This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Prior-guided Hierarchical Vector Quantized-VAE for Digital Subtraction Angiography Generation

Yunbi Liu[1], Zheyun Yang[1], Enqi Tang[1], Shiyu Li[1], Yifei Tian[1], Shiteng Suo[2*], and Qingshan Liu[1*]

[1] School of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China
[2] Department of Radiology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai,200127, China

Abstract. Digital Subtraction Angiography (DSA) is the clinical gold standard for vascular imaging but is frequently degraded by motion artifacts due to patient physiological movement. While recent deep learning methods bypass traditional subtraction by learning a direct mapping from contrast images to vascular images, they often struggle with the complex anatomical structures in contrast images, leading to incomplete and discontinuous vascular reconstructions. Moreover, the inherent motion artifacts in the training data prevent these models from fully eliminating artifacts in the generated images. To address these challenges, we propose a Prior-guided Hierarchical Vector Quantized Variational Autoencoder (PH-VQVAE). ***First***, our model features a dual-branch encoder that extracts features from both the pre-contrast mask and contrast images at multiple scales. We introduce a window-based cross-attention mechanism at each level to efficiently capture local contrast differences and fuse anatomical priors, enabling precise "soft subtraction" robust to local misalignments. ***Second***, we abandon the continuous latent representations used in conventional generative models in favor of a hierarchical discrete modeling approach. By mapping multi-scale features to learned discrete codebooks, our method effectively filters out background noise and motion artifacts while preserving the topological integrity of vessels. Experimental results demonstrate that our approach reconstructs the vessels well and achieves superior reconstruction accuracy compared to the state-of-the-art generative models, offering a promising tool for safer and more efficient interventional imaging.

Keywords: Digital Subtraction Angiography · Motion Artifacts · Vector Quantized VAE.

1 Introduction

Digital Subtraction Angiography (DSA) plays a pivotal role in interventional radiology by providing clear visualization of the vascular anatomy, essential for

* Corresponding author(s).
Shiteng Suo (shtsuo@gmail.com) and Qingshan Liu (qsliu@njupt.edu.cn)

diagnosing stenosis, aneurysms, and guiding catheter placement [12, 13]. The process involves subtracting a pre-contrast "mask" image from subsequent contrast-enhanced "live" frames [9, 11]. However, patient involuntary movements—such as breathing, heartbeat, or swallowing—often cause spatial misalignment between the mask and the live frames. These misalignments result in significant motion artifacts, obscuring the target vessels and potentially leading to "junk films" that necessitate repeat procedures, thereby increasing the patient's radiation dose and contrast agent toxicity [8].

To mitigate these issues, researchers have traditionally explored image registration techniques; however, non-rigid registration is computationally intensive and often fails in complex anatomical regions [15, 10, 17, 3]. Recently, deep learning has emerged as a powerful alternative, aiming to learn the direct mapping from raw angiography sequences to vessel-only images [1, 16, 19, 14, 7, 6, 20]. While promising, existing deep learning approaches suffer from two limitations. **First, most models rely on continuous latent spaces (e.g., U-Nets, GANs, or Diffusion models) to encode image features.** In a continuous manifold, the boundary between the sparse, low-contrast vascular signal and the stochastic background noise is often ambiguous. Prior studies have reported that continuous representations can be vulnerable to error accumulation [2], while discrete learning (i.e., vector-quantized models) can prevent input interference, such as noise, artifacts, and field bias [4]. **Second, current methods often utilize only the "live" contrast frame, neglecting the explicit use of the pre-contrast mask image.** Without this anatomical prior, models are forced to "guess" vessel structures from scratch. This lack of reference makes it difficult to distinguish actual contrast updates from static bone structures, often resulting in discontinuous or fragmented vessel reconstructions, especially when the contrast concentration is low.

To address these dual challenges, we introduce a Prior-guided Hierarchical Vector Quantized Variational Autoencoder (PH-VQVAE) for DSA image generation. Our primary motivation is to leverage the discrete latent space of VQ-VAE to filter out stochastic noise and motion artifacts that continuous models fail to eliminate. Crucially, we upgrade the standard VQ-VAE to a hierarchical architecture, ensuring that both large vascular structures and tiny collateral vessels are accurately reconstructed. More importantly, to resolve the issue of vessel discontinuity, we propose a dual-branch encoder architecture that explicitly incorporates the pre-contrast mask as an anatomical prior. Unlike single-branch methods, this design allows the network to learn a "soft subtraction" strategy in the feature space: by comparing the contrast frame against the mask prior, the model can robustly identify invariant background structures and isolate the flowing contrast agent, thereby significantly enhancing the connectivity and completeness of the reconstructed vasculars. To fuse these features effectively, we introduce a window-based cross-attention mechanism at each level. Unlike global attention, window-based attention focuses on local neighborhoods, allowing the model to efficiently capture fine-grained contrast differences and correct local misalignments without the computational burden of global interactions.

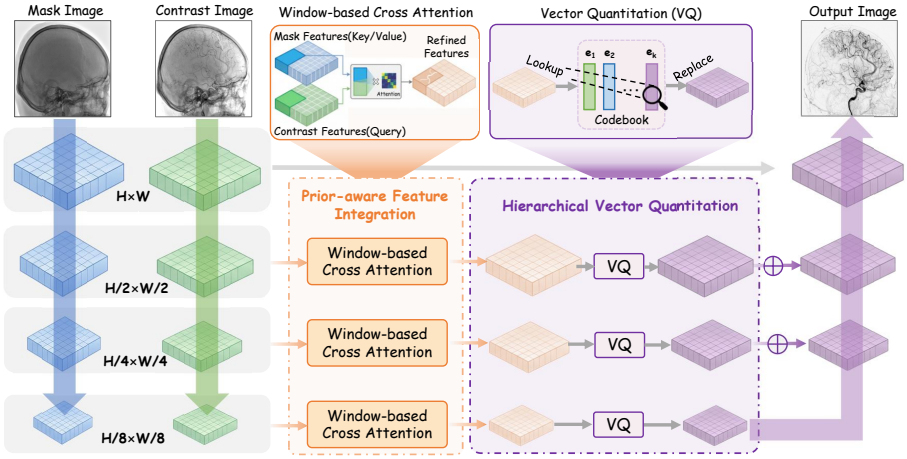

Fig. 1. Illustration of the proposed PH-VQVAE method for DSA image generation.

In summary, our main contributions are:

1. We propose, to the best of our knowledge, the first application of hierarchical VQ-VAE for high-fidelity DSA image generation. By leveraging a multi-scale discrete codebook, our framework effectively filters out stochastic motion artifacts while accurately reconstructing both large trunks and tiny collateral vessels, overcoming the limitations of continuous models.
2. We introduce a prior-guided dual-branch architecture with multi-scale window-based cross attention. This mechanism explicitly utilizes the pre-contrast mask as an anatomical prior, performing efficient "soft subtraction" at different resolutions to robustly recover vessels from complex backgrounds.
3. Extensive ablation studies and comparison experiments validate the validity and superiority of our method for generating high-fidelity DSA images, highlighting its potential for safer and more efficient interventional imaging.

2 Method

2.1 Overall Architecture

As illustrated in Fig. 1, our proposed framework is a Prior-guided Hierarchical Vector Quantized Variational Autoencoder (PH-VQVAE) designed to reconstruct high-fidelity DSA images. The architecture begins with a dual-branch encoder that independently extracts multi-scale features from the mask image (I_{mask}) and the contrast image ($I_{contrast}$). To robustly separate vascular signals from complex backgrounds, we introduce a *Prior-aware Feature Integration* module. Instead of simple concatenation, this module employs a window-based cross-attention mechanism at each scale, where mask features serve as keys/values to guide the refinement of contrast features (queries). This process performs a “soft

subtraction" in the feature space, utilizing the anatomical prior to filter out static structures and highlight local contrast differences. Subsequently, the refined features are passed through a *Hierarchical Vector Quantization* bottleneck. Here, continuous features are mapped to discrete codebooks, which act as structural filters to eliminate stochastic noise and motion inconsistencies that lack compact representations in the discrete space. Finally, a decoder aggregates these multi-scale quantized features to reconstruct the DSA image $\hat{Y}$, ensuring both topological integrity and precise vessel connectivity.

2.2 Method Details

Dual-branch Encoder with Prior-aware Feature Integration To effectively separate vascular signals from complex backgrounds, we propose a dual-branch encoder that leverages the mask image as an anatomical prior. Since I_{mask} and $I_{contrast}$ are spatially misaligned, direct concatenation of images or features is suboptimal. We introduce a Prior-aware Feature Integration module to align and refine features. Let E_m and E_c denote the encoders for the mask and contrast images, respectively. They independently extract hierarchical feature maps Z_m and Z_c at multiple scales. Within this module, we employ a window-based cross-attention mechanism to perform "soft subtraction" in the feature space. We treat the contrast features Z_c as the Query (Q), while the mask features Z_m serve as both Key (K) and Value (V). By computing the attention within non-overlapping local windows, the model learns the correlation between the dynamic contrast flow and the static background anatomy:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Z_c W_Q (Z_m W_K)^T}{\sqrt{d_k}}\right)(Z_m W_V) \tag{1}$$

where W_Q, W_K, W_V are learnable projection matrices. This mechanism allows the model to leverage the anatomical prior from the mask to suppress background features in the contrast image, yielding a refined feature representation $Z_{refined}$.

Hierarchical Vector Quantization Continuous latent spaces often encode stochastic motion artifacts and high-frequency noise indistinguishably from the vascular signal. To address this, we impose a discrete constraint using a *Hierarchical Vector Quantization (VQ)* bottleneck. A single-scale quantization at the deepest level (e.g., 1/8 resolution) would be insufficient, as the aggressive downsampling risks obliterating fine vascular structures and distal vessel branches. Therefore, we perform quantization at multiple scales (1/2, 1/4, and 1/8). We define a learnable codebook $C = \{e_k\}_{k=1}^{K} \subset \mathbb{R}^D$ at each scale, where K is the codebook size and D is the embedding dimension. For each spatial vector z_{ij} in the refined feature map $Z_{refined}$, we perform a nearest-neighbor lookup to replace it with the closest code vector e_k from the codebook:

$$z_q^{(i,j)} = e_k, \quad \text{where} \quad k = \arg\min_n \|z_{ij} - e_n\|_2 \tag{2}$$

Here, z_q represents the quantized feature map. This multi-scale discretization process acts as a structural filter: it preserves global vascular topology at coarse scales while retaining intricate vessel details at fine scales.

Decoder Reconstruction and Optimization The decoder G aggregates the multi-scale quantized discrete features z_q to reconstruct the high-fidelity DSA image $\hat{Y} = G(z_q)$. To enable end-to-end training despite the non-differentiable quantization step, we utilize the straight-through estimator, which copies gradients from the decoder input directly to the encoder output during backpropagation. The framework is optimized using a compound loss function comprising a reconstruction loss and a codebook alignment loss. The reconstruction term L_{rec} employs L1 loss to ensure pixel-wise fidelity between the output and the ground truth. To train the codebook, we minimize the vector quantization loss L_{vq}, which forces the encoder outputs to commit to the nearest code vectors and updates the codebook to match the encoder's representation:

$$L_{vq} = \|\text{sg}[Z_{refined}] - z_q\|_2^2 + \beta\|Z_{refined} - \text{sg}[z_q]\|_2^2 \tag{3}$$

where $\text{sg}[\cdot]$ denotes the stop-gradient operator and β is a hyperparameter balancing the commitment loss. The total objective is summarized as $L_{total} = L_{rec} + L_{vq}$.

2.3 Implementation

The proposed framework was implemented using PyTorch and trained on a workstation equipped with a single NVIDIA GeForce RTX 3090 GPU. Input images were resized to 512×512 pixels, with data augmentation techniques—including random scaling, rotation, and cropping—applied to enhance model robustness. The network architecture is configured with a hidden dimension of 128, employing a codebook size of $K = 512$ with an embedding dimension of $D = 128$. To stabilize the discrete latent representation, we utilized an exponential moving average decay of 0.99 and a commitment cost of 0.25, while the multi-head attention mechanism incorporates 8 heads with a gate initialization of -2.0. The model was trained for 200 epochs with a batch size of 2 using the AdamW optimizer, with both the learning rate and weight decay set to 1×10^{-4}. A linear warmup strategy was applied for the first 3 epochs to ensure convergence stability, and the final objective function integrates a vector quantization loss ($\lambda_{vq} = 0.5$) and an entropy regularization term ($\lambda_{entropy} = 0.02$) to encourage codebook diversity. For reproducibility, our source code and pre-trained models will be made publicly available.

3 Experiments

3.1 Experimental Data

The dataset used in this study was retrospectively collected from 211 patients who underwent cerebral DSA examinations at one hospital. To ensure the model

Table 1. Quantitative ablation study results on the test set. **Bold** indicates the best performance.

Method	SSIM ↑	PSNR (dB) ↑	RMAE ↓	LPIPS ↓
Contrast-only	0.880 ± 0.058	28.11 ± 5.38	2.00 ± 0.010	0.181 ± 0.054
Dual-branch concatenation	0.936 ± 0.024	32.33 ± 4.37	1.29 ± 0.006	0.102 ± 0.034
P-VQVAE (1/8)	0.886 ± 0.056	29.66 ± 5.14	1.74 ± 0.008	0.157 ± 0.051
PH-VQVAE (1/4+1/8)	0.911 ± 0.040	31.64 ± 4.71	1.41 ± 0.006	0.125 ± 0.042
PH-VQVAE (Proposed Tri-scale)	0.940 ± 0.022	**33.07 ± 4.12**	**1.17 ± 0.005**	0.095 ± 0.035
PH-VQVAE (Full-scale)	**0.946 ± 0.019**	32.26 ± 4.38	1.21 ± 0.005	**0.085 ± 0.028**

learns robust vessel features without being biased by severe misregistration, our training strategy primarily utilizes image pairs exhibiting mild motion artifacts. Specifically, the training set (derived from 184 patients) contains 2,242 pairs of high-quality data with minimal motion, comprising 1,043 anteroposterior (AP) and 1,199 lateral (LA) projection views. Similarly, the validation set (derived from 11 patients) includes 223 mild artifact pairs (115 AP and 108 LA) to monitor model performance on high-quality data. In contrast, to rigorously assess the model's generalization capability across varying degrees of motion, the independent test set (16 patients) includes a diverse range of severities, encompassing mild (173 pairs), moderate (902 pairs), and severe (170 pairs) motion artifacts. In total, the full dataset consists of 3,710 image pairs, strictly partitioned at the patient level to prevent data leakage.

3.2 Ablation Study

Experimental setup. We conducted a systematic ablation study to isolate the contributions of two core components: the Mask Prior Integration strategy and the Hierarchical Vector Quantization (VQ) configuration. To validate the necessity of the anatomical prior, we first compared our full model against a Contrast-only baseline, which removes the dual-branch encoder and cross-attention module, relying solely on self-extraction of features from the contrast image. Subsequently, to demonstrate the superiority of our Prior-aware Feature Integration, we evaluated a Dual-branch Concatenation variant, where features are extracted independently from the mask and contrast images via a dual-branch encoder, but instead of using cross-attention, the features from both branches are simply concatenated at each scale before passing through the hierarchical VQ and decoder. Furthermore, to determine the optimal quantization granularity for preserving vascular details while suppressing noise, we analyzed the impact of multi-scale VQ. We constructed four variants with increasing quantization depth: (1) a Single-scale model applying quantization only at the deepest bottleneck layer (1/8 resolution); (2) a Dual-scale model utilizing quantization at 1/8 and 1/4 resolutions; (3) our Proposed Tri-scale model, which applies quantization at 1/8, 1/4, and 1/2 resolutions; and (4) a Full-scale model that extends quantization to the original resolution (1/1).

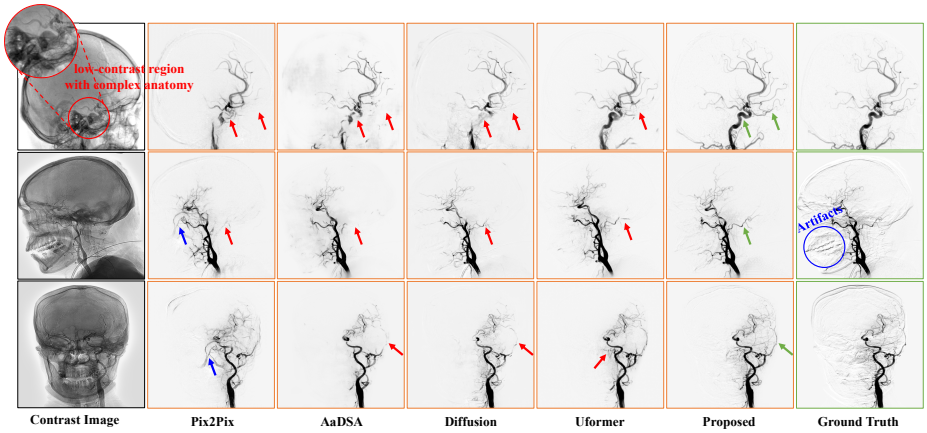

Fig. 2. Visual results of different comparison methods. Red and blue circles highlights the low-contrast region with complex anatomy in the contrast image and the artifacts in the ground truth DSA images, respectively. Red, blue and green arrows indicate loss of vessels, hallucinatory artifacts, and well-reconstructed vessels, respectively.

Table 2. Quantitative comparison with state-of-the-art methods on the test set.

Method	**SSIM** ↑	**PSNR** (dB) ↑	**RMAE** ↓	**LPIPS** ↓
Pix2Pix(Yonezawa et al., 2022)	0.860 ± 0.068	26.46 ± 6.42	2.67 ± 0.015	0.213 ± 0.068
AaDSA (Liu et al., 2025)	0.898 ± 0.043	27.35 ± 4.90	2.26 ± 0.011	0.235 ± 0.063
Diffusion (Ho et al., 2020)	0.893 ± 0.051	28.06 ± 5.53	2.05 ± 0.010	0.262 ± 0.055
Uformer (Wang et al., 2022)	0.906 ± 0.04	29.28 ± 4.92	1.76 ± 0.009	0.147 ± 0.053
PH-VQVAE (Ours)	**0.940 ± 0.022**	**33.07 ± 4.12**	**1.17 ± 0.005**	**0.095 ± 0.035**

Quantitative Analysis. The quantitative results of the ablation study are summarized in Table 1. As summarized in Table 1, the ablation study validates the effectiveness of our core components. The Contrast-only baseline exhibits the poorest performance, confirming the necessity of anatomical guidance. While the Dual-branch concatenation strategy improves structural similarity by incorporating mask features, its direct concatenation leads to suboptimal PSNR compared to our attentive integration. Regarding the quantization scale, performance generally improves with resolution. The Single-scale (1/8) and Dual-scale (1/4+1/8) models fail to capture fine vessel terminals. Our proposed Tri-scale (1/2+1/4+1/8) configuration achieves the optimal trade-off, reaching the highest PSNR (33.07 dB) and high SSIM (0.940). Although the Full-scale model yields marginally better SSIM, its PSNR drops slightly likely due to the preservation of high-frequency noise at the finest scale. Consequently, we selected the Tri-scale configuration as the optimal solution, striking the best balance between reconstruction fidelity and model efficiency.

3.3 Comparative Study

Experimental setup. To comprehensively evaluate the proposed Prior-guided Hierarchical VQ-VAE (PH-VQVAE), we designed a comparative study against state-of-the-art generative models. Our comparison includes representative methods from four categories: (1) GAN-based frameworks, specifically Pix2Pix, which has been adopted in previous DSA generation studies for cerebral [16] and abdominal [19] vasculature; (2) Diffusion-based models, implementing a conditional Denoising Diffusion Probabilistic Model (DDPM) to represent current state-of-the-art generative fidelity [5]; (3) Transformer-based image restoration architectures, specifically Uformer [18], representing powerful general-purpose models for robust artifact removal; and (4) the state-of-the-art DSA generation method, *i.e.,* AaDSA [7, 6]. Note that the original AaDSA framework relies on an artifact mask guidance module designed for a motion-corrupted training dataset. Given the dataset with minimal artifacts used in our work, we adapted its implementation by omitting the explicit artifact mask guidance while retaining its core generative architecture. All models were trained and tested on the same dataset splits to ensure consistency.

Quantitative and qualitative results. Quantitative results in Table 2 demonstrate that PH-VQVAE significantly outperforms those competing methods, achieving the highest structural fidelity (SSIM: 0.940±0.022) and peak signal-to-noise ratio (PSNR: 33.07±4.12 dB) with the lowest perceptual error (LPIPS: 0.095). Visual comparisons in Fig. 2 further validate this superiority, particularly in challenging low-contrast regions (red circle). While comparison methods suffer from vessel discontinuities (red arrows) or hallucinatory artifacts (blue arrow), our approach successfully reconstructs continuous and complete vascular trees (green arrows). Moreover, our model exhibits robustness by generating clean backgrounds even when the ground truth contains moderate motion artifacts (blue circle), confirming that the learned anatomical priors effectively filter out noise inherent in the training targets rather than memorizing artifacts.

4 Conclusion

In this work, we presented the Prior-guided Hierarchical Vector Quantized Variational Autoencoder (PH-VQVAE) to address the persistent challenges of discontinuous vessel reconstruction and motion artifacts in deep learning-based DSA imaging. Specifically, our dual-branch encoder with attentive feature integration leverages the mask prior to enhance faint vessel signals, effectively resolving the issue of vessel discontinuity even under low-contrast conditions. Meanwhile, the hierarchical vector quantization (VQ) robustly eliminates stochastic motion artifacts while preserving intricate vascular details that single-scale approaches often miss. Comprehensive experiments demonstrate that PH-VQVAE significantly outperforms state-of-the-art transformer-based and diffusion-based approaches, yielding high-fidelity DSA reconstructions with superior vessel connectivity and artifact suppression. These results highlight the clinical potential

of our framework to rescue non-diagnostic "junk films", reducing the need for repeated radiation exposure and contrast injection in interventional procedures.

5 Acknowledgment

This work is funded by the National Natural Science Foundation of China (No.62401280), Natural Science Research Start up Foundation of Recerruiting Talents of Nanjing University of Posts and Telecommunications (No.NY224024) and Natural Science Foundation of Shanghai (No.24ZR1445300).

6 Disclosure of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Gao, Y., Song, Y., Yin, X., Wu, W., Zhang, L., Chen, Y., Shi, W.: Deep learning-based digital subtraction angiography image generation. International Journal of Computer Assisted Radiology and Surgery **14**, 1775–1784 (2019)
2. Gu, S., Chen, D., Bao, J., Wen, F., Zhang, B., Chen, D., Yuan, L., Guo, B.: Vector quantized diffusion model for text-to-image synthesis. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10696–10706 (2022)
3. Guan, S., Wang, T., Sun, K., Meng, C.: Transfer learning for nonrigid 2d/3d cardiovascular images registration. IEEE Journal of Biomedical and Health Informatics **25**(9), 3300–3309 (2020)
4. Han, L., Tan, T., Zhang, T., Wang, X., Gao, Y., Lu, C., Liang, X., Dou, H., Huang, Y., Mann, R.: Non-adversarial learning: vector-quantized common latent space for multi-sequence mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 481–491. Springer (2024)
5. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
6. Liu, Y., Du, D., Liu, Y., Tu, S., Yang, W., Han, X., Suo, S., Liu, Q.: Subtraction-free artifact-aware digital subtraction angiography image generation for head and neck vessels from motion data. Computerized Medical Imaging and Graphics **121**, 102512 (2025)
7. Liu, Y., Du, D., Tu, S., Yang, W., Suo, S., Han, X.: Artifact-aware digital subtraction angiogram image generation for head and neck vessels. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 3527–3530. IEEE (2024)
8. Meijering, E.H., Niessen, W.J., Bakker, J., van der Molen, A.J., de Kort, G.A., Lo, R.T., Mali, W.P., Viergever, M.A.: Reduction of patient motion artifacts in digital subtraction angiography: evaluation of a fast and fully automatic technique. Radiology **219**(1), 288–293 (2001)

9. Meijering, E.H., Niessen, W.J., Viegever, M.: Retrospective motion correction in digital subtraction angiography: a review. IEEE Transactions on Medical Imaging **18**(1), 2–21 (1999)
10. Nejati, M., Sadri, S., Amirfattahi, R., et al.: Nonrigid image registration in digital subtraction angiography using multilevel b-spline. BioMed research international **2013** (2013)
11. Nikolau, E.: Development of a prototype 2D dual-energy subtraction angiography system for image-guided interventions. The University of Wisconsin-Madison (2024)
12. Okamoto, K., Ito, J., Sakai, K., Yoshimura, S.: The principle of digital subtraction angiography and radiological protection. Interventional Neuroradiology **6**(1_suppl), 25–31 (2000)
13. Shaban, S., Huasen, B., Haridas, A., Killingsworth, M., Worthington, J., Jabbour, P., Bhaskar, S.M.M.: Digital subtraction angiography in cerebrovascular disease: current practice and perspectives on diagnosis, acute treatment and prognosis. Acta Neurologica Belgica pp. 1–18 (2021)
14. Su, R., van der Sluijs, M., Cornelissen, S., van Zwam, W., van der Lugt, A., Niessen, W., Ruijters, D., van Walsum, T., Dalca, A.: Angiomoco: Learning-based motion correction in cerebral digital subtraction angiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 770–780. Springer (2023)
15. Sundarapandian, M., Kalpathi, R., Manason, V.D.: Dsa image registration using non-uniform mrf model and pivotal control points. Computerized medical imaging and graphics **37**(4), 323–336 (2013)
16. Ueda, D., Katayama, Y., Yamamoto, A., Ichinose, T., Arima, H., Watanabe, Y., Walston, S.L., Tatekawa, H., Takita, H., Honjo, T., et al.: Deep learning–based angiogram generation model for cerebral angiography without misregistration artifacts. Radiology **299**(3), 675–681 (2021)
17. Wang, G., Wang, P., Li, Y., Su, T., Liu, X., Wang, H.: A motion artifact reduction method in cerebrovascular dsa sequence images. International Journal of Pattern Recognition and Artificial Intelligence **32**(08), 1854022 (2018)
18. Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.U.: A general u-shaped transformer for image restoration. 2022 ieee. In: CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 17662–17672 (2022)
19. Yonezawa, H., Ueda, D., Yamamoto, A., Kageyama, K., Walston, S.L., Nota, T., Murai, K., Ogawa, S., Sohgawa, E., Jogo, A., et al.: Maskless 2-dimensional digital subtraction angiography generation model for abdominal vasculature using deep learning. Journal of Vascular and Interventional Radiology **33**(7), 845–851 (2022)
20. Zhao, H., Bai, Y., Chen, L., Ma, J., Lei, Y., Sun, T., Wu, L., Zhang, R., Xu, Z., Liang, X., et al.: Generative ai-based low-dose digital subtraction angiography for intra-operative radiation dose reduction: a randomized controlled trial. Nature Medicine pp. 1–9 (2026)