

3D CORONARY VESSEL RECONSTRUCTION FROM BI-PLANE ANGIOGRAPHY USING GRAPH CONVOLUTIONAL NETWORKS

*Kit Mills Bransby*¹ *Vincenzo Tufaro*^{1,2} *Murat Çap*^{1,2}
*Greg Slabaugh*¹ *Christos Bourantas*^{1,2} *Qianni Zhang*¹

¹ Queen Mary University of London, United Kingdom

² Department of Cardiology, Barts Health NHS Trust, London, United Kingdom

ABSTRACT

X-ray coronary angiography (XCA) is used to assess coronary artery disease and provides valuable information on lesion morphology and severity. However, XCA images are 2D and therefore limit visualisation of the vessel. 3D reconstruction of coronary vessels is possible using multiple views, however lumen border detection in current software is performed manually resulting in limited reproducibility and slow processing time. In this study we propose 3DAngioNet, a novel deep learning (DL) system that enables rapid 3D vessel mesh reconstruction using 2D XCA images from two views. Our approach learns a coarse mesh template using an EfficientB3-UNet segmentation network and projection geometries, and deforms it using a graph convolutional network. 3DAngioNet outperforms similar automated reconstruction methods, offers improved efficiency, and enables modelling of bifurcated vessels. The approach was validated using state-of-the-art software verified by skilled cardiologists.

Index Terms— deep learning, 3d reconstruction, angiography

1. INTRODUCTION

XCA is a standard procedure in the assessment of coronary artery disease, where an injection of radiopaque contrast medium into vessels enables visualisation of lumen morphology [1]. Regions of narrowing (stenosis) caused by a build up of atherosclerotic plaque that restrict blood flow to the heart can be identified, aiding treatment such as stent placement. Quantitative coronary angiography (QCA) was introduced to provide precise quantification of plaque lesions and disease progression. However, QCA only uses a 2D representation of the lumen, and therefore limits visualisations in cases of vessel overlap and foreshortening where the full morphology of the vessel is either obscured or distorted [1]. In order to overcome these problems, 3D-QCA methods have been developed to reconstruct coronary artery segments in three dimensions using two or more angiographic views. 3D-QCA models are of high clinical relevance due to their application in computational fluid dynamics where the flow of blood through coronary vessels is simulated. These approaches allow for visualisation of vessel geometry and stenosis in 3D, however they are typically semi-autonomous, time-consuming, and require manual correction of vessel segmentation or multi-step input from clinicians. Bifurcation points where the vessel splits into multiple branches affect the flow and velocity of blood, however reconstructing main and side branches in a single 3D-QCA model has been a particular challenge in past investigations.

Since the inception of the ShapeNet dataset [2], mesh deformation DL algorithms such as 3DR2N2 [3], Pixel2Mesh [4], Pixel2Mesh++ [5] have demonstrated state-of-the-art results for 3D

reconstruction from 2D images. This is especially applicable in medical imaging where data is often limited to one or few views and has led to 3D reconstruction of the heart in HeartFFDNet [6], liver in Xray2Shape [7] and lungs in DeepOrganNet [8]. Such methods offer advantages over traditional approaches as they are able to reconcile surface morphology in information poor areas, have higher performance and are computationally efficient. Despite this, no attempts have been made to reconstruct coronary vessels using these advanced DL techniques.

Our Contribution. Based on the challenges discussed, we propose and validate a novel deep learning methodology for 3D reconstruction of coronary artery segments called 3DAngioNet. To the best of our knowledge this is the first deep learning paper to automate coronary segment reconstruction with bifurcation points and without the need for clinical correction.

2. METHOD

2.1. System Overview

3DAngioNet uses a coarse-to-fine approach based on three modules. Firstly, a Mesh Initialisation (MI) module creates a coarse mesh using an EfficientNetB3-UNet segmentation algorithm and back-projection from 2D to 3D using stereo geometry. Secondly, a Surface Refinement (SR) module adds fine-grained detail by deforming the initial mesh using image features sampled from an encoder using a graph convolutional network (GCN). Third, in cases of bifurcation an additional step is carried out where main and side branches are stitched together using a simple Boolean operation. The model takes a pair of angiographic images as input, with corresponding acquisition geometries and segment of interest (SOI) start and end points provided by an expert clinician, and outputs a 3D mesh surface. The full framework is presented in Fig. 1.

2.2. Mesh Initialisation Module

2D Vessel Segmentation. As segmentation of a specific SOI is required, there are a number of challenges to overcome when using deep learning segmentation algorithms. Firstly, as the SOI is similar to other vessels in shape, texture and pixel values, it is likely a DL segmentation algorithm will try to segment the full coronary tree resulting in a high number of false positives. This was confirmed experimentally and overcome by cropping the X-ray around the SOI. Despite this, overlapping vessels in the cropped patch may still restrict learning as the choice of vessel can be ambiguous. This was mitigated by rotating images so the start and end points of the SOI are fitted to the x-axis. This strategy ensures the algorithm learns to segment a vessel only if it is orientated in a horizontal position,

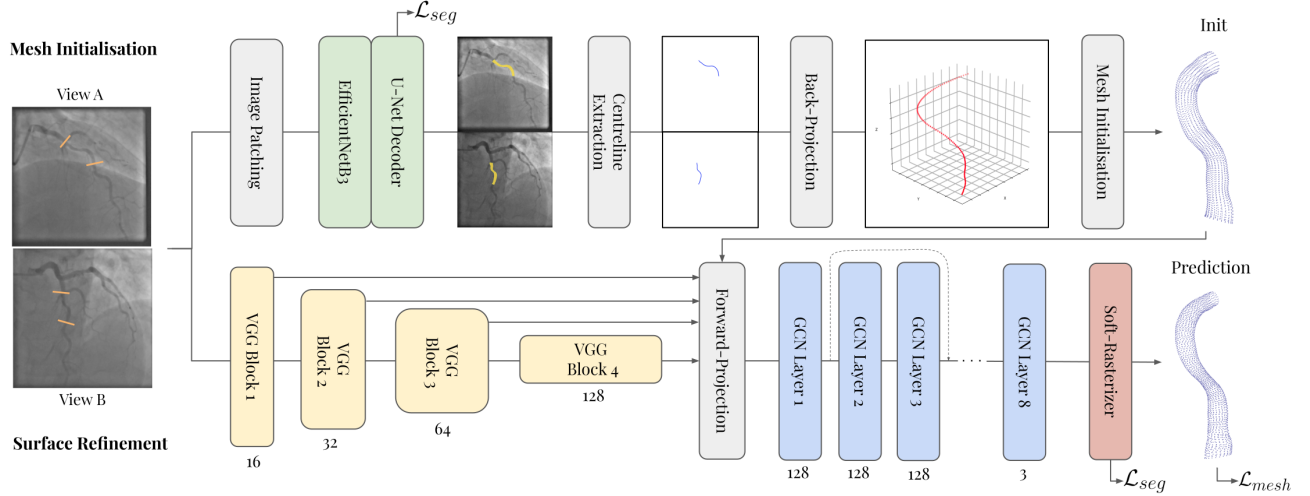


Fig. 1. 3DAngioNet framework.

vastly reducing false positives. A U-Net model [9] with an EfficientNetB3 [10] encoder backbone initialised with ImageNet weights is used as the segmentation architecture. The training of the network is supervised by a Binary Cross Entropy Dice loss (\mathcal{L}_{seg}) between the predicted probability map and a binary ground truth map.

3D Centreline Extraction. Contour coordinates from vessel segmentation predictions are extracted and the 2D centreline determined. As the geometries of multiple views are known, it is possible to back-project a 2D point (u,v) to the 3D plane (x,y,z) using triangulation [11].

Mesh Estimation. A mesh is constructed from the 3D centreline vector $P \in R^{100,3}$ and mean lumen radius vector $R \in R^{100,1}$ in Blender software. For each 3D centreline point P_i and lumen radius R_i , a ring of 60 vertices V_i is created where the distance from P_i to each new vertex $V_{i,j}$ is R_i . The ring is placed on the normal to the 3D centreline, where $P_i \vec{P}_{i+1} \perp P_i \vec{V}_{i,j}$. Each resulting coronary vessel mesh has 6,000 vertices connected by 11,940 edges to form 5,940 quad faces.

2.3. Surface Refinement

Image Features. A VGG16 [12] network encodes image features from each view, and returns the outputs from the first four convolutional blocks to the projection layer. This lightweight encoder is preferred over EfficientNet, as feature map resolution is maintained in the early encoding layers which helps to preserve the vessel of interest which only occupies a small area of the original image. High level feature maps from late in the network contain more coarse representations of vessel morphology such as curvature, thickness, shape, while low level features from early in the network contain finer details such as edges. The combination of high and low level features ensures both coarse and fine details are leveraged for improved deformation.

Forward-Projection. Image feature maps and the initialised mesh are passed to a forward-projection layer which maps relevant pixels to vertices on the mesh template. In forward-projection, a 3D point (x,y,z) on the initial mesh is projected to the 2D image feature plane (u,v) for each view. Four adjacent pixels are sampled using bi-linear interpolation for every 2D coordinate on each of the four feature maps, and the output vectors concatenated between views. The 3D

positional encoding values for the initial mesh are also concatenated with the output vector resulting in a feature vector of size 483 for each vertex.

Graph Convolutional Network. Our graph network receives graph features as input from the forward-projection layer where 3D shape information for each vertex of the initial mesh is encoded. Features are passed through a block of 8 GCN layers [13], where information is aggregated from neighbouring nodes, eventually regressing the 3D location for each vertex. Residual connections are added to minimise the over-smoothing effect common in graph networks. The graph convolutional operation is formalised by:

$$H^{(l+1)} = \sigma(\hat{D}^{-1/2} \hat{A} \hat{D}^{-1/2} H^{(l)} W^{(l)}) \quad (1)$$

Where \hat{A} is the adjacency matrix with self-loops of the graph G , $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$ is the diagonal degree matrix of G , $W^{(l)}$ are trainable weights for each layer and σ is a ReLU activation. $H^{(l)} \in R^{N \times D}$ is output for layer l where N is the number of nodes in the graph and D is the dimension of the hidden layer given to be 128 [13].

Differentiable Rendering. We implement a differentiable rendering pipeline based on Soft-Rasterizer [14] which provides an additional 2D supervisory signal during training. At each iteration our 3D mesh prediction is projected to the angiographic image space using the camera geometry of each of two views. Alignment between projection silhouette and vessel mask is calculated using the same \mathcal{L}_{seg} loss from the MI module.

Losses. The SR module is supervised by a point-wise mean squared error loss \mathcal{L}_{MSE} to minimise the distance between predicted \hat{v} and ground truth v mesh vertices.

$$\mathcal{L}_{MSE} = \frac{1}{|v|} \sum_{i=1}^{|v|} (v_i - \hat{v}_i)^2 \quad (2)$$

However, a MSE loss alone is not sufficient to produce realistic vessel shapes with smooth and even surfaces, hence a set of 3 regularization losses must be used to avoid undesirable morphological artefacts. A normal loss \mathcal{L}_{Norm} ensures faces are pointing in the correct direction by penalising large angles between the predicted

vertex normals $n_{\hat{v}_i}$ and ground truth vertex normals n_{v_i} .

$$\mathcal{L}_{Norm} = 1 - \sum_i \|n_{\hat{v}_i} \cdot n_{v_i}\| \quad (3)$$

$$n_{v_i} = \frac{1}{N(f)} \sum_{j=1}^{N(f)} f_{i,j} \quad (4)$$

Where n_{v_i} is the normal at a vertex v_i , $N(f)$ is the number of faces adjacent to the vertex v_i , $f_{i,j}$ refers to the normal of the j th face surrounding the i th vertex. The face normal is simply calculated as the cross product between any two vectors made from vertices that make up a face.

An edge loss \mathcal{L}_{Edge} is used to restrict the length of the edge from its initial state and leads to a visually appealing surface,

$$\mathcal{L}_{Edge} = \frac{1}{|E|} \sum_{v_i, v_j \in E} \|v_i - v_j\|_2^2 \quad (5)$$

where E is a set of graph edges.

A Laplacian loss \mathcal{L}_{Lap} is used to penalise overlapping vertices, ensuring a smooth mesh surface. The discrete Laplacian L at vertex v_i and the Laplacian loss \mathcal{L}_{Lap} are calculated as follows:

$$L(v_i) = \frac{1}{N(v_i)} \sum_{j \in N(v_i)} (v_i - v_j) \quad (6)$$

$$\mathcal{L}_{Lap} = \frac{1}{n} \sum_{i=1}^n \|L(v_i) - L(\hat{v}_i)\|_2^2 \quad (7)$$

Where $N(v_i)$ is the number of neighbouring vertices v_j to the vertex v_i that are connected by an edge, and $L(v_i)$ refers to the discrete Laplacian at a given vertex v_i . The losses are combined using Eq. 8, where weights are found through extensive experimentation.

$$\mathcal{L} = \mathcal{L}_{MSE} + 0.01\mathcal{L}_{Norm} + 2.5\mathcal{L}_{Edge} + 100\mathcal{L}_{Lap} + 0.0002\mathcal{L}_{Seg} \quad (8)$$

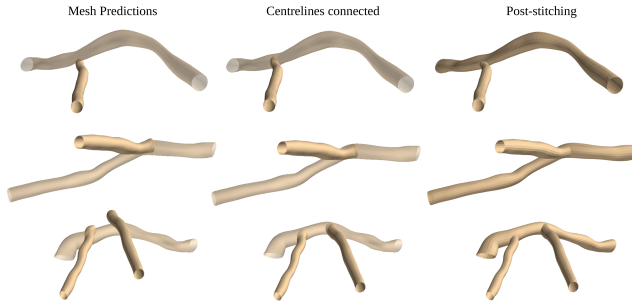


Fig. 2. Stitching Process.

2.4. Stitching

Main and side branches are reconstructed separately using the above method and then unnormalised to give their original positions in 3D space. It is assumed that the start point of side branch centreline should intersect somewhere along the main branch centreline. However, due to errors in acquisition or reconstruction, there is often a small distance between these. To fix this, side branch(es) are translated in 3D space so that their start point shares the same coordinates with the nearest main branch centreline point. The main and

side branch(es) are then combined using a Boolean Union operation which removes any area which is shared between the two mesh, and creates a new mesh from the unshared areas. This is formalised by $A \cup B := \{x \in R^3 \mid x \in A \text{ and } x \in B\}$ where A and B are the main branch and side branch meshes. The full stitching process is illustrated in Fig 2.

3. EXPERIMENTS AND RESULTS

3.1. Dataset

Overview A dataset of angiographic studies of 414 patients with atherosclerotic lesions was obtained using four different XCA machines. Coronary reconstruction was performed on a segment of intermediate significance detected on coronary angiography (90% > Diameter Stenosis >30%). In total 489 segments in these angiographic images fulfilled the inclusion criteria and were included in the final analysis. The dataset was split into 70% train, 15% val and 15% test examples, where each example has a pair of angiographic images and acquisition geometry information.

Ground Truth. 2D segmentation masks and 3D ground truth meshes were created from two precisely calibrated views using commercially validated QAngio XA 3D RE software (Medis Medical Imaging Systems, Leiden, The Netherlands), with manual correction performed by two expert cardiologists. For bifurcated vessels, only 2D ground truth masks are available. Each mesh is standardised to 6000 vertices, 5940 faces and 11940 edges.

Pre-Processing. Grayscale angiographic images (512 x 512) from each view were normalised to the pixel values [0,1]. 3D initialisation and ground truth mesh vertices were normalised by a translation to the origin, rotated in 3D to fit the z-axis and normalised to the interval [0,1]. Camera extrinsic and intrinsic parameters found in the acquisition DICOM file were used to roughly calibrate views. Calibration was then optimised to match the vessel start and end point correspondence given by clinicians using a Levenberg–Marquardt algorithm [15].

3.2. Training

The MI module was trained using an Adam optimizer with an initial learning rate of 0.001 with an epoch decay factor of 0.96. A batch size of 8 was used and the best validation loss was saved during training over 100 epochs. The SR module was trained using the same set up for 500 epochs, a decay factor of 0.99 and batch size of 1. 3DAngioNet is implemented with PyTorch using an NVIDIA 1080Ti GPU.

3.3. Comparison to Current Methods

We validate our framework by comparing test set predictions to the 3D ground truth using mean absolute error (MAE) and Hausdorff distance (HD). Precision and recall are also calculated by determining the percentage of points that can find a nearest neighbour within a threshold value τ [4]. F-score is the harmonic mean between precision and recall, and τ set at 0.0005. In addition, we project the 3D mesh predictions back to the 2D angiographic image space and calculate alignment with the 2D ground truth using Dice and Jaccard Index.

Results in Fig. 3, Fig. 4 and Table 1 demonstrate that our model is able to reconstruct vessels with minimal error and high resemblance to those created using state-of-the-art Medis software. By automating the vessel reconstruction step, our framework offers improvements in efficiency with reconstructions taking approximately

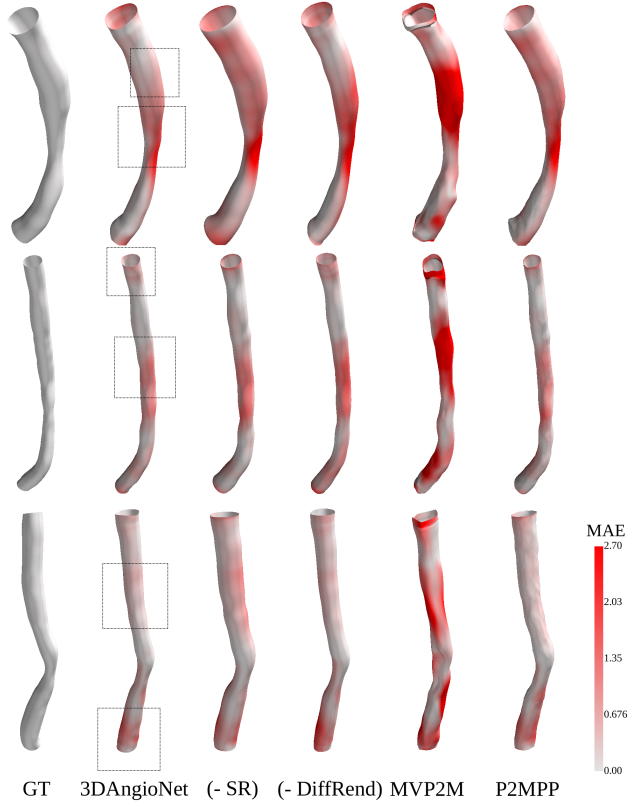


Fig. 3. 3D Qualitative Comparison. Bounding boxes indicate regions of best performance.

| | MAE↓ (mm) | HD↓ (mm) | F-score↑ | Dice↑ | Jaccard↑ |
|--------------|---------------|---------------|--------------|--------------|--------------|
| 3DAngioNet | 0.3459 | 1.1884 | 82.60 | 87.59 | 78.57 |
| (-) SR | 0.3658 | 1.8536 | 76.64 | 86.62 | 77.01 |
| (-) DiffRend | 0.3570 | 1.6283 | 81.63 | 86.57 | 76.95 |
| MVP2M | 0.4266 | 1.7612 | 81.54 | 86.15 | 76.25 |
| P2MPP | 0.3545 | 1.6520 | 81.90 | 87.40 | 78.18 |

Table 1. 3D and 2D Quantitative Comparison.

5s (1080Ti GPU, Batch size=1), compared to an estimated 4 minutes for Medis.

Further comparison between our system and other existing methodologies is difficult as our dataset is private, and no other public datasets or source code for similar vessel reconstruction methods is available. We compare to general purpose multi-view 3D reconstruction methods Multi-View-Pixel2Mesh (MVP2M) [5] and Pixel2Mesh++ (P2MPP) [5] networks. As these networks initialise a mesh from an ellipsoid they are incapable of creating hollow tubular structures, therefore we compare performance while using the mesh initialisation from 3DAngioNet. Table 1 and Fig. 3 demonstrates that our approach has more accurate and consistent reconstructions with fewer errors in lumen surface.

Table 1 presents an ablation study which compares performance using a number of different strategies: (1) Removing the SR module, and creating a mesh using the MI module only, and (2) Removing the differentiable rendering unit and additional 2D supervision. Models



Fig. 4. 2D Qualitative comparison between 2D projections from Fig 3 and the 2D ground truth (2-3x Zoom). Forth row compares a bifurcated vessel prediction after stitching. This comparison does not appear in Fig.3 as only 2D ground truth is available for bifurcated vessels.

trained using differentiable rendering and SR module outperform those without, justifying our system design choices.

4. CONCLUSION

In this work we propose an effective method for 3D coronary vessel reconstruction. We exploit segmentation and graph learning to automate this process, with minimal error. Experimental results demonstrate that our method is superior to general purpose 3D reconstruction algorithms, while producing mesh vessels with high resemblance to those created manually using state-of-the-art software. This work enables streamlining of time-consuming tasks while opening up future research avenues in atherosclerotic plaque progression. Our model is validated on data from four different XCA machines, demonstrating that our approach is generalizable to new data. In future work we aim to fully automate 3DAngioNet by learning the SOI region rather than using clinical guidance, and improve lumen resolution by co-registering QCA with intravascular imaging modalities.

5. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available by Barts Health NHS Trust. Ethical ap-

proval was not required as confirmed by the license attached with the data.

6. ACKNOWLEDGMENTS

This research is part of AI-based Cardiac Image Computing (AICIC) funded by the faculty of Science and Engineering at Queen Mary University of London. We would like to thank Medis Medical Imaging Systems for access to their QAngio XA 3D RE software which was used to create ground truth standards.

7. REFERENCES

- [1] Paolo Garrone, Giuseppe Biondi-Zoccai, Ilaria Salvetti, Noemi Sina, Imad Sheiban, Peter R Stella, and Pierfrancesco Agostoni, “Quantitative coronary angiography in the current era: principles and applications,” *Journal of interventional cardiology*, vol. 22, no. 6, pp. 527–536, 2009.
- [2] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu, “ShapeNet: An Information-Rich 3D Model Repository,” *Stanford University — Princeton University — Toyota Technological Institute at Chicago*, 2015.
- [3] Christopher B Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, and Silvio Savarese, “3d-r2n2: A unified approach for single and multi-view 3d object reconstruction,” in *European conference on computer vision*. Springer, 2016, pp. 628–644.
- [4] Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang, “Pixel2mesh: Generating 3d mesh models from single rgb images,” in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 52–67.
- [5] Chao Wen, Yinda Zhang, Zhuwen Li, and Yanwei Fu, “Pixel2mesh++: Multi-view 3d mesh generation via deformation,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1042–1051.
- [6] Fanwei Kong, Nathan Wilson, and Shawn Shadden, “A deep-learning approach for direct whole-heart mesh reconstruction,” *Medical Image Analysis*, vol. 74, pp. 102222, 2021.
- [7] Fei Tong, Megumi Nakao, Shuqiong Wu, Mitsuhiro Nakamura, and Tetsuya Matsuda, “X-ray2shape: reconstruction of 3d liver shape from a single 2d projection image,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 1608–1611.
- [8] Yifan Wang, Zichun Zhong, and Jing Hua, “Deeporgannet: on-the-fly reconstruction and visualization of 3d/4d lung models from single-view projections by deep deformation network,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 960–970, 2019.
- [9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [10] Mingxing Tan and Quoc Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.
- [11] Negar Chabi, Domenico Iuso, Oliver Beuing, Bernhard Preim, and Sylvia Saalfeld, “Self-calibration of c-arm imaging system using interventional instruments during an intracranial biplane angiography,” *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–12, 2022.
- [12] Karen Simonyan and Andrew Zisserman, “Very deep convolutional networks for large-scale image recognition,” *3rd International Conference on Learning Representations, ICLR 2015*, 2014.
- [13] Thomas N Kipf and Max Welling, “Semi-supervised classification with graph convolutional networks,” *International Conference on Learning Representations*, 2017.
- [14] Shichen Liu, Tianye Li, Weikai Chen, and Hao Li, “Soft rasterizer: A differentiable renderer for image-based 3d reasoning,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 7708–7717.
- [15] Jorge J Moré, “The levenberg-marquardt algorithm: implementation and theory,” in *Numerical analysis*. 1978, pp. 105–116, Springer.