A Deep Learning Approach for Direct Mesh Reconstruction of Intracranial Arteries

Major Research Project
MSc Medical Imaging

Raquel González López

Daily supervisors: Iris N. Vos and Dr. Jelmer M. Wolterink
Examiner: Dr. ir. Koen L. Vincken
Second reviewer: Dr. Alberto de Luca

Utrecht, 10th March 2024

UMC Utrecht

Universiteit Utrecht

A Deep Learning Approach for Direct Mesh Reconstruction of Intracranial Arteries

Raquel González López

Image Sciences Institute
University Medical Center Utrecht (UMCU)
r.gonzalezlopez@students.uu.nl

Abstract—The Circle of Willis (CoW) is a group of vessels connecting major circulations of the brain. Its vascular geometry is believed to influence the onset and outcome of serious neurovascular pathologies. These geometric features can be obtained from surface meshes to capture vessel topology and morphology. A recent deep learning technique to handle non-Euclidean data, such as meshes, is Geometric Deep Learning (GDL). To this end, this study aimed to explore a GDL-based approach to directly reconstruct surface meshes of the CoW from magnetic resonance angiography images, thereby eliminating the traditional postprocessing steps required to obtain such a mesh from volumetric representations. The network architecture includes both convolutional and graph convolutional layers, allowing it to operate with images and meshes at the same time. It takes as input an image volume and a template mesh and outputs a 3D surface mesh. Experiments were performed on five crops representing different vessels and bifurcations to capture both stability and variability within the CoW. The results showed that anatomy-specific template input meshes and enhancement of the image feature representation increase the accuracy of the reconstruction. Moreover, incorporating the curvature characteristics of the meshes showed promising capability of handling complex geometries and sharp edges. However, achieving a consistent performance across CoW regions remains a challenge.

Index Terms—Circle of Willis, magnetic resonance angiography, deep learning, vascular geometry, surface mesh reconstruction, graph convolutional networks

I. INTRODUCTION

The Circle of Willis (CoW) is a group of connected vessels located in the base of the brain forming a circulatory anastomosis [1]. Several arteries comprise the CoW, including the anterior communicating, anterior cerebral, internal carotid, posterior communicating, posterior cerebral, and basilar arteries. The CoW has the ability to redistribute cerebral blood flow in case of impaired or decreased flow through one or more of its proximal feeding vessels [2]. This ability is considerably affected by its morphology and topology, which presents anatomical variations among 70% of the population [3]. The different configurations of the CoW and sharper bifurcation angles between its arteries have previously been identified as morphological imaging markers predictive of various neurovascular pathologies, such as aneurysms [4]. Recent studies have focused on understanding how the vascular geometry of the CoW influences the onset and outcome of these pathologies [5, 6].

These geometric features that capture vessel topology and shape can be obtained using Geometric Deep Learning (GDL). Rather than grid-structured data, such as images, GDL is a technique capable of handling non-Euclidean data, for example, graphs or meshes. These meshes can be roughly described as piece-wise planar approximations of a surface composed of nodes, edges, and faces. These components are often discretized as triangular meshes and used for modeling 3D objects or structures [7]. Using such mesh representations, we can derive characteristics for perceptual parsing of 3D surfaces, in this case, vessel surfaces [8]. Meshes have been successfully applied to classification and segmentation problems because of their ability to handle topology and shape information. Overall, these 3D representations are increasingly applied in clinical applications, including physics-based simulations, to study vascular hemodynamics [9].

Recent studies have shown promising precision in reconstructing meshes, but they are highly dependent on the performance of voxel-wise segmentation and resolution of the image data. This segmentation is then usually followed by surface reconstruction and postprocessing techniques, such as smoothing or morphological operations. Segmentation of the intracranial arteries is typically performed using magnetic resonance angiography (MRA) images, which is a non-invasive imaging modality to visualize patient-specific vasculature [10–13]. However, these segmentations may still include disconnected regions, such as missing small vessels, which results in an incorrect surface topology representation or vessels merging. Furthermore, the presence of noise caused by staircase artifacts prevents segmentations from having contiguous or sub-voxel accuracy [14]. In current literature, there is little focus on converting the segmented volumes to mesh; therefore, improving mesh quality could enhance the topology and shape information.

The next step in reconstructing meshes, surface reconstruction, also faces challenges, particularly when applied to intracranial arteries. Currently, there are two main approaches for surface reconstruction: centerline-based and template-based. The first approach involves a two-step process that reconstructs the vascular surface from centerlines. As centerlines are usually extracted from segmentations, similar issues arise. This limitation echoes the challenges mentioned earlier in segmentations [15, 16]. The second approach, the template-based, gained

recent interest with new neural network architectures enabling direct 3D mesh reconstruction from volumetric images. These architectures are based on the diffeomorphic deformation of an input template to the target surfaces, meaning that the template mesh and image are taken as input and the mesh is altered iteratively by learning the deformation field of its vertices [17–20]. While these methods have been applied to simpler structures such as abdominal organs (i.e., the liver) or the hippocampus, the anatomy of the intracranial arteries presents a more complex and variable challenge.

To overcome the previously outlined shortcomings, such as segmentations and centerlines inaccuracies, we explore using a GDL-based approach to predict surface meshes of the CoW bifurcations directly. To achieve this, we will use the template-based approach, consisting of a graph convolutional network (GCN) to predict mesh deformation given a specified template input mesh. More specifically, this GCN is combined with a convolutional neural network (CNN) based image encoder to extract the features from the MRA images. To analyze the model's adaptability, it will be applied to various bifurcations within the CoW.

II. METHODS

A. Dataset

To obtain information from the different components of the CoW, Time of Flight MRA images were used. The open-source dataset was provided by the "Topology-Aware Anatomical Segmentation of the Circle of Willis for CTA and MRA" challenge or TopCoW for short [21]. It was organized in association with the Medical Imaging and Computer Assisted Intervention 2023. The data cohort is composed of patients admitted to the Stroke Center of the University Hospital Zurich. MRA images were acquired by SIEMENS Skyra or Avanto Fit model, with: magnetic field strength of 3 or 1.5 Tesla, TOF-3D or TOF-3D multi slab mode, voxel size between 0.29 to 0.35 mm in X-Y dimension, and between 0.5 to 0.6 in the Z dimension.

The patients of the data cohort were admitted for or recovering from a stroke-related neurological disorder. The dataset consists of 90 patients with manual multilabel annotations. These annotations only included vessel components and regions necessary to assess the CoW anatomy and geometry. These vessel components are the left and right internal carotid artery (ICA), left and right anterior cerebral artery (ACA), left and right anterior communicating artery (Acom), left and right posterior communicating artery (Pcom), left and right posterior cerebral artery (PCA), and basilar artery (BA). A schematic representation of the CoW is shown in Fig. 1.

As previously mentioned, the CoW morphology and topology are highly variant. This results in certain principal artery components being hypoplastic or absent [2]. The TopCoW dataset includes a wide range of CoW variants, among which: with or without Acom, double Acom, with or without Pcom, triple ACA, aplastic or hypoplastic vessel segments of the ACA and PCA, vessel fenestrations, etc. This last refers to splits
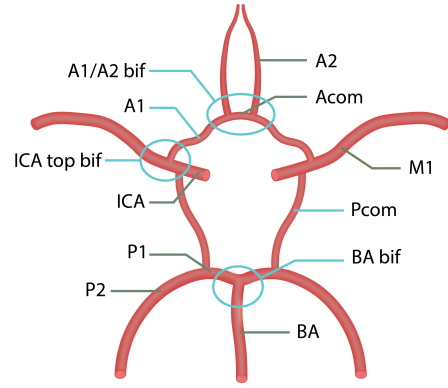

Fig. 1. Schematic representation of the Circle of Willis. The abbreviations are as follows: A1/A2 are segments of the anterior cerebral artery (ACA); Acom is the anterior communicating artery; M1 is the segment of the middle cerebral artery at 50% of total length; P1/P2 are segments of the posterior cerebral artery (PCA); Pcom is the posterior communicating artery; ICA is the internal carotid artery; BA is the basilar artery. The vessels of interest (A1 and Pcom) and bifurcations of interest (A1/A2, ICA top, and BA top) are marked by blue circles.

or openings in the artery walls that can significantly affect vascular flow dynamics [21].

B. Preprocessing

All images were preprocessed by padding to a uniform size of 576x640x256 to standardize image dimensions, and Z-score normalization to adjust for intensity variations. To acquire different representations of the CoW and address the anatomical variability and complexity of the CoW, the MRA images were cropped into different segments. First, two types of vessels were selected to represent these anatomical differences: the lower segment of the ACA (A1), which exhibits minimal anatomical variation across individuals, and the Pcom, which shows large variability. Second, three bifurcations were chosen to capture both stability and variability within the CoW. The ICA top and BA top bifurcations were selected for their consistent anatomical presence across individuals, while in contrast, the A1/A2 bifurcation was used for its anatomical variation and as an indicator of the complex interconnections within the CoW (see Fig.1).

To analyze each of the selected vessels and bifurcations, the preprocessed images were cropped at the specific locations of interest. To extract the vessel locations, a skeletonization method based on a distance transform was utilized [22]. The resulting vessel skeletons were converted into graph representations with nodes. The coordinates of these nodes were employed to compute the vessel center of mass and create a crop of 32x32x32 around that location. A similar approach was used for the bifurcations, instead of using the center of mass of the nodes in the graph, the top node of the graph was used as the center of the crop. Then, a connected components analysis was applied to discard small voxel clusters from close

vessels. This process was performed in MevisLab (version 3.4.2) [23].

Next, the dataset was split into three subsets: training, validation, and testing. The first 80% of the data (72 cases) was allocated for training, while the subsequent 10% (9 cases) was used for validation and the remaining 10% (9 cases) for testing. For the A1 and Pcom vessels, and the ICA top bifurcation, both left and right structures were considered. It should be noted that due to the data variety, such as missing vessels, not all sets included the totality of the dataset.

C. Network architecture

Following the extraction and preprocessing of the vessel structures, the next section delineates the employed network architecture. The adopted baseline framework is based on the Voxel2Mesh [17]. It takes as input a 3D image volume, such as the MRA crops, its corresponding ground truth, and a sphere template mesh. It consists of two components: (1) a CNN that operates on voxels and extracts features from the input volume, and (2) a GCN that predicts the deformation of the input sphere mesh vertices. Both networks are combined by sampling the features extracted from the CNN into vertex locations of the GCN.

1) Voxel encoder and decoder: The CNN consists of a voxel encoder and decoder that are based on a standard 3D U-Net architecture [24]. Briefly, it has four resolution steps in both the voxel decoder and encoder. Each layer in the encoder contains two 3x3x3 convolutions, with batch normalization (BN), and a rectified linear unit (ReLU), which are followed by a 2x2x2 max pooling. The decoder consists of a 2x2x2 transpose convolution, followed by the same convolutions, BN, and ReLu on each resolution step.

2) Mesh deformation: The next module of the architecture consists of a GCN that iteratively deforms the input mesh vertices in four steps by using the features extracted from the CNN voxel decoder. The input of the mesh decoder is a 3D sphere mesh. Formally, each graph convolution could be defined as:

$$\mathbf{f'} = w_1\mathbf{f} + \frac{1}{\mathcal{N}(\mathbf{v}_l)} \sum_{\mathbf{v}_l^i \in \mathcal{N}(\mathbf{v}_l)} \mathbf{f}^i w_2 e^{-d_i^2/\sigma^2} \quad (1)$$

where $\mathbf{f'}$ and $\mathbf{f}$ represent the feature vector associated with the 3D vertex coordinates $\mathbf{v}_l$ after each mesh decoder block l. $\mathcal{N}(\mathbf{v}_l)$ is the set of vertex neighbors and $\mathbf{f}^i$ is the feature vector of its corresponding neighbor vertex $\mathbf{v}_l^i$. d_i represents the Euclidean distance between $\mathbf{v}_l^i$ and $\mathbf{v}_l$. w_1, w_2, and σ are weights learned during training.

Then, the two decoders communicate in all four resolution levels, progressively enhancing precision. To transition features extracted from the voxel decoder in continuous 3D space, the feature maps are interpolated trilinearly from discrete voxels. The network also includes adaptive mesh sampling and unpooling mechanisms. This means that only the necessary mesh vertices and their neighbors were used to refine the output mesh. Therefore, the network learns to obtain optimum sampling locations. In this way, the mesh refinement is more precise and densely sampled in high-curvature areas but not elsewhere, thereby reducing the amount of memory needed for the computation.

3) Architecture modifications: In this study, several modifications were made to the baseline architecture to improve the feature representation and adaptability to different types of data. Briefly, the voxel decoder and encoder were enhanced with residual blocks, which allow the gradient propagation during training through the network more easily and help improve generalization. Each block consists of 3x3x3 convolutions and a dropout layer before the last convolutional layer. In addition, Leaky ReLU nonlinearities were used for all feature map convolutions [25–27]. Moreover, based on Rundo *et al.* [28], Squeeze-and-Excitations (SE) blocks were included to exploit the adaptive channel-wise feature recalibration to boost the generalization performance [29]. Lastly, the feature maps from both the voxel decoder and encoder were connected with the mesh deformation module. This enables the network to identify relevant information throughout the training phase, which is not achievable with a limited set of feature maps [19]. Fig. 2, depicts the modified architecture.

D. Mesh initialization

Current techniques for explicit mesh reconstruction techniques begin with a generic template mesh, such as a 3D sphere. Alternatively, to enhance the reconstruction of CoW vessels and bifurcation crops, anatomy-specific templates were used. This approach tailors the initial mesh to the specific anatomical features of interest, as demonstrated in studies [19, 20]. This could help to reduce the distance between the predictions and ground truths while also avoiding large deformation during the initial stages of training. To obtain a general anatomy of the CoW, the Forkert *et al.* [30] MRA atlas was used. A vesselness filter was applied [31], which indicates how similar a structure is to a tube, and in this way detects the bright vessels from the averaged background. The center locations of the desired vessels and bifurcations were selected in the atlas by visual inspection and used as the center to create a 32x32x32 crop of the image. Next, a threshold interval region growing algorithm was used to extract the mask. However, this was not possible for the Pcom vessel, since it is often underdeveloped or missing. Consequently, the intensity values of the MRA atlas for this region were blended with those of the background. In this specific case, the mask was extracted by manual segmentation. Then, an isosurface of each of the desired structures was generated with a Neighboring Cells algorithm [23] and the number of vertices was reduced until approximately 400 vertices to reduce computational memory. Laplacian smoothing was applied to the resulting surfaces until the surface did not change anymore [32]. Lastly, the templates were normalized by translating their centroid to the origin (0, 0, 0), and scaled to fit inside a unit sphere [33].
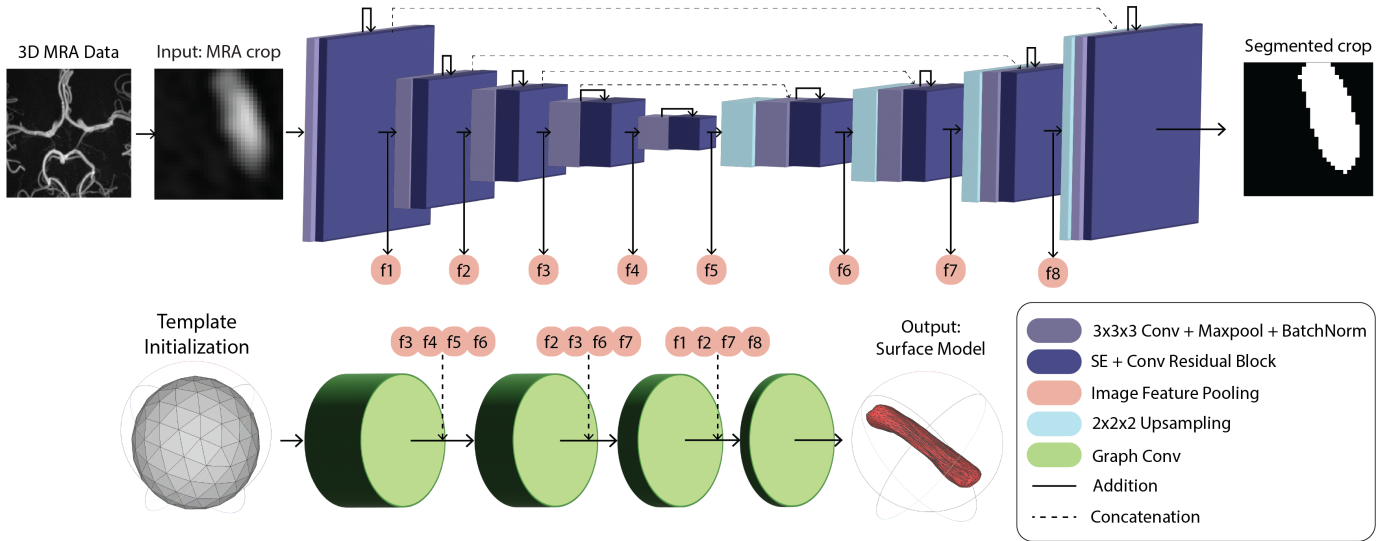
Fig. 2. Modified Voxel2Mesh network pipeline. The architecture takes as input a 3D cropped vessel or bifurcation and a template mesh. It predicts a voxel-wise segmentation and surface meshes. It is composed of a CNN that extracts image features and communicates at each level to a GCN decoder to deform the template mesh. At each step of the mesh decoding, it receives features from both the encoder and decoder of the CNN. The mesh is deformed by adding vertices only where needed.

E. Loss functions

The training procedure was supervised by the binary ground truth segmentations from the TopCow dataset and their corresponding 3D ground truth meshes. These meshes were obtained by applying the marching cube algorithm [28]. Several loss functions were used during training. Firstly, the cross-entropy loss $\mathcal{L}_{ce}$ was used for the segmentation module and can be defined as:

$$\mathcal{L}_{ce}(y_P, y_G) = -(y_G \log(y_P) + (1 - y_G) \log(1 - y_G)) \quad (2)$$

where y_G denotes the ground truth binary segmentation and y_P the predicted probability map. This loss returns a probability value between 0 and 1, in which the higher the loss, the higher the predicted probability diverges from the ground truth.

Secondly, the mesh loss is a combination of geometry consistency losses, which include Chamfer distance and normal consistency losses, and regularization losses, which are the edge length and Laplacian losses. These losses help to improve convergence and smooth the output mesh. Regarding the Chamfer loss, it is a popular evaluation criterion to determine the similarity between mesh vertices. Chamfer loss assigns the nearest vertex in the other point cloud for a point from the ground truth or the predicted mesh [34]. Then, all the distances are averaged to get the similarity between the two meshes. It can be defined as:

$$\mathcal{L}_{cf}(P_i, G_i) = \sum_{p \in P_i} \min_{g \in G_i} ||p - g||_2^2 + \sum_{g \in G_i} \min_{p \in P_i} ||p - g||_2^2 \quad (3)$$

where p and q are vertex from the predicted mesh P_i and ground truth mesh G_i, respectively.

A variation of the standard Chamfer loss was introduced by Bongratz *et al.* [19] to reduce the smoothing effect of the other regularization loss terms that can lead to lower geometric accuracy. To solve this, they proposed a curvature-weighted Chamfer loss to emphasize high-curvature regions. For this, consider $\kappa(p) \in [1, \kappa_{max}]$ to be the point weights based on the local discrete mean curvature [35, 36], where $\kappa_{max} = 5$. Then, the curvature-weighted Chamfer loss can be described as:

$$\mathcal{L}_{\mathbf{cwcf}}(P_i, G_i) = \sum_{p \in P_i} \kappa(u) \min_{g \in G_i} ||p - g||_2^2 + \sum_{g \in G_i} \kappa(\tilde{u}) \min_{p \in P_i} ||p - g||_2^2 \quad (4)$$

where $\tilde{u} = argmin||p - r||^2$. In this regard, the loss pulls predicted points in high-curvature regions closer to their correct position than points in low-curvature regions. Only ground truth weights were considered to avoid possible inaccurate curvature from the predictions, which could be deceptive.

Next, the normal consistency loss computes the surface normal, which is obtained from the cross product between two edges (in this case, p_1 and p_2) of a face connected to vertex p. This effectively aligns surface normals with ground truth observations to improve geometric consistency, even on non-planar surfaces. It is defined as follows:

$$\mathcal{L}_n(P_i, G_i) = \sum_{p \in P_i : \arg\min_{g \in G_i} ||p-g||_2^2} ||(p_1-p) \times (p_2-p) - n_g||_2^2 \quad (5)$$

where n_g is the observed surface normal from the ground truth.

Regarding the regularization losses, the Laplacian loss prevents mesh self-intersection by restricting excessive vertex

mobility. This loss serves as a local detail-preserving operator by ensuring a consistent movement among its neighboring vertices. To compute this loss, the difference between a vertex p location and the mean location of this neighboring vertices k_p are used as follows:

$$\mathcal{L}_{lap}(P_i) = \sum_{p \in P_i} \left\| p - \sum_{k_p \in \mathcal{N}(p)} \frac{1}{\|\mathcal{N}(p)\|} k_p \right\|_2^2 \quad (6)$$

The last term, the edge length loss is used to penalize flying vertices and to encourage a more uniform predicted mesh density. It is calculated from the difference between each edge length between a vertex p and its neighboring vertices k_p for both the ground truth and the predicted mesh. The average edge length is then the ratio between these meshes, such as:

$$\mathcal{L}_{edge}(P_i) = \sum_{p \in P_i} \sum_{k_p \in \mathcal{N}(p)} \|p - k_p\|_2^2 \quad (7)$$

As a result, the total loss is a weighted sum of all five losses described, which can be written as:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{ce} + \sum_{l=1}^{L} \lambda_2 \mathcal{L}_{cf}^l + \lambda_3 \mathcal{L}_n^l + \lambda_4 \mathcal{L}_{lap}^l + \lambda_5 \mathcal{L}_{edge}^l \quad (8)$$

where L is the number of stages in the mesh decoder. Each loss term is assigned a hyperparameter λ, $\lambda_{1,2,3} = 1$, and $\lambda_{4,5} = 0.1$, which were based on literature [17].

F. Experiments

Different experiments were performed to assess the adaptability of the model to the CoW morphology and topology. The Voxel2Mesh network is used as a baseline for the experiments, described in II-C1 and II-C2.

1) Comparing mesh initialization templates: The first experiment is to compare if using an anatomy-specific template improves the model performance, in comparison to the basic 3D sphere. For each desired vessel and bifurcation, their corresponding smoothed templates from II-D were introduced to the Voxel2Mesh network, instead of the 3D sphere.

2) Comparing model modifications: The second experiment is used to assess whether the modifications from the baseline architecture, described in II-C3, improve the overall performance for each of the desired vessels and bifurcation. Especially, the change in information exchange between the models. The combination of a CNN and GCN raises the dilemma of how to transfer information from one subnetwork to another, in this case from the CNN to the GCN. It is possible to employ feature maps derived from either the encoder [18], the decoder [21], or both [19].

3) Comparing curvature-weighted Chamfer loss: The third, and last experiment, consists of using the curvature-weighted Chamfer loss, explained in II-E, to assess if it helps to obtain a more accurate reconstruction in high-curvature areas. For this, the discrete mean curvature of the ground truth meshes is obtained from the cotangent Laplacian and used as point weights of the loss. The modified model is used for this case.

G. Training strategy

To perform fair comparisons, all models were trained with the same parameters. The baseline model was trained for three different random initializations to assess its effect on the training procedure. Then, a fixed seed was employed for both experiments. All cases were trained for a maximum of 9000 epochs, with 300 iterations each, evaluating every 100 iterations. Training is done using the Adam optimizer [37], with an initial learning rate of 10^{-4}. To optimize the learning, a learning rate scheduler was used, reducing the learning rate by a factor of 0.1 upon no improvement in validation loss. Furthermore, early stopping is enforced to cease training upon no improvement in validation loss to ensure efficiency and prevent overfitting. The best epoch was selected based on the validation set.

Data augmentation techniques, based on Wickramasinghe *et al.* [17], were used during training to increase the robustness of the model to fluctuations in input images. These techniques included: axis permutation, flips, quaternion-based 3D rotations, translations, and scaling. They were randomly applied with a 50% probability, resulting in a composite transformation.

H. Evaluation metrics

To evaluate the performance of the different models, the Dice Similarity Coefficient (DSC), Jaccard Index (JAC), and Chamfer weighted symmetric distance, from now on Chamfer distance, were calculated between the ground truth segmentations and the predicted meshes. The DSC quantifies the overlap between two volume segmentations and is defined as twice the intersection of the volumes divided by their union. JAC is then the intersection between the volume segmentations divided by their union [38]. Both DSC and JAC measure the overlap between volumes, in this context, they were both used to be able to compare with the baseline and current literature results.

I. Statistical Analysis

To evaluate the differences between the models, a statistical analysis was performed. First, the Shapiro-Wilk test was used to confirm the non-parametric distribution of the samples [39]. Since the data is paired, a Wilcoxon-ranked t-test [40] is suitable. To test the null hypothesis, the significance level established was $\alpha = 0.05$, meaning that if the p-value is lower than this, the null hypothesis would be rejected. Since multiple model comparisons with the Wilcoxon-ranked t-test are being performed, Bonferroni adjustment was used. It is applied to limit the possibility of getting a statistically significant result when testing multiple hypotheses.

J. Implementation details

All models are based on PyTorch (v1.11.0) [41], PyTorch3d (v0.7.2) [42]. They were trained and evaluated on Nvidia GP102 Titan X (12GB) using CUDA version 11.3. The experiments were tracked with Weights and Biases [43].

III. Results

In this section, the results of the experiments are presented. Only the test results are reported. Table I summarizes the results in terms of DSC, JAC, and Chamfer distance. As a reference, the performance of the baseline model for the CHAOS dataset [17, 44], with the same reduced settings as the rest of the models, had a mean JAC of 0.86 and Chamfer distance of $0.26 \cdot 10^{-2}$. Then, the median DSC of the three random initializations for each data structure was in the following range: 0.55-0.61 for the A1 vessel, 0.47-0.55 for the Pcom vessel, 0.70-0.75 for the A1/A2 bifurcation, 0.85-0.86 for the ICA top bifurcation and 0.84-0.87 for the BA top bifurcation. The ICA top and BA top bifurcations have a more stable performance than the rest of the structures.

A. Comparing mesh initialization templates

The 3D sphere and the generated atlas templates are shown in Fig. 3. The atlas templates have a considerably higher number of vertices, increasing from 162 for the sphere to approximately 400. This difference also impacted the training time, which increased notably. Regarding the analytical results, as shown in Table I, the atlas template achieved a higher DSC overall compared to the sphere templates, except for the A1 vessel, where the DSC was significantly lower ($p = 0.02$). However, the ICA top bifurcation resulted in a significantly higher DSC ($p < 0.001$). These differences can also be seen in Fig. 4. Correspondingly, the JAC values followed a similar trend. In addition, the Chamfer distance was consistently lower for the atlas templates across all vessels except for the A1.
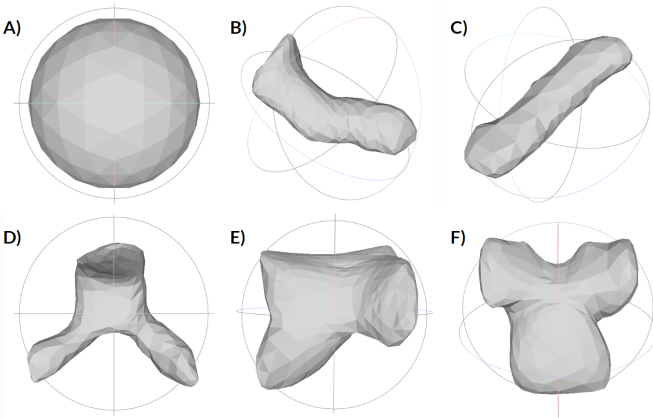

Fig. 3. General and anatomy-specific mesh initialization templates: A) 3D Sphere, B) A1 and C) Pcom vessels, D) A1/A2, E) ICA top, and F) BA top bifurcations.
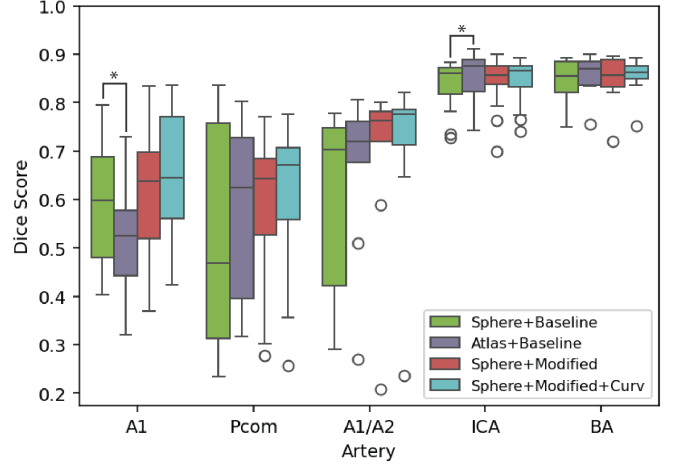

Fig. 4. Boxplot of the DSC values for the different models. '*' indicates that the model to the right is significantly better than the model to the left ($p < 0.05$).

B. Comparing model modifications

The modified model outperformed the baseline model for all cases, except for the ICA top bifurcation where the DSC median remains unchanged at 0.86, although the interquartile range increased slightly (see Table I). There is a considerable increase in performance for the A1 vessel (0.64), Pcom vessel (0.64), and A1/A2 bifurcation (0.76), which was also higher than the atlas template baseline model. The best and worst performances of the modified model can be seen in Fig. 5. Similarly, the JAC also experienced an increase in the cases where DSC improved, stayed consistent for the BA top bifurcation at 0.75, and experienced a minor reduction for the ICA top bifurcation. In contrast, there was not a corresponding substantial reduction in the Chamfer distance for the first three scenarios.

C. Comparing curvature-weighted Chamfer loss

The modified model with the curvature-weighted Chamfer loss resulted in a higher performance in terms of DSC for all cases, except the BA top bifurcation which remained unchanged (0.86), as can be seen in Table I and visually in Fig. 4. This is in comparison to the same model utilizing the standard loss function. In Fig. 6, it can be seen that the curvature-weighted Chamfer loss model produces a closer curvature to the ground truth in areas of high curvature for the A1/A2 bifurcation. Moreover, the JAC values slightly increased for all cases. The Chamfer distance values were also reduced for each case when employing the curvature-weighted loss.

IV. Discussion

In the present study, we evaluate the efficacy and adaptability of a GDL-based approach in modeling surface meshes of selected CoW vessels and bifurcations. Our results highlight that using anatomy-specific templates can help reduce the distance between predictions and ground truths, preventing large deformations during the early stages of training, which

		A1	Pcom	A1/A2	ICA top	BA top
DSC ↑	*Sphere + Baseline*	0.60 [0.48, 0.69]	0.47 [0.31, 0.76]	0.70 [0.42, 0.75]	0.86 [0.82, 0.87]	0.85 [0.82, 0.89]
	Atlas + Baseline	0.53 [0.44, 0.58]	0.62 [0.40, 0.73]	0.72 [0.68, 0.76]	0.88 [0.82, 0.89]	0.87 [0.84, 0.89]
	Sphere + Modified	0.64 [0.52, 0.70]	0.64 [0.53, 0.68]	0.76 [0.72, 0.78]	0.86 [0.84, 0.88]	0.86 [0.83, 0.89]
	Sphere + Modified + Curv	0.65 [0.56, 0.77]	0.67 [0.59, 0.71]	0.78 [0.71, 0.79]	0.87 [0.83, 0.88]	0.86 [0.85, 0.88]
JAC ↑	*Sphere + Baseline*	0.43 [0.32, 0.53]	0.31 [0.19, 0.61]	0.54 [0.27, 0.60]	0.76 [0.69, 0.77]	0.75 [0.70, 0.79]
	Atlas + Baseline	0.36 [0.29, 0.41]	0.43 [0.25, 0.57]	0.56 [0.51, 0.62]	0.78 [0.70, 0.80]	0.77 [0.72, 0.80]
	Sphere + Modified	0.47 [0.35, 0.54]	0.47 [0.36, 0.52]	0.62 [0.56, 0.64]	0.75 [0.72, 0.78]	0.75 [0.71, 0.80]
	Sphere + Modified + Curv	0.48 [0.39, 0.63]	0.51 [0.39, 0.55]	0.64 [0.55, 0.65]	0.76 [0.71, 0.78]	0.76 [0.74, 0.78]
Cf $\cdot 10^{-2}$ ↓	*Sphere + Baseline*	1.67 [0.71, 2.82]	2.55 [0.97, 3.00]	1.31 [0.95, 2.84]	0.53 [0.39, 1.89]	0.61 [0.39, 1.11]
	Atlas + Baseline	2.28 [1.15, 3.95]	1.26 [0.61, 2.19]	0.79 [0.59, 1.04]	0.46 [0.27, 1.75]	0.40 [0.26, 1.29]
	Sphere + Modified	1.60 [0.94, 3.57]	2.25 [1.60, 2.85]	0.79 [0.60, 1.17]	0.49 [0.33, 1.53]	0.60 [0.41, 1.68]
	Sphere + Modified + Curv	1.00 [0.50, 2.81]	2.09 [1.10, 2.62]	0.63 [0.55, 1.00]	0.48 [0.35, 1.50]	0.54 [0.37, 0.65]

First, the baseline model (Sphere + Baseline), then the atlas template mesh initialization (Atlas + Baseline) experiment, the model modifications (Sphere + Modified) experiment, and lastly, adding the curvature-weighted Chamfer loss to the modified model (Sphere + Modified + Curv) experiment. The vessel and bifurcations abbreviations are as follows: A1 is the lower segment of the anterior communicating artery (ACA), Pcom is the posterior communicating artery, A1/A2 is the bifurcation between the two segments of the ACA, ICA top is the bifurcation at the top of the internal carotid artery, and the BA top is the bifurcation at the top of the basilar artery. Then, the evaluation metrics abbreviations are: DSC is the Dice Similarity Coefficient, JAC is the Jaccard Index, and Cf is the Chamfer distance.
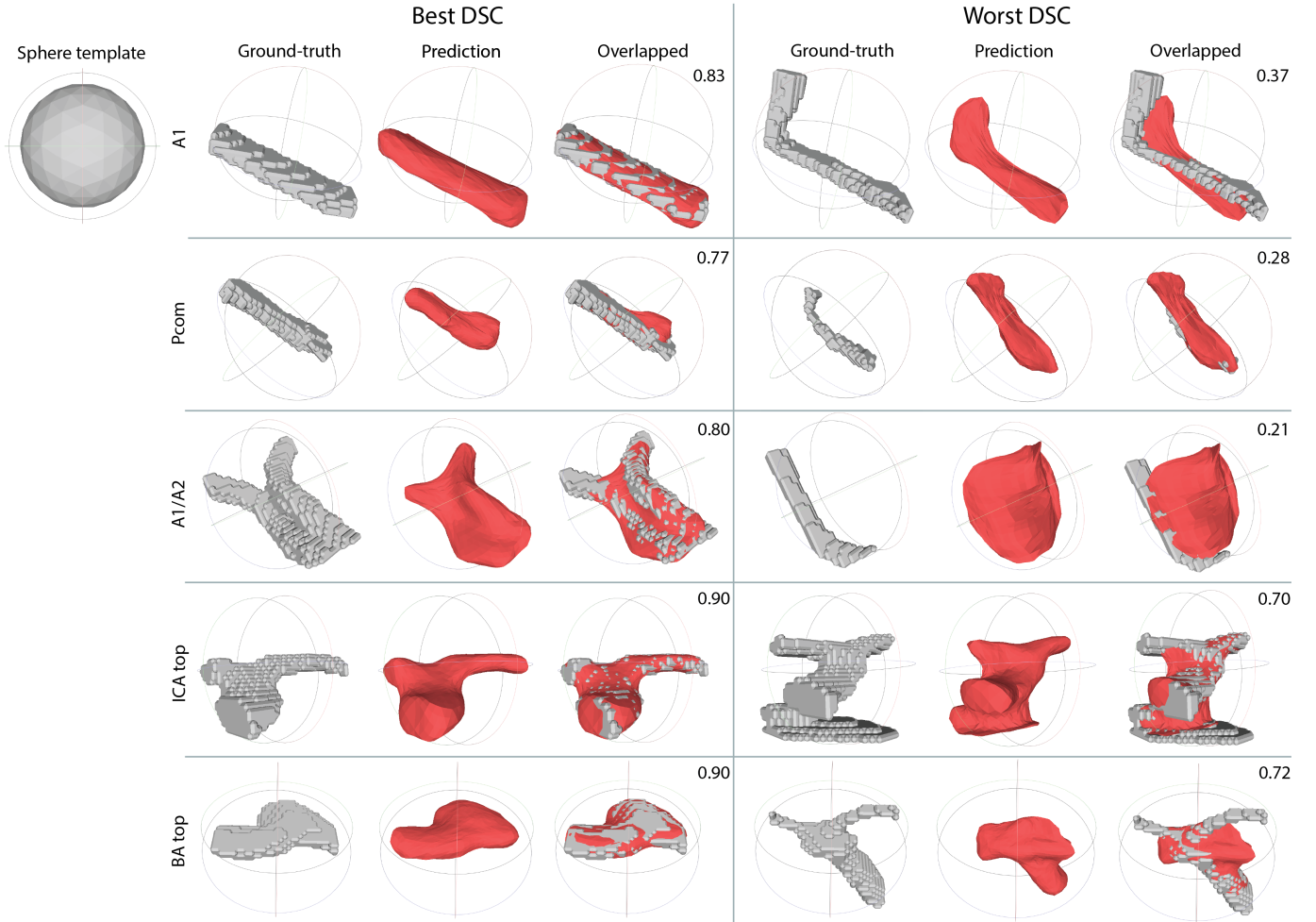

Fig. 5. Visualization of the best and worst reconstruction results of the modified model for each case, in terms of the Dice Score Coefficient (DSC). The sphere template is used as a reference. The vessel and bifurcations abbreviations are as follows: A1 is the lower segment of the anterior communicating artery (ACA), Pcom is the posterior communicating artery, A1/A2 is the bifurcation between the two segments of the ACA, ICA top is the bifurcation at the top of the internal carotid artery, and the BA top is the bifurcation at the top of the basilar artery.
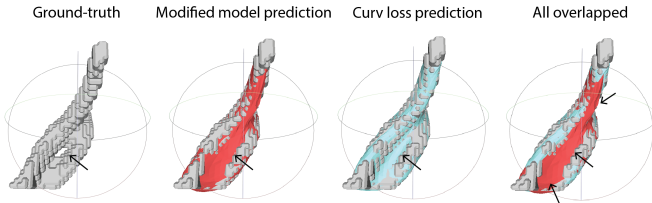
Fig. 6. Visualization of the A1/A2 bifurcation ground truth and its corresponding predicted mesh of the modified model with the standard Chamfer loss versus the modified model with the curvature-weighted Chamfer loss. Both predictions are shown overlapped on the ground truth to observe their adaptability to the bifurcation and compare their performance. The arrows highlight specific high-curvature regions and the presence of a fenestration.

can avoid local minimums and achieve faster convergence. Furthermore, enhancing the feature representation and providing features from both the voxel decoder and encoder is most effective, as was also suggested in other studies [20, 45].

Almost all cases yielded improvements with the atlas template, likely due to reduced initial deformations and a greater vertex count. However, the A1 template did not improve the results. This could be affected by the variability in the crops. Since the ACA vessel was labeled as a whole, no distinction was made for the A1 and A2 segments. This issue led to crops including parts of the highly curved A1/A2 bifurcation, as can be seen in the worst DSC case for A1 in Fig. 5. This suggests that the sphere template mesh could be beneficial for more variable data since it provides a more general shape for the initial deformation stages.

The overall performance of the baseline model was considerably lower for the single vessels. This could be due to the variability of the data. Notably, the single vessels were more sensitive to diameter changes and length. For instance, in the case of the Pcom, the diameter could be as low as 1mm, or as high as 2.5mm, which caused a huge range in metrics values. It should also be noted that the dataset used to train the Pcom was lower (88 cases) since it was missing in around 40% of the patients, and sometimes only present in one side. This did not happen for the main bifurcations, such as the ICA top and BA top. The observed variability for both bifurcations across patients was substantially lower than the Pcom, explaining the higher performance. Interestingly, there were twice as many data samples for the ICA top (left and right) compared to the BA top, but the achieved model performance was comparable. So, including more cases for the most variable structures, such as the A1 and Pcom vessels or the A1/A2 bifurcation, could boost model performance.

The modified model outperformed the baseline in Chamfer distance, with DSC scores either improving or remaining constant. In general, the model tends to round the sharp edges at the bifurcations, as shown in Fig. 5. In some cases, such as the A1/A2 bifurcation, the predicted output shape seems topologically different despite resulting in a high DSC value. This discrepancy suggests that a high DSC does not always correspond to an anatomically accurate representation.

Nevertheless, some cases deviate considerably from the median performance, which can be seen in the second column of Fig. 5. The worst DSC value for the A1 vessel seems to contain part of the A2 segment of the ACA in the crop. This resulted in a sharp edge, which was not present in the rest of the A1 training data. Similarly, in the A1/A2 worst DSC case, only a single vessel is present instead of both sides vessels. This, and the fact that it is far from the origin center, resulted in a bad prediction. It seems that the ACA was longer and more curved, which made them further from the bifurcation point. Since the connected component analysis was used for obtaining the crops, only one ACA was considered for the ground-truth crop. A better and optimized cropping strategy could have helped in these cases. Moreover, the model struggled with small diameter or hypoplastic vessels, particularly evident for the Pcom worst DSC, underscoring the impact of diameter variability and dataset size.

Introducing the curvature-weighted Chamfer loss slightly enhanced the performance of the modified model compared to using the standard loss, especially for the A1/A2 bifurcation (see Fig. 6), where higher curvatures were expected. This improvement is not only crucial for accurately capturing sharp edges in bifurcations, resulting in a more refined mesh, but it also helps mitigate the challenge of balancing accuracy and regularization terms in the loss function for highly curved regions in the field of deformation-based surface reconstruction methods [19, 45]. Furthermore, by incorporating the information of neighboring vertices, the GCN generates a mesh that is not only smoother but also more refined, enhancing the overall quality of the reconstructed surface.

This study presents several limitations. First, the model is unable to capture fenestrations, which can be seen in Fig. 6. This is because the genus-0 spherical template mesh and the generated anatomy-specific templates cannot capture holes in between the vessels. Second, this study has been performed on small crops of selected CoW vessels and bifurcations that were trained separately. The multi-site dataset contained a lot of variability with different CoW configurations. It would be interesting to assess the performance of the model on healthy controls and a broader dataset, to include more cases with smaller vessel diameters or hypoplastic vessels, which may improve its robustness. In this last case, other metrics like the Hausdorff distance could provide a more informative view than the DSC as it is sensitive to small disagreements in voxels [46]. Moreover, it would be interesting to train the modified model with the curvature-weighted Chamfer-loss and the anatomy-specific templates, since it was trained with the 3D sphere template mesh as input. Third, the computational resources were limited, since a lower number of vertices, fewer unpooling operations, and fewer sampling data points were used to train the model. For this, the baseline model with the CHAOS dataset was trained with this setting, and resulted in a similar JAC but a higher Chamfer distance, $0.26 \cdot 10^{-2}$, in comparison to $0.13 \cdot 10^{-2}$. Nevertheless, all of our models yielded a lower performance. Given these challenges, future

efforts aimed at addressing these limitations could potentially enhance the models' performance.

The assessment presented in this study represents the initial stages of a comprehensive effort to enhance 3D vascular modeling. Future directions could involve a previous detection of all the major CoW bifurcations, initialize different template meshes on those locations, and use a shared GCN to simultaneously deform and predict the vascular surface meshes on the multiple locations. This could also be followed by an algorithm to combine all generated meshes to obtain a whole mesh of the complete CoW [47]. In this way, the templates would only be initialized if a bifurcation is detected, which could help alleviate the problem of missing vessels. Some further modifications could be implemented into the modified model, such as deep supervision in the voxel decoder [10, 18, 19, 48], as well as attention gates to help focus on the target [49]. Another possibility would be to include gauge equivariance in the model, which could enhance the model's ability to accurately represent and reconstruct the complex geometrical features of the CoW by ensuring consistent and invariant responses to local transformations of the vascular structures [50].

V. CONCLUSION

In this study, we explored a GDL approach to directly predict surface mesh reconstructions of the vessels and bifurcations of the CoW, thereby eliminating the traditional postprocessing steps required to obtain such a mesh from volumetric representations. The architecture takes as input an image volume and a template mesh and outputs a 3D surface mesh. Our results highlight the model's adaptability to anatomy-specific templates and the effectiveness of enhancing the feature representation of the volumetric images. Furthermore, the employment of curvature-weighted Chamfer loss showed promising capability of handling complex geometries and sharp edges. Nevertheless, given the CoW's variability across regions, achieving consistent accuracy remains a challenge. Future work could potentially extend this approach with a CoW bifurcation detection model to initialize multiple template meshes and further create a complete surface reconstruction of the CoW vasculature. Such advancements could facilitate a deeper understanding of the CoW's geometric characteristics, crucial for assessing neurovascular pathologies.

REFERENCES

[1] B. J. Alpers, R. G. Berry, and R. M. Paddison, "Anatomical studies of the circle of willis in normal brain," *AMA Archives of Neurology Psychiatry*, vol. 81, no. 4, pp. 409–418, 1959.

[2] M. J. Krabbe-Hartkamp *et al.*, "Circle of willis: Morphologic variation on three-dimensional time-of-flight mr angiograms.," *Radiology*, vol. 207, no. 1, pp. 103–111, 1998, pmid:9530305.

[3] B. Eftekhar, M. Dadmehr, S. Ansari, M. Ghodsi, B. Nazparvar, and E. Ketabchi, "Are the distributions of variations of circle of willis different in different populations? – results of an anatomical study and review of literature," *BMC Neurology*, vol. 6, no. 1, p. 22, 2006, ID: Eftekhar2006. [Online]. Available: https://doi.org/10.1186/1471-2377-6-22.

[4] T. Horikoshi, I. Akiyama, Z. Yamagata, M. Sugita, and H. Nukui, "Magnetic resonance angiographic evidence of sex-linked variations in the circle of willis and the occurrence of cerebral aneurysms," *Journal of neurosurgery*, vol. 96, no. 4, pp. 697–703, 2002.

[5] J. Alastruey, K. H. Parker, J. Peiró, S. M. Byrd, and S. J. Sherwin, "Modelling the circle of willis to assess the effects of anatomical variations and occlusions on cerebral flows," *Journal of Biomechanics*, vol. 40, no. 8, pp. 1794–1805, 2007, ID: 271132. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021929006002946.

[6] R. Pascalau, V. A. Padurean, D. Bartos, A. Bartos, and B. A. Szabo, "The geometry of the circle of willis anatomical variants as a potential cerebrovascular risk factor," *Turk Neurosurg*, vol. 29, no. 2, pp. 151–158, 2019, pmid:29484629.

[7] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," *arXiv preprint arXiv:2104.13478*, 2021.

[8] H. Lei, N. Akhtar, M. Shah, and A. Mian, "Mesh convolution with continuous filters for 3-d surface parsing," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[9] A. G. Rahma and T. Abdelhamid, "Hemodynamic and fluid flow analysis of a cerebral aneurysm: A cfd simulation," *SN Applied Sciences*, vol. 5, no. 2, p. 62, 2023, ID: Rahma2023. [Online]. Available: https://doi.org/10.1007/s42452-023-05276-0.

[10] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "Nnu-net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature methods*, vol. 18, no. 2, pp. 203–211, 2021.

[11] F. Dumais *et al.*, "Eicab: A novel deep learning pipeline for circle of willis multiclass segmentation and analysis," *NeuroImage*, vol. 260, p. 119 425, 2022, ID: 272508. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811922005420.

[12] S. Avadiappan, S. Payabvash, M. A. Morrison, A. Jakary, C. P. Hess, and J. M. Lupo, "A fully automated method for segmenting arteries and quantifying vessel radii on magnetic resonance angiography images of varying projection thickness," *Frontiers in Neuroscience*, vol. 14, p. 537, 2020.

[13] R. Nader, R. Bourcier, and F. Autrusseau, "Using deep learning for an automatic detection and classification of the vascular bifurcations along the circle of willis," *Medical image analysis*, vol. 89, p. 102 919, 2023, ID: 272154. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841523001792.

[14] J. Kretschmer, C. Godenschwager, B. Preim, and M. Stamminger, "Interactive patient-specific vascular modeling with sweep surfaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2828–2837, 2013, ID: 1.

[15] H. Zakaria, A. M. Robertson, and C. W. Kerber, "A parametric model for studies of flow in arterial bifurcations," *Annals of Biomedical Engineering*, vol. 36, pp. 1515–1530, 2008.

[16] M. Decroocq, C. Frindel, P. Rougé, M. Ohta, and G. Lavoué, "Modeling and hexahedral meshing of cerebral arterial networks from centerlines," *Medical image analysis*, vol. 89, p. 102 912, 2023.

[17] U. Wickramasinghe, E. Remelli, G. Knott, and P. Fua, "Voxel2mesh: 3d mesh model generation from volumetric data," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23*, Springer, 2020, pp. 299–308.

[18] F. Kong, N. Wilson, and S. Shadden, "A deep-learning approach for direct whole-heart mesh reconstruction," *Medical image analysis*, vol. 74, p. 102 222, 2021, ID: 272154. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136184152100267X.

[19] F. Bongratz, A.-M. Rickmann, S. Pölsterl, and C. Wachinger, "Vox2cortex: Fast explicit reconstruction of cortical surfaces from 3d mri scans with geometric deep neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20 773–20 783.

[20] F. Bongratz, A.-M. Rickmann, and C. Wachinger, "Abdominal organ segmentation via deep diffeomorphic mesh deformations," *Scientific Reports*, vol. 13, no. 1, p. 18 270, 2023, ID: Bongratz2023. [Online]. Available: https://doi.org/10.1038/s41598-023-45435-2.

[21] *Topcow: Benchmarking topology-aware anatomical segmentation of the circle of willis (cow) for cta and mra*, 2023. arXiv: 2312.17670 [cs.CV].

[22] D. Selle, B. Preim, A. Schenk, and H.-O. Peitgen, "Analysis of vasculature for liver surgical planning," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1344–1357, 2002. DOI: 10.1109/TMI. 2002.801166.

[23] F. Ritter *et al.*, "Medical image analysis," *IEEE Pulse*, vol. 2, no. 6, pp. 60–70, 2011. DOI: 10.1109/MPUL.2011.942929.

[24] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, pp. 424–432.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, Springer, 2016, pp. 630–645.

[26] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*, Springer, 2018, pp. 287–297.

[27] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*, Ieee, 2016, pp. 565–571.

[28] L. Rundo *et al.*, "Use-net: Incorporating squeeze-and-excitation blocks into u-net for prostate zonal segmentation of multi-institutional mri datasets," *Neurocomputing*, vol. 365, pp. 31–43, 2019.

[29] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.

[30] P. Mouches and N. D. Forkert, "A statistical atlas of cerebral arteries generated using multi-center mra datasets from healthy subjects," *Scientific data*, vol. 6, no. 1, p. 29, 2019.

[31] Y. Sato *et al.*, "Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," *Medical image analysis*, vol. 2, no. 2, pp. 143–168, 1998, ID: 272154. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841598800091.

[32] J. Vollmer, R. Mencl, and H. Mueller, "Improved laplacian smoothing of noisy surface meshes," in *Computer graphics forum*, vol. 18, Wiley Online Library, 1999, ch. 3, pp. 131–138.

[33] Dawson-Haggerty et al., *Trimesh*, version 3.2.0, Dec. 8, 2019. [Online]. Available: https://trimesh.org/.

[34] P. Achlioptas, O. Diamanti, I. Mitliagkas, and L. Guibas, "Learning representations and generative models for 3d point clouds," in *International conference on machine learning*, PMLR, 2018, pp. 40–49.

[35] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr, "Discrete differential-geometry operators for triangulated 2-manifolds," in *Visualization and mathematics III*, Springer, 2003, pp. 35–57.

[36] A. Nealen, T. Igarashi, O. Sorkine, and M. Alexa, "Laplacian mesh optimization," in *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, 2006, pp. 381–389.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[38] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: Analysis, selection, and tool," *BMC medical imaging*, vol. 15, no. 1, pp. 1–28, 2015.

[39] S. S. SHAPIRO and M. B. WILK, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965. DOI: 10.1093/biomet/52.3-4.591. [Online]. Available: https://doi.org/10.1093/biomet/52.3-4.591.

[40] R. F. Woolson, "Wilcoxon signed-rank test," *Wiley encyclopedia of clinical trials*, pp. 1–3, 2007.

[41] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[42] N. Ravi *et al.*, "Accelerating 3d deep learning with pytorch3d," *arXiv:2007.08501*, 2020.

[43] L. Biewald, *Experiment tracking with weights and biases*, Software available from wandb.com, 2020. [Online]. Available: https://www.wandb.com/.

[44] A. Selver, A. E. Kavur, *et al.*, "Chaos-combined (ct-mr) healthy abdominal organ segmentation," in *The IEEE International Symposium on Biomedical Imaging (ISBI)*, 2019.

[45] F. Bongratz, A.-M. Rickmann, and C. Wachinger, "Neural deformation fields for template-based reconstruction of cortical surfaces from mri," *Medical Image Analysis*, vol. 93, p. 103 093, 2024.

[46] A. Reinke *et al.*, "Common limitations of image processing metrics: A picture story," *arXiv preprint arXiv:2104.05642*, 2021.

[47] D. Alblas, C. Brune, K. K. Yeung, and J. M. Wolterink, "Going off-grid: Continuous implicit neural representations for 3d vascular modeling," in *International Workshop on Statistical Atlases and Computational Models of the Heart*, Springer, 2022, pp. 79–90.

[48] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, and P.-A. Heng, "3d deeply supervised network for automatic liver segmentation from ct volumes," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, pp. 149–157.

[49] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[50] P. De Haan, M. Weiler, T. Cohen, and M. Welling, "Gauge equivariant mesh cnns: Anisotropic convolutions on geometric graphs," *arXiv preprint arXiv:2003.05425*, 2020.