ReScan-IA: A Spatially-Adaptive Diffusion Framework for Controllable 3D Intracranial Aneurysm Inpainting

Tzu I Chuang[1*], Orhun Utku Aydin[1*,†], Satoru Tanioka[1], Dietmar Frey[1], and Adam Hilbert[1]

CLAIM – Charité Lab for AI in Medicine, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Charitéplatz 1, 10117 Berlin, Germany
tichuang0115@gmail.com; orhunutkuaydin@gmail.com; satoru.tanioka.ns@icloud.com; dietmar.frey@charite.de; adam.hilbert@charite.de

Abstract. Intracranial aneurysms exhibit substantial variability in size, morphology, and anatomical location, limiting the availability of diverse and balanced training data for CTA-based learning tasks such as detection and segmentation. Existing synthesis and augmentation approaches either lack explicit spatial control or struggle to generate anatomically consistent aneurysms in 3D angiography volumes. We propose ReScan-IA, a spatially-adaptive 3D diffusion framework for controllable aneurysm synthesis. We frame aneurysm inpainting as a virtual re-scan, where a clinically plausible aneurysm is introduced into an existing CTA volume as if it had been captured in a repeat acquisition. ReScan-IA integrates spatially-adaptive conditioning to control aneurysm placement and morphology. A binary vessel prior promotes vascular continuity, while multi-class Circle of Willis labels are used during synthetic sampling to select anatomically corresponding vessel territories, insertion sites, and aneurysm templates. We train on 1,223 multi-center CTA volumes from the LargeIA and RSNA datasets and evaluate utility on two independent external datasets under real-only, synthetic-only, and augmented training settings. Augmenting 250 real cases with synthetic samples improves voxel-wise Dice on External A from 41.1% to 62.2%, while reducing false positives per case from 5.07 to 1.83 and improving recall over the real-only baseline across both external cohorts. Implementation is available at: https://github.com/9xEzreaL/ReScan-IA

Keywords: Intracranial Aneurysm · Medical Image Synthesis · Diffusion Models.

[*] Equal contribution; each may list themselves as lead co-first author.
[†] Corresponding author: Orhun Utku Aydin, orhunutkuaydin@gmail.com.

1 Introduction

Intracranial aneurysms (IAs) exhibit substantial variability in size, morphology, and anatomical location [9]. Accurate detection and segmentation in CTA volumes require diverse and well-balanced training data. However, annotated IA datasets are limited, highly imbalanced, and subject to cross-center domain shift [1, 5, 12], restricting the generalization of learning-based models.

Generating synthetic training samples has emerged as a promising solution to address data scarcity and improve model robustness in medical imaging [2, 3, 10]. However, existing data augmentation strategies, including geometric transformations and intensity perturbations, fail to introduce novel aneurysm morphologies [4, 14]. GAN-based synthesis methods often operate in 2D and lack explicit spatial controllability [16]. Recent diffusion-based inpainting approaches improve visual fidelity but do not enforce vascular structural consistency [13, 15], leading to unrealistic pathological insertions in 3D CTA volumes.

To address these limitations, we propose ReScan-IA, a spatially-adaptive 3D diffusion framework for controllable aneurysm synthesis. We formulate aneurysm insertion as a re-scanning process, where a clinically plausible aneurysm is introduced into an existing CTA volume while preserving surrounding vascular anatomy. The model incorporates dual conditioning: (1) a vessel segmentation prior to enforce anatomical continuity, and (2) a spatial control mask that enables precise localization and morphology control via decoder-level spatial modulation.

Our contributions are: (1) We introduce a spatially-adaptive diffusion framework for structure-aware 3D aneurysm pathology synthesis. (2) We design a dual-conditioning architecture that combines input-level anatomical priors with feature-level spatial modulation. (3) We demonstrate that controllable synthetic augmentation significantly improves downstream IA segmentation and detection under low-data and cross-dataset settings.

2 Methods

2.1 Problem Formulation

As illustrated in Fig. 1(a), we formulate aneurysm synthesis as anatomically guided inpainting in 3D CTA volumes. Let $x_0 \in R^{1\times H\times W\times D}$ denote an original 3D CTA scan. Given an inpainting mask m, our goal is to generate a modified volume $\hat{x}$ that preserves vascular anatomy outside the mask while synthesizing a plausible aneurysm inside.

Structural guidance is provided by a binary vessel mask $v \in \{0,1\}^{1\times H\times W\times D}$ encoding vascular geometry. The model is further conditioned on a spatial control mask $s \in \{0,1\}^{1\times H\times W\times D}$ that specifies aneurysm location and morphology. During training, s corresponds to the ground-truth aneurysm region; during inference, it serves as a user-defined control signal. The generative process of ReScan-IA is formulated as learning the conditional probability

$$p_\theta(\hat{x} \mid x_0, m, v, s). \tag{1}$$

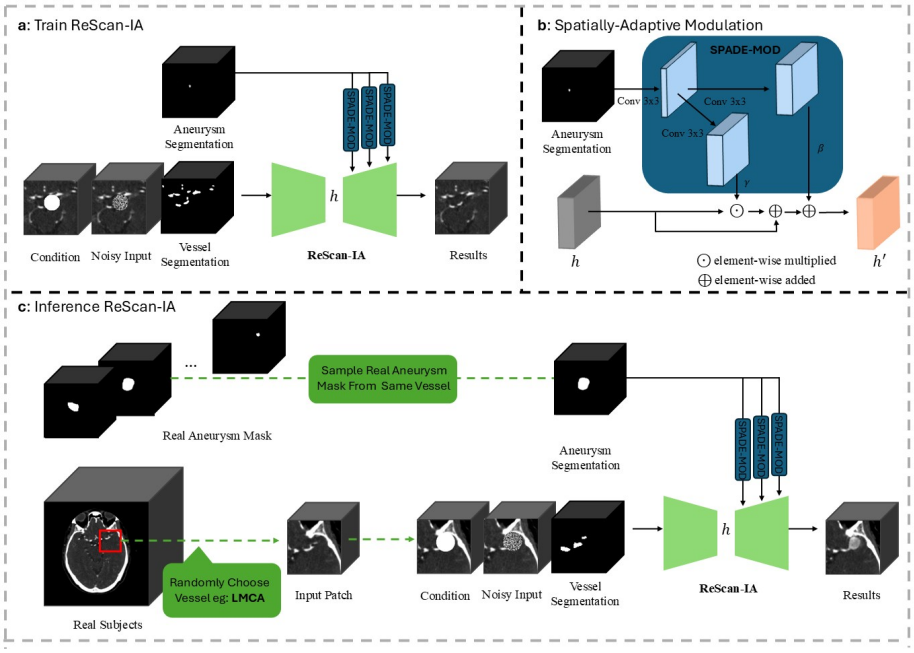

Fig. 1. Overview of the ReScan-IA framework. (a) Training procedure of the diffusion model with vessel-aware and spatial conditioning. (b) Illustration of the spatially-adaptive feature modulation mechanism. (c) Downstream synthetic data generation and mask sampling strategy during inference.

During synthesis, context preservation is enforced by

$$\hat{x} = (1 - m) \odot x_0 + m \odot \tilde{x}, \tag{2}$$

where $\tilde{x}$ denotes the denoised content within the masked region. This formulation preserves anatomical context outside the mask while enabling controllable aneurysm synthesis.

2.2 Dual-Conditional Diffusion Architecture

To model the conditional distribution $p_\theta(\hat{x} \mid x_0, m, v, s)$, we adopt a masked 3D diffusion-based inpainting framework. ReScan-IA employs a 3D U-Net denoiser ϵ_θ to predict the Gaussian noise injected during the forward diffusion process [6]. The forward process progressively perturbs x_0 to obtain x_t at timestep $t \in \{1, \ldots, T\}$.

At each denoising step, the network receives

$$C_{in} = \text{concat}(x_t \odot m + x_0 \odot (1 - m),\, x_{\text{bg}},\, v), \tag{3}$$

where

$$x_{\text{bg}} = (1 - m) \odot x_0. \tag{4}$$

Here, the masked noisy input $x_t \odot m + x_0 \odot (1 - m)$ enables diffusion progressive refinement within the target region, while x_{bg} explicitly preserves the surrounding vascular anatomy. The vessel mask v provides structural guidance to maintain vascular continuity.

Let

$$h_e = \text{Encoder}(C_{in}, t), \tag{5}$$

denote the hierarchical latent representations extracted by the encoder, where t is encoded using sinusoidal timestep embeddings. While input-level conditioning enforces global anatomical consistency, explicit morphology and location control are introduced via feature-level spatial modulation in the decoder (Sec. 2.3).

2.3 3D Spatially-Adaptive Feature Modulation

The proposed spatially-adaptive modulation mechanism is illustrated in Fig. 1(b). To achieve fine-grained spatial control over aneurysm placement, we integrate a 3D spatially-adaptive feature modulation mechanism within the decoder of the U-Net. Inspired by SPADE [11], we extend spatial modulation to volumetric feature maps using 3D convolutions. Let $h \in R^{C \times D \times H \times W}$ denote an intermediate decoder feature map. Given a binary spatial control mask $s \in \{0, 1\}^{1 \times D \times H \times W}$, resampled to match the resolution of h, we compute spatially varying modulation parameters $\gamma(s)$ and $\beta(s)$ via lightweight 3D convolutional layers. Specifically, a shared embedding is first extracted as

$$A = \text{ReLU}\big(\text{Conv}_{3\times3\times3}(s)\big), \tag{6}$$

followed by

$$\gamma(s) = \text{Conv}_{\gamma}(A), \quad \beta(s) = \text{Conv}_{\beta}(A). \tag{7}$$

The feature map is then modulated through an affine transformation:

$$h' = h \odot \big(1 + \gamma(s)\big) + \beta(s), \tag{8}$$

where $\odot$ denotes element-wise multiplication. This modulation injects spatial control signals directly into decoder features, allowing the network to adaptively synthesize aneurysm structures at user-specified locations.

Notably, spatial modulation is applied exclusively in the decoder, forming an asymmetric design: the encoder extracts stable anatomical representations from C_{in}, while the decoder integrates pathological structures guided by s. This design ensures precise spatial control while preserving global vascular context.

2.4 Training Objective

The diffusion model is trained with an ℓ_1 masked denoising objective to predict injected Gaussian noise. To focus the generative capacity on the aneurysm synthesis region while preserving the surrounding healthy anatomy, we define a spatially masked loss:

$$\mathcal{L} = E_{x_0,\epsilon,t}\left[\left\|m \odot \left(\epsilon - \epsilon_\theta(C_{in}, t, s)\right)\right\|_1\right], \tag{9}$$

where m denotes the inpainting mask and $\odot$ represents element-wise multiplication. By restricting the loss computation to voxels where $m = 1$, gradients are propagated only within the designated synthesis region, thereby preserving anatomical structures outside the mask.

2.5 Diffusion Training Details

The diffusion model was trained for 100,000 iterations on an NVIDIA L40 GPU using Adam (lr 5×10^{-5}) with $T = 1000$ timesteps and a linear schedule [13]. Training used 96^3 patches (batch size 1), and inference used 128^3 patches for increased spatial context.

2.6 Datasets and Preprocessing

ReScan-IA was trained on 1,223 multi-center CTA volumes, including 287 subjects from the RSNA Intracranial Aneurysm Detection challenge dataset [12] and 936 subjects from LargeIA [1]. An additional 250 held-out LargeIA subjects were excluded from ReScan-IA training and used as the internal pool for downstream segmentation and detection experiments. Cross-dataset generalization was assessed on two non-overlapping LargeIA external test cohorts, External A (n=71) and External B (n=67), which were distinct from all subjects used for ReScan-IA training, synthetic sample generation, and model selection. All volumes were resampled to isotropic 0.5 mm spacing, skull-stripped using SynthStrip [7], clipped to [-50, 450] HU, and linearly normalized to [-1, 1]. The same preprocessing pipeline was applied to all datasets.

2.7 Downstream Experimental Setup

We evaluated the utility of ReScan-IA for data augmentation through controlled downstream experiments on aneurysm segmentation and detection tasks using nnU-Net [8]. All models were trained with the 3d_fullres configuration (patch size = $96 \times 96 \times 96$) under a 5-fold cross-validation setup for 100 epochs with identical hyperparameter settings. The voxel spacing and preprocessing pipeline followed Sec. 2.6.

We defined four controlled training regimes based on subjects from the LargeIA internal validation set:

(1) Real-10: 10 real CTA subjects were used for training.

Table 1. Segmentation and detection performance on External Datasets A and B. Best results per dataset are highlighted in bold.

Methods	**Segmentation**		**Detection**		
	Dice (Voxel-wise)	Dice (Case-wise)	FP / case	Precision	Recall
External A					
Real-10	35.90	10.24 (4.10–16.38)	1.93 (1.42–2.44)	9.26 (3.36–15.16)	30.00 (16.26–43.74)
Synthetic-250	42.83	21.95 (15.45–28.45)	11.71 (10.50–12.92)	5.39 (4.12–6.67)	**88.89** (79.34–98.44)
Real-250	41.11	11.17 (4.82–17.51)	5.07 (4.49–5.66)	3.81 (1.73–5.88)	30.00 (16.26–43.74)
Real+Synthetic (250+250)	**62.24**	**24.65** (16.27–33.03)	**1.83** (1.49–2.16)	**19.02** (12.22–25.82)	58.89 (44.11–73.67)
External B					
Real-10	19.46	9.88 (4.39–15.37)	**1.72** (1.15–2.28)	12.81 (5.35–20.28)	27.78 (14.36–41.20)
Synthetic-250	**35.20**	**18.80** (13.12–24.49)	12.16 (11.30–13.03)	5.15 (3.93–6.36)	**87.04** (77.32–96.76)
Real-250	28.58	12.00 (6.47–17.53)	5.28 (4.67–5.90)	4.85 (2.61–7.08)	38.89 (24.61–53.17)
Real+Synthetic (250+250)	28.06	16.57 (9.48–23.67)	1.82 (1.51–2.14)	**18.08** (11.17–24.98)	51.11 (36.26–65.96)

(2) Synthetic-250: For each of the 10 subjects, 25 synthetic volumes were generated using ReScan-IA, yielding 250 synthetic-only samples. Synthetic morphologies were sampled from real aneurysm masks and inserted into anatomically corresponding vessels.

(3) Real-250: 250 real CTA subjects were used for training.

(4) Real+Synthetic (250+250): The 250 real subjects from (3) were augmented with the 250 synthetic volumes generated in (2).

All trained models were evaluated on two independent external test cohorts (External A and External B) to assess cross-dataset generalization performance. For segmentation evaluation, we report both case-wise Dice and voxel-wise Dice. Case-wise Dice was computed per subject and averaged across the test set with 95% confidence intervals. Voxel-wise Dice was computed by summing true positives, false positives, and false negatives across all cases before applying the Dice formula, corresponding to voxel-wise aggregation over the entire dataset.

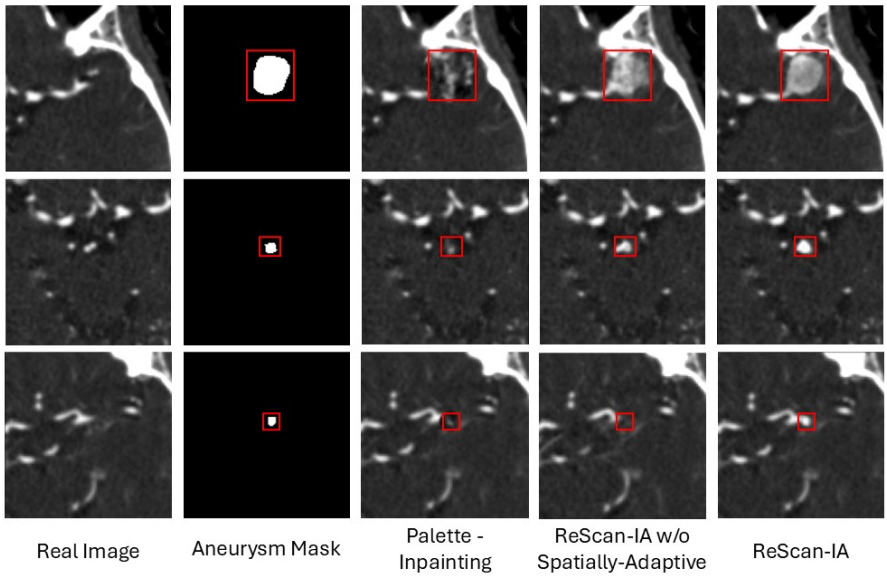

Fig. 2. Qualitative ablation comparison. From left to right: real CTA image, aneurysm mask, Palette diffusion inpainting, ReScan-IA without spatially-adaptive modulation, and the proposed ReScan-IA. Palette fails to preserve vascular continuity, while removing spatially-adaptive modulation leads to poor morphological control. ReScan-IA achieves anatomically consistent and controllable synthesis.

3 Results

3.1 Qualitative Ablation

Qualitative comparisons between Palette diffusion, ReScan-IA without spatially-adaptive modulation, and the full ReScan-IA are presented in Fig. 2.

Palette produces discontinuities along vessels and unrealistic aneurysm structures. Removing spatially-adaptive modulation results in plausible textures but inaccurate adherence to the intended spatial mask. The full ReScan-IA preserves vascular continuity and generates aneurysms aligned with the specified control region.

3.2 Quantitative Results

Segmentation Performance. On External A, Real+Synthetic achieved the highest voxel-wise Dice (62.24%), compared to Real-250 (41.11%) and Synthetic-250 (42.83%). Case-wise Dice followed a similar pattern (24.65 vs. 11.17 and 21.95). On External B, Synthetic-250 obtained the highest voxel-wise Dice (35.20%), exceeding Real-250 (28.58%) and Real+Synthetic (28.06%).

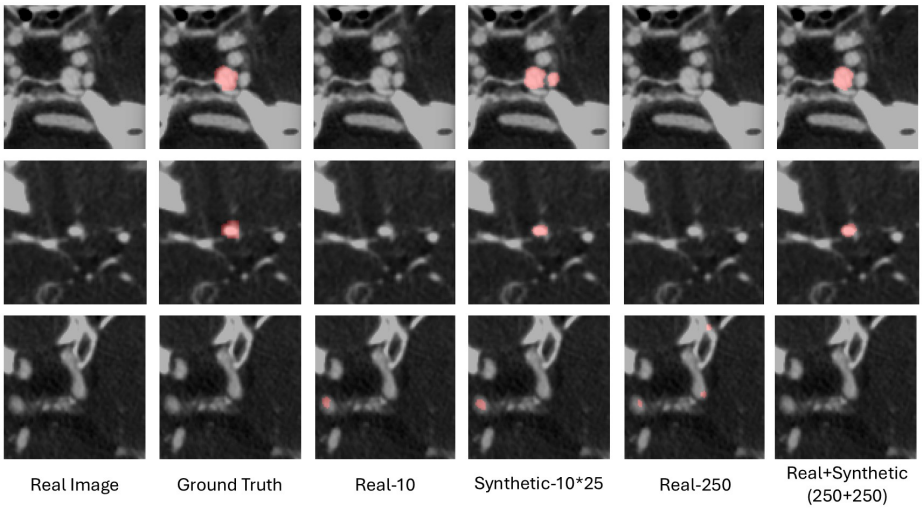

Fig. 3. Qualitative comparison of downstream segmentation task. Synthetic-only training increases sensitivity but yields more false positives, while real+synthetic training improves boundary accuracy and suppresses spurious detections.

Detection Performance. On External A, Real+Synthetic achieved the lowest false positives per case (1.83) and the highest precision (19.02%), with recall of 58.89%. Synthetic-250 achieved the highest recall (88.89%) but resulted in higher false positives (11.71). On External B, Synthetic-250 achieved the highest recall (87.04%), while Real+Synthetic reduced false positives (1.82) and achieved higher precision (18.08%). Across both datasets, Synthetic-250 consistently achieved the highest recall, whereas Real+Synthetic yielded the highest precision.

3.3 Qualitative Segmentation Comparison

Qualitative comparisons across training regimes reveal distinct detection behaviors (Fig. 3). Models trained with only 10 real cases exhibit limited sensitivity. Synthetic-only training improves lesion detection but introduces more false positives. Training with 250 real cases reduces spurious detections, yet remains susceptible to boundary under-segmentation. In contrast, the combined Real+Synthetic regime achieves more accurate delineation with fewer false positives.

4 Discussion and Conclusion

We presented ReScan-IA, a spatially-adaptive 3D diffusion framework for controllable aneurysm synthesis in CTA volumes. By combining vessel-aware conditioning with decoder-level spatial modulation, the model enables anatomically

consistent and location-specific aneurysm inpainting for structured synthetic data generation. This local inpainting formulation is motivated by the inherently localized nature of intracranial aneurysms, making 3D diffusion training more practical despite its computational cost while improving data efficiency by focusing model capacity on the pathology region rather than the full volume.

Across two external cohorts, controllable synthetic augmentation improved downstream segmentation and detection under limited-data and domain-shift settings. Synthetic-only training increased sensitivity, whereas combining real and synthetic data reduced false positives and improved precision, resulting in more balanced detection behavior.

The performance differences between External A and B may reflect variations in aneurysm size distribution. Since voxel-wise Dice is computed via dataset-level voxel aggregation, larger lesions contribute more strongly to the overall score due to greater voxel support. In contrast, case-wise Dice reflects per-subject segmentation consistency and is more sensitive to small or difficult lesions.

Limitations include dependence on accurate vessel/control masks and patch-based synthesis that may limit global context. Future work will explore stronger geometric priors and broader validation across additional datasets. Overall, ReScan-IA highlights the potential of controllable pathology-aware diffusion models to enhance the robustness and generalization of clinical IA detection systems under data scarcity and domain shift.

Disclosure of Interests. The authors have no competing interests.

References

1. Bo, Z.H., Qiao, H., Tian, C., Guo, Y., Li, W., Liang, T., Li, D., Liao, D., Zeng, X., Mei, L., et al.: Toward human intervention-free clinical diagnosis of intracranial aneurysm via deep neural network. Patterns **2**(2) (2021)
2. Chalcroft, L., Crinion, J., Price, C.J., Ashburner, J.: Domain-agnostic stroke lesion segmentation using physics-constrained synthetic data. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 163–173. Springer (2025)
3. Deleat-besson, R., Goujat, C., Bernard, O., Croisille, P., Viallon, M., Duchateau, N.: Controllable latent diffusion model to evaluate the performance of cardiac segmentation methods. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 100–109. Springer (2025)
4. Ding, W., Ji, K., Castro, S., Luo, Y., Roi, D., Yap, C.H.: Two-stage generative model for intracranial aneurysm meshes with morphological marker conditioning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 595–604. Springer (2025)
5. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. IEEE Transactions on Biomedical Engineering **69**(3), 1173–1185 (2022)
6. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
7. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)

8. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
9. Keedy, A.: An overview of intracranial aneurysms. Mcgill Journal of Medicine **9**(2), 141–146 (2006)
10. Luna, M., Baek, J., Kim, W.H., Son, W.G., Lee, K.M., Kim, H.J., Kim, J.: Improved tumor segmentation using selective synthetic augmentation for enhanced surgical planning in breast mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 315–324. Springer (2025)
11. Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y.: Semantic image synthesis with spatially-adaptive normalization. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2337–2346 (2019)
12. Rudie, J., Calabrese, E., Ball, R., Chang, P., Chen, R., Colak, E., Correia de Verdier, M., Prevedello, L., Richards, T., Saluja, R., Zaharchuk, G., Sho, J., Vazirabad, M.: RSNA Intracranial Aneurysm Detection. Kaggle, https://www.kaggle.com/competitions/rsna-intracranial-aneurysm-detection (2025), accessed 22 February 2026
13. Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., Norouzi, M.: Palette: Image-to-image diffusion models. In: ACM SIGGRAPH 2022 conference proceedings. pp. 1–10 (2022)
14. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. Journal of big data **6**(1), 1–48 (2019)
15. Wolleb, J., Sandkühler, R., Bieder, F., Valmaggia, P., Cattin, P.C.: Diffusion models for implicit image segmentation ensembles. In: International conference on medical imaging with deep learning. pp. 1336–1348. PMLR (2022)
16. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: A review. Medical image analysis **58**, 101552 (2019)