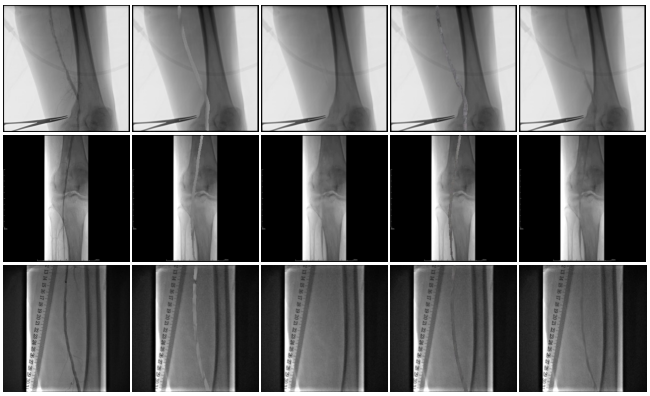

Fig. S5. **Qualitative Ablation Results of the Vessel-Focused Design.** Compared with the ablated variants, VeCAS synthesizes clearer and more continuous vessels.

S3. Additional Results

Qualitative Comparisons of Stage I. Fig. S3 provides qualitative comparisons of different methods for vascular structure localization. Generative models produce unstable predictions with poor anatomical consistency. Pix2Pix [1] and CycleGAN [2] fail to recover most vascular regions, while ControlNet [3] introduces false-positive responses in some cases. Segmentation models and cross-modality interaction baselines can generally localize the major vessels, but their predictions often suffer from discontinuous centerlines and inaccurate boundaries. In contrast, VeCAS produces vessel masks that are more consistent with the ground truth, with fewer false positives and better preservation of vascular continuity.

Qualitative Comparisons of the Vessel-Focused Design. Fig. S5 illustrates the effect of the vessel-focused design. When both the vessel-focused loss and vessel-focused sampling are removed, the synthesized vessels remain visible, but the contrast enhancement appears highly unnatural, with unrealistic intensity distributions and poor consistency with

TABLE S2

ABLATION STUDY OF MULTI-SCALE LATENT PREDICTION. FILLED CIRCLES INDICATE THE SCALES AT WHICH LATENT PREDICTION IS APPLIED. THE BEST AND SECOND-BEST RESULTS ARE HIGHLIGHTED IN **BOLD** AND UNDERLINED, RESPECTIVELY.

Scale					DSC (%) ↑	clDice (%) ↑
1	2	3	4	5		
○	○	○	○	●	51.79 ± 3.34	59.87 ± 4.65
○	○	○	●	●	51.80 ± 3.09	59.87 ± 4.28
○	○	●	●	●	$\mathbf{52.56 \pm 3.32}$	$\mathbf{60.62 \pm 4.42}$
○	●	●	●	●	51.66 ± 3.07	59.63 ± 4.08
●	●	●	●	●	$\underline{52.02 \pm 3.88}$	$\underline{60.08 \pm 4.42}$

TABLE S3

ABLATION STUDY ON THE TRAINING STRATEGIES OF $\mathcal{E}_{\mathbf{A}}$ AND $\mathcal{D}_{\mathbf{A}}$. THE BEST AND SECOND-BEST RESULTS ARE HIGHLIGHTED IN **BOLD** AND UNDERLINED, RESPECTIVELY.

Pre-train	$\mathcal{E}_{\mathrm{A}}$	$\mathcal{D}_{\mathrm{A}}$	DSC (%) ↑	clDice (%) ↑
✗	🔥	🔥	48.38 ± 3.96	56.08 ± 4.86
✓	🔥	🔥	51.75 ± 3.49	60.07 ± 4.65
✓	❄	🔥	$\underline{51.98 \pm 3.65}$	$\mathbf{60.93 \pm 5.36}$
✓	❄	❄	$\mathbf{52.56 \pm 3.32}$	$\underline{60.62 \pm 4.42}$

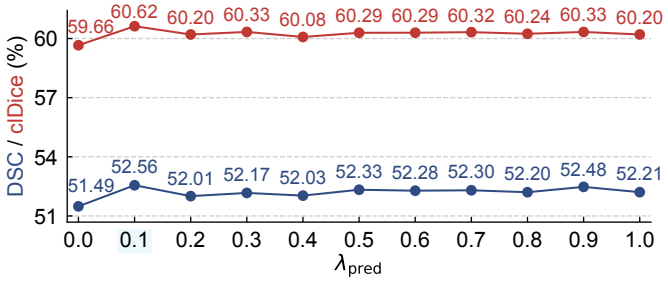

Fig. S6. **Ablation Study on the Hyperparameter λ_{pred}.** VeCAS maintains stable performance across a broad range of λ_{pred} values, demonstrating its robustness to this hyperparameter.

TABLE S4

ABLATION STUDY ON THE NUMBER OF SAMPLING STEPS. THE BEST AND SECOND-BEST RESULTS ARE HIGHLIGHTED IN **BOLD** AND UNDERLINED, RESPECTIVELY.

Steps	DSC (%) ↑	clDice (%) ↑	PSNR (dB) ↑	SSIM ↑
10	36.31 ± 6.18	40.74 ± 7.51	36.63 ± 0.84	0.9530 ± 0.0180
20	40.04 ± 5.33	45.19 ± 6.33	36.77 ± 0.87	0.9536 ± 0.0183
50	$\underline{41.29 \pm 4.47}$	$\underline{46.48 \pm 5.30}$	$\mathbf{36.78 \pm 0.87}$	$\mathbf{0.9538 \pm 0.0180}$
100	$\mathbf{42.27 \pm 4.99}$	$\mathbf{47.70 \pm 6.29}$	$\underline{36.78 \pm 0.91}$	$\underline{0.9537 \pm 0.0183}$
200	23.21 ± 5.03	25.47 ± 6.03	35.48 ± 0.72	0.9480 ± 0.0178

the surrounding background. When the vessel-focused loss is removed, the vascular structures become weakened or even invisible, suggesting that this loss is essential for maintaining vessel visibility. In contrast, removing only the vessel-focused sampling still allows the model to recover visible vessels. However, the contrast opacification remains less natural than that of the full model, although it is more natural than that produced by the variant containing neither component. By jointly using the vessel-focused loss and vessel-focused sampling, VeCAS produces clearer vascular structures and more natural contrast opacification.

Multi-Scale Prediction. Table S2 reports the results of applying latent feature prediction at different scales. Using only the deepest feature at scale 5 achieves 51.79% DSC and 59.87% clDice, whereas applying the predictor to scales 3, 4, and 5 further improves performance. This indicates that intermediate high-level features provide useful structural information for vascular localization. However, extending latent prediction to shallower scales does not yield additional gains, possibly because low-level features are more sensitive to modality-specific appearance differences between non-contrast X-ray images and X-ray angiograms.

Training Strategies for $\mathcal{E}_{\mathrm{A}}(\cdot)$ and $\mathcal{D}_{\mathrm{A}}(\cdot)$. Table S3 compares different training strategies. In the setting without pre-training on X-ray angiograms, $\mathcal{E}_{\mathrm{A}}(\cdot)$ and $\mathcal{D}_{\mathrm{A}}(\cdot)$ are jointly trained from scratch together with the non-contrast branch. To ensure a fair comparison, an additional segmentation loss is applied to the angiogram branch. However, this setting still yields limited performance, suggesting that jointly learning the angiogram-domain representation from scratch fails to provide a sufficiently stable and vessel-sensitive latent space for distillation. Pre-training $\mathcal{E}_{\mathrm{A}}(\cdot)$ and $\mathcal{D}_{\mathrm{A}}(\cdot)$ clearly improves performance, demonstrating the importance of constructing a reliable angiogram-domain representation before distillation. Further freezing the pre-trained $\mathcal{E}_{\mathrm{A}}(\cdot)$ while keeping $\mathcal{D}_{\mathrm{A}}(\cdot)$ trainable yields the highest clDice, indicating that a fixed angiogram-domain latent space facilitates stable latent feature prediction from non-contrast X-ray images. Our default setting achieves the best DSC with comparable clDice, providing more balanced overall performance.

Hyperparameter Analysis. (1) λ_{pred}: As shown in Fig. S6, the performance remains stable over a broad range of λ_{pred} values. Notably, all non-zero settings consistently outperform the baseline with $\lambda_{\text{pred}} = 0.0$, indicating the effectiveness of the latent prediction loss $\mathcal{L}_{\text{pred}}$. The best performance is achieved at $\lambda_{\text{pred}} = 0.1$, while other non-zero values yield comparable results, suggesting that the proposed method is not sensitive to this hyperparameter. **(2) Sampling Steps:** As shown in Table S4, increasing the number of sampling steps from 10 to 100 consistently improves DSC and clDice, indicating that a sufficient denoising trajectory is necessary for recovering reliable vascular structures. The best performance is achieved with 100 sampling steps, which yields the highest DSC and clDice while maintaining comparable PSNR and SSIM. However, further increasing the number of steps to 200 leads to a noticeable decrease in DSC and clDice, despite relatively stable PSNR and SSIM, suggesting that excessive sampling mainly affects vascular structural fidelity rather than global image quality. This may be because repeated fusion between the synthesized vascular region and the non-contrast background can accumulate small errors around thin vessels and mask boundaries, resulting in vessel discontinuities or slight spatial shifts. Therefore, we set the default number of sampling steps to 100 in all experiments.

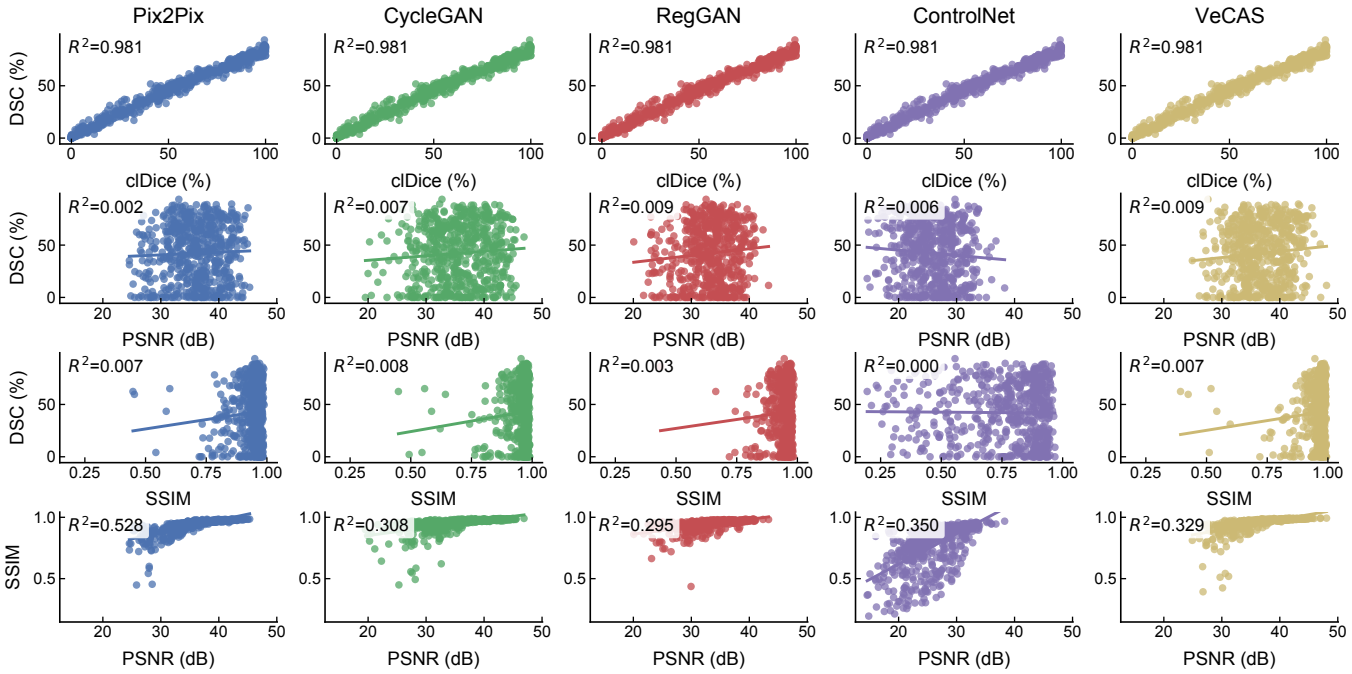

Fig. S7. **Correlation Analysis of Quantitative Metrics.** DSC and clDice show a strong correlation, indicating their consistency in assessing vascular structural fidelity. In contrast, PSNR and SSIM exhibit weak correlations with DSC, suggesting that whole-image fidelity metrics alone may not reliably reflect vascular structural fidelity and should be interpreted together with vessel-specific metrics.

S4. Discussion

Clinical Role of VeCAS. VeCAS is designed as an auxiliary computational visualization tool for vascular interventions, rather than a replacement for X-ray angiography. The synthesized angiograms aim to provide roadmap-like visualization of likely vascular trajectories from non-contrast X-ray images. Therefore, they should be interpreted as angiogram-like guidance images rather than direct evidence of true lumen patency, stenosis severity, occlusion status, or contrast-flow dynamics. This distinction is important because vascular structures may be weakly visible or even invisible in non-contrast X-ray images. As a result, VeCAS mainly learns image-conditioned anatomical priors from the training data and can generate vascular structures that are anatomically plausible. However, for cases with severe stenosis, complete occlusion, abnormal branching, collateral circulation, or other lesion-specific abnormalities, the synthesized angiogram may not faithfully reflect the true pathological condition of the patient. Therefore, VeCAS should be used only as a supplementary visual aid and should be interpreted together with conventional fluoroscopy, intermittent contrast injections, and physician judgment.

Whole-Image Metrics. PSNR and SSIM are commonly used to evaluate global image fidelity. However, because vascular regions are thin and occupy only a small portion of the image, whole-image metrics may be dominated by the non-vascular background and may be insensitive to local vascular errors, such as missing vessels, discontinuities, or inaccurate branches. As shown in Fig. S7, DSC and clDice are highly correlated, indicating that they consistently reflect vascular structural fidelity. In contrast, PSNR and SSIM show weak correlations with DSC, suggesting that high whole-image similarity does not necessarily imply accurate vascular synthesis. The limited correlation between PSNR and SSIM further indicates that they capture different aspects of image fidelity. Therefore, in contrast-free angiogram synthesis, PSNR and SSIM should not be interpreted in isolation but should be considered together with vessel-specific metrics.

Limitations and Future Work. Despite the encouraging results, this study has several limitations. First, the dataset was collected from a single medical center and focused on lower-limb vascular interventions. Although patient-level five-fold cross-validation was adopted, the generalizability of VeCAS to different institutions, imaging systems, acquisition protocols, and patient populations still requires multicenter validation. Second, the final synthesis quality depends on the accuracy of vascular structure localization in Stage I. Missed vessels, false-positive branches, boundary inaccuracies, or spatial misalignment in the predicted mask may be propagated to Stage II, leading to incomplete or misplaced vascular synthesis. Further improving the robustness of vessel localization remains an important direction for future work. Third, VeCAS focuses on synthesizing anatomically plausible vascular structures within the corresponding anatomical region, rather than inferring or simulating disease-specific abnormalities. Therefore, the synthesized angiograms should not be interpreted as direct evidence of patient-specific lumen patency or lesion severity. Future work will address this limitation by incorporating explicit pathological annotations and disease-aware modeling, enabling the synthesized angiograms to better reflect patient-specific vascular lesions.

References

[1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. CVPR*, 2017, pp. 1125–1134.

[2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. ICCV*, 2017, pp. 2223–2232.
[3] L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in *Proc. ICCV*, 2023, pp. 3836–3847.