Two Projections Suffice for Cerebral Vascular Reconstruction

Alexandre Cafaro[1,3,4], Reuben Dorent[1] Nazim Haouchine[1], Vincent Lepetit[5], Nikos Paragios[3], William M. Wells III.[1,2], and Sarah Frisken[1]

[1] Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA
[2] Massachusetts Institute of Technology, Cambridge, MA, USA
[3] TheraPanacea, Paris, France
[4] Paris-Saclay University, Gustave Roussy, Inserm 1030, Villejuif, France
[5] LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France

Abstract. 3D reconstruction of cerebral vasculature from 2D biplanar projections could significantly improve diagnosis and treatment planning. We introduce a novel approach to tackle this challenging task by initially backprojecting the two projections, a process that traditionally results in unsatisfactory outcomes due to inherent ambiguities. To overcome this, we employ a U-Net approach trained to resolve these ambiguities, leading to significant improvement in reconstruction quality. The process is further refined using a Maximum A Posteriori strategy with a prior that favors continuity, leading to enhanced 3D reconstructions. We evaluated our approach using a comprehensive dataset comprising segmentations from approximately 700 MR angiography scans, from which we generated paired realistic biplanar DRRs. Upon testing with held-out data, our method achieved an 80% Dice similarity w.r.t the ground truth, superior to existing methods. Our code and dataset are available at https://github.com/Wapity/3DBrainXVascular.

1 Introduction

Digital Subtraction Angiography (DSA) plays an important role in the planning and treatment of neurovascular diseases providing surgeons with rich information about the brain angioarchitecture and hemodynamics [15]. Although 3D MRA, CTA, or rotational DSA exist, 2D DSA remains the gold standard, due to its high resolution and clinical availability. DSA is commonly acquired as a set of biplanar anterior-posterior (AP) and lateral (L) projections of the vascular network [16,9]. Unlike 3D rotational scanners, which are not suitable for real-time interventions, and single DSAs, which are limited to simpler tasks, biplanar DSAs offer an optimal balance of speed, anatomical constraints, cost efficiency, and reduced radiation exposure. Yet, projection onto 2D images causes vessel overlap, which makes it difficult for surgeons to confidently localize lesions, understand their shapes and morphologies, or distinguish between vessels feeding and draining malformations when the number of 2D views is limited [16,9]. Thus, 3D reconstruction becomes critical.

Reconstructing cerebral vasculature from biplanar projections is a heavily ill-posed problem. The dense and intricate arrangements of blood vessels overlap and intertwine onto 2D projections, raising major ambiguities. While few attempts have been made to tackle this challenging problem, most of them focus on simpler vascular structures, like main coronary arteries, and typically require manual adjustments for vessel endpoints and bifurcations. To obtain more complex reconstructions, other techniques rely on pre-existing 3D models of patient's vasculature to add constraints or simulate flow [6]. However, the availability of 3D imaging cannot be guaranteed in clinical practice. A non-learning approach [7] relies on structural and temporal constraints but requires perfect tedious semi-manual annotations of segmented vessel centerlines from DSAs, limiting its use for real-time intervention. Alternatively, various deep learning techniques have been proposed, including self-supervised approaches [18], Neural Radiance Fields (NeRF) techniques [12], denoising approach [17] and Generative Adversarial Networks (GANs) [19]. These models typically aim at learning a prior to disambiguate DSAs and performing direct prediction. However, none of these techniques reached a good level of performance when only 2 projections are available. 3D backprojected volumes offers geometrical cues for reconstruction, albeit as noisy and ambiguous representations of the actual volume. Unlike the Denoiser approach [17], we show that a deep learning network can significantly clarify these volumes by learning priors on vascular patterns, resulting in closely matching reconstructions.

Contribution. We propose in this paper a novel method for 3D reconstruction of DSA from only two projections. Our method follows a two step process; the first step involves a disambiguating reconstructor from back-projected volumes, built upon the Denoiser model. We enhanced the model with improved architecture and design to tackle the inherent complexities of the task. In the next step, we refine our initial predictions to find the Maximum A Posteriori (MAP) estimate given the projections. This refinement occurs through iterative optimization on a voxel grid, starting from our preliminary estimation of the vasculature. To improve the vascular network's structural integrity and connectivity, we introduce a connectivity prior inspired by Ising prior [5]. We conduct several experiments to benchmark our method against the existing state-of-the-art and show that our method delivers a high level of reconstruction accuracy, closely matching the target vasculatures. It marks a significant improvement over existing techniques, suggesting that as few as two projections might be sufficient for disambiguating structures for accurate 3D reconstruction. Our method is a promising approach, paving the way for validation on real data and potential clinical translation.

2 Method

Our proposed methodology uses a two-step pipeline, as illustrated in Figure 1. First, a 3D U-Net predicts an initial 3D vasculature from the back-projected vol-

ume. Second, Maximum A Posteriori estimation is employed to refine the initial 3D model with a newly introduced connectivity loss that encourages closing.

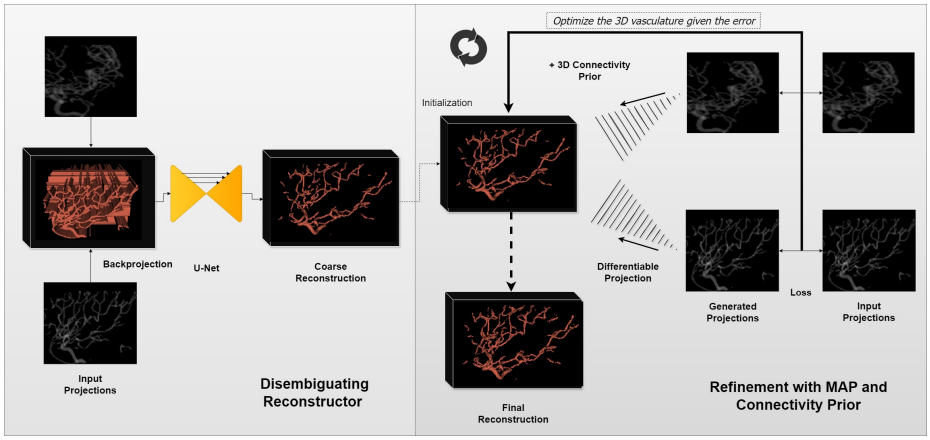

Fig. 1. Our pipeline. Initially, we employ a 3D U-Net to generate a preliminary vasculature model from back-projected volumes. This model is subsequently refined through Maximum A Posteriori estimation, with the introduction of a connectivity prior, aimed at enhancing the model's structural cohesiveness and closure.

2.1 Disambiguating Reconstructor

Our goal is to address the challenging and ill-posed problem of reconstructing 3D brain vasculature from biplanar projections (I_0, I_{90}), acquired simultaneously on a bi-plane scanner, available in most interventional radiology suites. To create an initial 3D model of the brain vasculature, we propose to employ a 3D nnUNet [10] model denoted as U_θ parametrized by the weights θ. Specifically, we first convert 2D images into a 3D volume V_b using back-projection. Then, the 3D nnUNet aims to disambiguate the back-projection to create a filtered volume corresponding to a 3D brain vasculature.

Given the absence of real paired projection-3D volume datasets, we propose to train our 3D nnUNet model using synthetic projections from ground truth volumes, which were then converted into back-projected 3D volumes. Specifically, we model blood vessels within a 3D space, assigning them a binary map with an intensity value of 1. Then we simulate biplanar projection using Mean Intensity Projection along rays, corresponding to summing up log attenuation coefficients. This approach mirrors the accumulation of attenuation that occurs as X-rays pass through vessels. Finally, we create a backprojected volume V^b by extending these 2D projections into 3D along their original rays. Note that projection and backprojection are fast and differentiable. Overall, this enables us to create a paired dataset of back-projected 3D volumes and ground truth GT predictions.

To optimize parameters θ of our model, we use a combination of Dice and cross-entropy loss:

$$\mathcal{L}(U_\theta(V_b), GT) = \lambda_D \cdot \mathcal{L}_{Dice}(U_\theta(V_b), GT) + \lambda_{CE} \cdot \mathcal{L}_{CE}(U_\theta(V_b), GT), \quad (1)$$

where λ_D and λ_{CE} are weighting coefficients, $U_\theta(V_b)$ represents the predicted reconstruction for a given back-projected volume V_b and GT denotes the ground truth 3D vasculature.

2.2 Refinement of 3D Vasculature with MAP Estimate.

Building on our initial prediction, we then proceed to a refinement phase. This phase focuses on fine-tuning a 3D voxel-grid initialized with our deep learning output, $V_{coarse} = U_\theta(V_b)$, to improve 3D reconstruction alignment with the original projections, as well as connectivity.

MAP Estimation. The process employs MAP estimation, using the pseudo-probability V_{coarse} as our initialization, i.e. $V_0 = V_{coarse}$. Our goal is to iteratively adjust this volume toward a MAP estimate V^*, achieving an optimal vasculature configuration that matches observed projections while enforcing structural integrity and connectivity, leading to the following optimization problem:

$$V^* = \underset{V}{\operatorname{argmax}} \log \mathcal{P}(V|I_0, I_{90}) = \underset{V}{\operatorname{argmax}} \mathcal{L}(V, I_0, I_{90}) + \mathcal{R}(V) \quad (2)$$

Here, $\mathcal{P}(V|I_0, I_{90})$ denotes the posterior probability of the vasculature V given the observed projections I_0, I_{90}, where $\mathcal{L}(V, I_0, I_{90})$ and $\mathcal{R}(V)$ respectively represents the log-likelihood of the current estimate to the observed projections, and a regularization term that integrates a specially designed connectivity prior.

Connectivity Prior. We introduce a connectivity prior inspired by the Ising model [5] to enhance voxel interconnectivity and create a more cohesive vascular network. In the context of vasculature, continuity is expected within vessels, except at termination points. The proposed reguarlization term encourages neighboring elements to exhibit similar values for spatial coherence, while allowing for natural discontinuities at boundaries or edges. To achieve this, we formulate a loss function that acts as an energy function:

$$\mathcal{R}_c(V) = -\frac{1}{N} \sum_{w=1}^{W} \sum_{h=1}^{H} \sum_{d=1}^{D} \sum_{x \in \mathcal{N}(w,h,d)} V_{w,h,d} \cdot V_x,$$

where $N = W \times H \times D$ and $\mathcal{N}(w, h, d)$ represents the set of all 26 neighboring voxels of (w, h, d). This formula calculates the negative sum of the product of each voxel with its 26 neighbors within the 3D grid, normalized by the number of points $N = W \times H \times D$. Minimizing this energy function encourages connecting voxels by growing connections and filling in gaps, thus improving structural integrity and coherence.

Optimization. The refinement begins with V_{coarse} and the optimization aims to balance data fidelity with the regularization informed by our connectivity prior. Eq. 2 refines as :

$$V^* = \underset{v}{\operatorname{argmin}} \left\| A_0 \circ V - I_0 \right\|_2 + \left\| A_{90} \circ V - I_{90} \right\|_2 + \lambda_c \mathcal{R}_c(V) \tag{3}$$

where the λ_c is fixed.

A_i are the projector operators under view i.

3 Experiments and Results

We evaluate our method for our target application, 3D vasculature reconstruction from biplanar projections. In this section, we introduce our dataset, our model architecture, and present both quantitative and qualitative comparisons with state-of-the-art methods. We also include an ablation study to evaluate the contribution of our regularization prior.

3.1 Dataset and Preprocessing

Voxel-based Binary Vasculature Maps Creation. Given the scarcity of 3D vasculature segmentations, we developed a comprehensive approach to generate a substantial dataset for training our deep learning model. Utilizing the publicly available TubeTK [3] dataset, we processed MRAs from 100 healthy patients, including 43 with detailed ground truth segmentations. To augment the dataset further, we trained a nnUNet model on these binary maps, achieving a validation Dice score of 0.75. Using this model we segmented additional MRAs from the publicly available IXI dataset [8], which comprises 580 Time-of-Flight (TOF) MRAs of healthy patients.

Clinical Realism. MRA typically images vessels in both brain hemispheres. In contrast, DSA is typically used to image one arterial branch at a time. Thus, DSA has 1) less vascular complexity than MRA and 2) vasculature restricted to one hemisphere. To mimic this clinical reality, we partitioned MRA images along the brain mid-plane. As a positive side effect, as we could use both hemispheres independently, this doubled the size of our dataset from 680 to 1360.

Computational Constraints. Given the substantial size of vasculature volumes, we optimized GPU efficiency by downsampling all volumes to a resolution of $0.8 \times 0.94 \times 0.94$. We employed Signed Distance Fields (SDFs) to maintain the integrity of thin vessels during downsampling, preserving thin vessel structures and reducing artifacts, unlike binary images which can cause these structures to break apart or disappear. We used FastGeodis [1] with truncation at 14 and level at 0.03. Additionally, by identifying the minimal bounding grid for the vasculature, we standardized the volumes to $112 \times 80 \times 128$.

Model Training. We created Digitally Reconstructed Radiographs (DRRs) from these volumes to serve as projections. These projections were then used to generate backprojected volumes as input for our model. These were resampled to a size of $128 \times 96 \times 128$ and Z-Score normalized. Our dataset was divided into $1,029$ training cases and 257 validation cases, split between different patients.

3D Reconstruction. In the test phase, we utilized 70 distinct cases from the IXI dataset [8] to assess our reconstruction methodology.

3.2 Implementation Details

Disambiguating Reconstructor. Our model, built with PyTorch [14] and tailored for the NVIDIA GeForce RTX 3090 GPU, adopts a nnUNet design with 6 encoding and 5 decoding blocks. It features asymmetric downsampling—5 times in larger dimensions and 4 times in the smallest—enhancing feature extraction across scales with channels increasing in the encoder (32, 64, 128, 256, 320, 320) and decreasing in the decoder (320, 256, 128, 64, 32). Skip connections improve information flow over simple encoder-decoders. Pooling mainly uses 2x2x2 kernel sizes, and convolutions are performed with 3x3x3 kernels. Running on a batch size of 3, the model starts with a 0.01 learning rate, using SGD with Nesterov momentum, a weight decay of 3e-5, and a polynomial rate scheduler. The loss function equally mixes dice and cross-entropy. Robustness is improved by extensive data augmentation, including spatial, noise, and contrast modifications. The model was trained in under a day for 500 epochs.

Optimization. Our reconstruction optimizes a full-resolution voxel grid, initialized with the model's preliminary predictions from backprojected volumes. It is done on the same GPU using the Adam [11] optimizer at a learning rate of 1e−1. Optimal weights ($\lambda_2 = 1$, $\lambda_c = 0.0002$) were determined through grid search. The process, taking 500 iterations, completes in about 20 seconds, but excluding connectivity loss cuts it down to 3 seconds.

3.3 Results and Discussion

Baselines. There are very few papers working on this specific task. We compared our approach against simple backprojection, learning-based supervised state-of-the-art Denoiser [17], voxel-grid optimization without prior, and unsupervised gan-based reconstruction model X2Vision [4]. As no implementation of Denoiser was available, we reimplemented it.

Metrics. For evaluation, we employed metrics including dice, cldice [13], and balanced Hausdorff distance (HD) [2]. Compared to dice, cldice provides a balanced view, especially valuing the retrieval of branching patterns and minimizing the bias towards larger vessels.

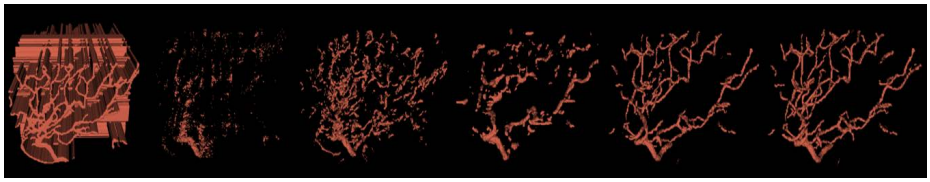

Fig. 2. Visual comparison of 3D reconstruction from biplanar projections by our model and baselines. Backprojection produces highly noisy and ambiguous reconstructions. Grid optimization captures only broad structures and non-ambiguous segments. X2Vision offers slightly clearer reconstructions, capturing the main artery and some structures but remains fuzzy. The Denoiser reconstructs parts of the vessels but struggles with connectivity and complex branching areas. In contrast, our reconstruction closely mirrors the actual data, showcasing better well-defined patterns and accurately connected, complex vessel networks.

3D Reconstruction from 2 Projections. Figure 2 visually presents our reconstructions with baseline methods, while Table 1 details our quantitative results. Our approach significantly outperforms others, closely matching the actual vessel geometry and enhancing the precision of vessel proximity to targets, achieving impressive results.

Backprojection retrieves all possible locations of vessel presence, resulting in very noisy reconstructions with numerous false positives. Direct voxel-grid optimization, similar to our refinement process but without proper initialization, only reconstructs unambiguous areas and misses detailed structural nuances, leading to many false negatives. This underscores the complexity of reconstruction without prior knowledge.

Table 1. Comparison with State-of-the-art. Standard deviations in parentheses.

Method	Dice ↑	clDice ↑	Balanced HD ↓
Voxel-Grid	0.22(±0.11)	0.16(±0.11)	1.81(±0.26)
X2Vision [4]	0.26(±0.03)	0.20(±0.03)	3.42(±0.46)
Denoiser [17]	0.34(±0.05)	0.28(±0.06)	1.94(±0.24)
Ours (coarse)	**0.77**(±**0.04**)	**0.75**(±**0.04**)	**0.42**(±**0.09**)
Ours (coarse w/ refinement)	**0.80**(±**0.04**)	**0.78**(±**0.04**)	**0.34**(±**0.09**)

X2Vision, which uses unsupervised GANs, struggles to learn the intricate and sparse nature of vascular structures. This weak prior allows it to identify the main artery but fails to produce continuous, realistic reconstructions. The Denoiser model, constrained by its simple architecture and lack of skip connections, only provides coarse vessel outlines and cannot capture complex vessel branching, showing limited improvement. We improved the Denoiser model by

adopting a more complex architecture, including skip connections, using a U-Net instead of a simple encoder-decoder, and combining dice and cross-entropy losses. We also introduced data augmentation and a refinement step. These enhancements significantly improved performance, leading to reconstructions that closely match the target vessel structures and branchings. Our refinement step further enhances the reconstruction. By introducing a MAP estimate refinement paired with a connectivity loss, we not only improve vessel connectivity but also refine vessel shapes and fill in missing structures, achieving more precise and closed vessel representations in the final volume.

Ablation Study. Our ablation study, summarized in Figure 3 and Table 2, reveals that initial reconstructions may contain gaps, but MAP refinement and incorporating connectivity loss significantly enhance the quality. MAP refinement aligns structures with projections, while connectivity loss improves cldice scores by improving capture of vessel centerlines, outperforming grid optimization. It clarifies complex junctions and promotes interconnected high-probability voxels, filling gaps and enhancing structural integrity and coherence.

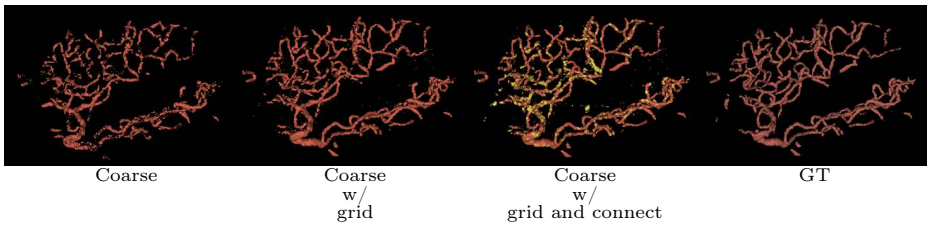

Fig. 3. Visual Ablation Study. The initial results from our model are promising but exhibit gaps. Refinement with MAP enhances closure and refinement. Adding the connectivity prior further strengthens these improvements, allowing the model to closely replicate the intricate complexity of actual vasculature. The yellow shows the additional closing when introducing the connectivity loss.

Table 2. Ablation Study. Standard deviations are provided in parentheses.

Method	Dice ↑	clDice ↑	Balanced HD ↓
Grid	0.22(±0.11)	0.16(±0.11)	1.81(±0.26)
Grid w/ Connectivity	0.29(±0.05)	0.23(±0.05)	2.47(±0.26)
Coarse	0.77(±0.04)	0.75(±0.04)	0.42(±0.09)
Coarse w/ Grid	0.79(±0.04)	0.76(±0.04)	0.35(±0.08)
Coarse w/ Grid + Connectivity	**0.80**(±**0.04**)	**0.78**(±**0.04**)	**0.34**(±**0.09**)

4 Conclusion and Future Work

We introduced a new method for the highly ill-posed 3D cerebral vascular reconstruction from biplanar DSAs. Our two-step approach starts with a disambiguating reconstructor, followed by refinement through MAP estimation and a connectivity prior. This method marks a notable advancement over existing methods, suggesting for the first time that two projections could effectively disambiguate complex vascular structures. Further improvements are expected with the integration of additional vascular properties. Due to GPU hardware capabilities, our work was limited to using downsampled volumes, restricting our method to vessels with diameters larger than 1-2mm. Additionally, we used synthetic DSA generated from automatically segmented MRA images due to the lack of paired MRA/DSA datasets. We are currently assembling such a dataset to validate our method on real data towards clinical translation.

Acknowledgments. This study was funded in part by the National Institutes of Health grants: R01EB034223, R03EB033910, and K25EB035166.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Asad, M., Dorent, R., Vercauteren, T.: Fastgeodis: Fast generalised geodesic distance transform. Journal of Open Source Software **7**(79), 4532 (Nov 2022). https://doi.org/10.21105/joss.04532, http://dx.doi.org/10.21105/joss.04532
2. Aydin, O.U., Taha, A.A., Hilbert, A., Khalil, A.A., Galinovic, I., Fiebach, J.B., Frey, D., Madai, V.I.: On the usage of average hausdorff distance for segmentation performance assessment: hidden error when used for ranking. European radiology experimental **5**, 1–7 (2021)
3. Aylward, S.R., Bullitt, E.: Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. IEEE transactions on medical imaging **21**(2), 61–75 (2002)
4. Cafaro, A., Spinat, Q., Leroy, A., Maury, P., Munoz, A., Beldjoudi, G., Robert, C., Deutsch, E., Grégoire, V., Lepetit, V., Paragios, N.: X2Vision: 3D CT Reconstruction from Biplanar X-Rays with Deep Structure Prior. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (2023)
5. Cipra, B.A.: An introduction to the ising model. The American Mathematical Monthly **94**(10), 937–959 (1987)
6. Copeland, A.D., Mangoubi, R.S., Desai, M.N., Mitter, S.K., Malek, A.M.: Spatiotemporal data fusion for 3d+ t image reconstruction in cerebral angiography. IEEE transactions on medical imaging **29**(6), 1238–1251 (2010)
7. Frisken, S., Haouchine, N., Du, R., Golby, A.J.: Using temporal and structural data to reconstruct 3d cerebral vasculature from a pair of 2d digital subtraction angiography sequences. Computerized Medical Imaging and Graphics **99**, 102076 (2022)
8. Hammersmith Hospital London: IXI Dataset: Brain Development. https://brain-development.org/ixi-dataset/

9. Haouchine, N., Juvekar, P., Xiong, X., Luo, J., Kapur, T., Du, R., Golby, A., Frisken, S.: Estimation of high framerate digital subtraction angiography sequences at low radiation dose. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VI 24. pp. 171–180. Springer (2021)
10. 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)
11. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. In: arXiv (2014)
12. Maas, K.W., Pezzotti, N., Vermeer, A.J., Ruijters, D., Vilanova, A.: Nerf for 3d reconstruction from x-ray angiography: Possibilities and limitations. In: VCBM 2023: Eurographics Workshop on Visual Computing for Biology and Medicine. pp. 29–40. Eurographics Association (2023)
13. Paetzold, J.C., Shit, S., Ezhov, I., Tetteh, G., Ertürk, A., Munich, H.Z., Menze, B.: cldice—a novel connectivity-preserving loss function for vessel segmentation. In: Medical Imaging Meets NeurIPS 2019 Workshop (2019)
14. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Others: PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: NeurIPS (2019)
15. Ruedinger, K., Schafer, S., Speidel, M., Strother, C.: 4d-dsa: development and current neurovascular applications. American Journal of Neuroradiology **42**(2), 214–220 (2021)
16. Settecase, F., Rayz, V.L.: Advanced vascular imaging techniques. Handbook of Clinical Neurology **176**, 81–105 (2021)
17. Wu, S., Kaneko, N., Mendoza, S., Liebeskind, D.S., Scalzo, F.: 3d reconstruction from 2d cerebral angiograms as a volumetric denoising problem. In: International Symposium on Visual Computing. pp. 382–393. Springer (2023)
18. Zhao, H., Zhou, Z., Wu, F., Xiang, D., Zhao, H., Zhang, W., Li, L., Li, Z., Huang, J., Hu, H., et al.: Self-supervised learning enables 3d digital subtraction angiography reconstruction from ultra-sparse 2d projection views: a multicenter study. Cell Reports Medicine **3**(10) (2022)
19. Zuo, J.: 2d to 3d neurovascular reconstruction from biplane view via deep learning. In: 2021 2nd International Conference on Computing and Data Science (CDS). pp. 383–387. IEEE (2021)