GRAPH NEURAL NETWORK BASED FUTURE CLINICAL EVENTS PREDICTION FROM INVASIVE CORONARY ANGIOGRAPHY

Xiaowu Sun[1], *Theofilos Belmpas*[1], *Ortal Senouf*[1], *Emmanuel Abbé*[1], *Pascal Frossard*[1], *Bernard De Bruyne*[3] , *Olivier Muller*[2], *Stéphane Fournier*[2], *Thabo Mahendiran*[2], *Dorina Thanou*[1]

[1]EPFL, Switzerland [2]CHUV, Switzerland [3]OLV Hospital, Aalst, Belgium

ABSTRACT

Early prediction of future clinical events from invasive coronary angiography (ICA) remains a daily challenge in clinical routine practice. In this study, we hypothesize that stenosis's geometry information could benefit the prediction of future events from ICA. To address this, we propose a framework that employs graph neural networks (GNNs) to exploit geometry information from ICA and integrates it with clinical information to predict the occurrence of events at the stenosis level. The proposed model can be extended to predict events using two-view imaging data as well. The performance is compared to classical baseline models on a dataset comprising 1551 stenosis, out of which 414 exhibited an event in the following two years. The results illustrate that the proposed approach outperforms other models, with F1-scores of 0.57 and 0.59 for one-view and two-view data, respectively. To the best of our knowledge, this is the first work that investigates the importance of the geometry information for future events prediction in a learning context. The code is available at https://github.com/xsunn/eventsPre.

Index Terms— Graph neural network, future clinical events, invasive coronary angiography

1. INTRODUCTION

Cardiovascular disease (CVD) stands as the foremost cause of global mortality, accounting for 32% of all deaths [1]. Therefore, early prediction of future clinical events is of paramount importance. In clinical practice, diameter stenosis (DS) and fractional flow reserve (FFR) are used to estimate the severity of coronary artery stenosis and the need for intervention therapy to avoid CVD. DS is defined as the percentage that indicates how much the artery is narrowed by plaque or other obstructions [2]. FFR is the gold standard for evaluating whether a particular stenosis is responsible for a clinically significant reduction in blood flow within the coronary arteries [3]. However, as many as 8% of patients with mild to moderate stenosis as observed in the ICA images, which FFR deems as not severe, still present events within the following two years [4].

Predicting future clinical events through angiography alone is a complex task, as the prediction relies on numerous dynamic factors such as plaque composition, inflammation, and clot formation. These factors are not always fully captured by angiography, making accurate prediction of future events challenging. Several studies have been dedicated to integrating patient data, FFR and conventional machine learning (ML) methods as prognostic indicators for events prediction [5, 6]. However, ML models require manual feature extraction, and certain features cannot be acquired through the models' self-learning processes.

In recent years, deep learning-based methods have been applied and achieved immense success in angiography images analysis including vessel segmentation and stenosis detection [7, 8]. However, the potential capability of deep learning methods for future events prediction has not been fully investigated yet. The first related work used a convolutional neural network (CNN), ResNet [9] as the backbone, and extracted patches centered on artery stenosis to predict myocardial infarction (MI) at a lesion level within 5 years [10]. An anatomy-informed multi-task deep learning framework was proposed to predict future MI events in both patients and lesion level, combining different views [11]. However, none of these existing approaches considered explicitly the geometric information, which offers a holistic perspective on coronary artery structure and the effects of stenosis on blood flow within angiography images.

In this work, we aim to extract and use the stenosis related geometric features, such as stenosis severity and overall anatomy of arteries, for the occurrence of future events prediction at a stenosis level. Graph neural networks (GNNs), a recent advancement in deep learning techniques [12], are capable of capturing geometric information from complex data. GNN-based models have demonstrated promising applications in the analysis of ICA images, such as coronary artery segmentation and reconstruction [13, 14]. In this study, we focus on a more challenging future prediction task, and propose a GNN-based model to explore the lesion geometric information to predict future events within 2 years at a stenosis level. The contribution of this work is three-fold: (1) We present, to the best of our knowledge, the first study to introduce GNNs for incorporating the lesion geometry in predicting future events from ICA. (2) We propose a GNN-based framework that is highly extensible, capable of integrating both geometry information from multiple views and addi-

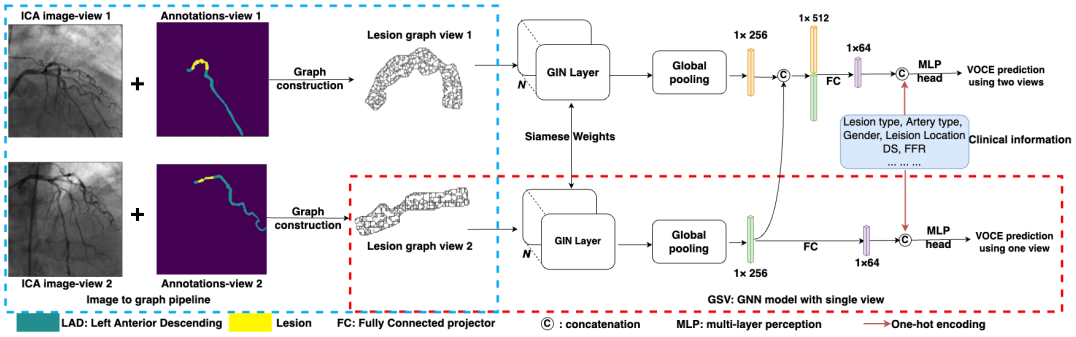

Fig. 1. Proposed future events prediction framework which contains a graph construction pipeline and a GNN based prediction network which can be extended for two views. ***GSV***: GNN network using single view as input.

tional clinical measurements, such as FFR and DS, for events prediction. (3) We conduct a comparative analysis between our proposed approach and clinical methods, as well as purely image-based methods, demonstrating the importance of geometry information in events prediction.

2. METHODOLOGY

In this section, we first describe how to model the geometry of the stenosis by extracting a graph from an ICA image. We then propose a novel learning framework that exploits the geometry together with clinical information as input of a GNN model to predict future events.

2.1. Image to graph pipeline

The left module in Fig. 1 describes the pipeline of lesion graph construction from an ICA image. Given that the lesion is a common contributor to future events, this study uses lesion annotations to exclude the region out of the lesion. Each lesion pixel is mapped to a node of the graph. Two approaches, namely K Nearest Neighbor (KNN) and Delaunay triangulation (DT), are introduced to construct the edges between neighboring nodes. KNN connects the K closest nodes in terms of proximity, while DT constructs a triangulation of the nodes such that no node is inside the circumcircle of any triangle. Fig. 2 illustrates three graph examples where the four columns represent a raw ICA image, the corresponding annotations, point cloud of node coordinates and a graph representation of these points, respectively. Compared to the raw image, the graph representation of the artery clearly preserves the geometry, while filtering out the pixel noise from the background.

Given a graph ***G=(V, E)*** constructed from an ICA image, where ***V*** is the set of nodes and ***E*** denotes the edges between them, each node features are represented by both the artery's geometry and its texture. Texture features are extracted from the gray-level intensity in the original raw image, while geometric features encompass the pixel positions in the Euclidean space of the raw image and their respective distances from the artery's center-line ***L*** which can be extracted using the lesion annotation. The corresponding coordinate in the raw image for the ith node in graph ***G*** is defined as $PN_i = (x_i, y_i)$, similarly, the corresponding coordinate in the raw image for the jth pixel on the center-line ***L*** is denoted as $PL_j = (X_j, Y_j)$. The distance between the ith node PN_i in graph ***G*** and the jth pixel PL_j on ***L*** can be derived using the following equation:

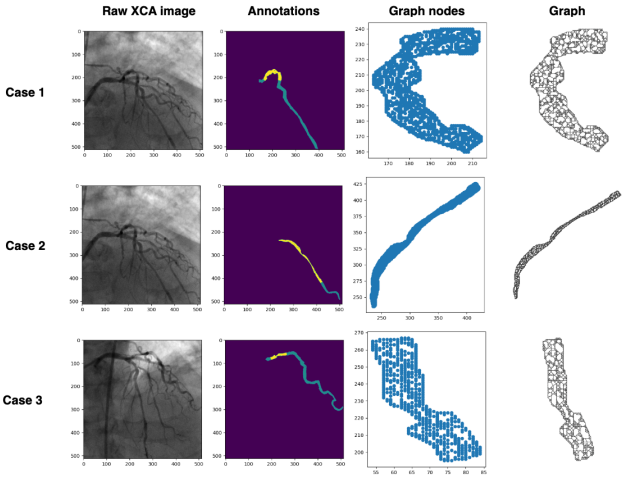

Fig. 2. Graph constructed from examples of ICA images. Those four columns are raw image, anatomical annotations, graph node coordination and graph edge connection.

$$d(PN_i, PL_j) = \sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2}. \quad (1)$$

The distance between the ith graph node PN_i and the center-line is defined as:

$$D(PN_i, L) = min(\{d(PN_i, PL_j)\}_{j=1,2,...,M}), \quad (2)$$

where ***M*** is the pixel number on ***L***, ***min*** is the minimum value.

Given a set of images represented as a graph ***G*** and the clinical information ***C***, the goal is to design a model ***F(G,C)*** that learns to predict the occurrence of future events, i.e., 1 for events and 0 for non-events.

2.2. GNN based Network Architecture

***GSV*: *G*NN network for a *S*ingle-*V*iew ICA image**. Given a graph representation of an ICA image as shown in Fig.1, the GSV model first employs *N* Graph Isomorphism Network (GIN) layers [15] as the backbone to learn an embedding for each node within the graph. Then, a global pooling layer is

used to aggregate the node embeddings into a comprehensive global representation of dimension 256, followed by a fully connected (FC) layer of dimension 64. Additional clinical information derived from patient and lesion level data are concatented to the graph features, forming a joint representation. The classification head is implemented by a Multi-layer Perceptron (MLP) comprising four hidden layers that continuously reduce the dimension of the final embedding to one. Additionally, the input to each layer is normalized by Batch normalization (BN) and dropout is applied for regularization. The non-linearity function at the ouput of each layer is set to be the rectified linear unit (ReLU).

***GTV*: siamese *G*NN for *T*wo-*V*iews ICA image.** In clinical practice, cardiologists often need to integrate information from multiple views of arteries to reach the final diagnosis. Our framework can be easily extended to incorporate complementary information from two views by using two shared GSV as backbone. As illustrated in Fig.1, comparing with GSV, GTV adds another branch to extract geometry from the second view while keeping the other structures the same as the GSV. Those two branches also employ a siamese *N* GIN layers as backbone, the same as that in GSV. The global representations from two views are concatenated and projected to a lower dimension of 64 by a FC layer. Following that, the clinical information is fused using another FC layer. Given that considering more than two views would reduce the dataset, in this study, we initially predict events using two views.

3. EXPERIMENTS AND RESULTS

In this section, we quantify the performance of our method. We first introduce the dataset and discuss baseline models including clinical parameter-based approaches, classical ML models and ResNet-based models. Second, we investigate the effect of the graph connectivity on the predictive performance. The performance of our proposed models is then compared to that of those baseline models. Finally, an ablation study on the clinical data is also performed. The model's performance is evaluated using the classification metrics including Accuracy (ACC), Recall (Rec), Precision (Prec), F1-score (F1) and Specificity (Spec).

Five-fold cross-validation is applied to evaluate the performance of different models. In each run, one fold is held out as the testing fold, while the remaining four folds are utilized for training and validation. The optimal hyper-parameters of GIN layers, learning rate and weight decay are determined using optuna [16] with 20 trials based on the best F1 score on training and validation set. They are set to 6, 0.00115 and 0.08, respectively. With these hyper-parameters, we train a model on the training and validation data using cross-entropy as the loss function. The testing data is only used for the final evaluation of a model. The averaged values and standard deviation of testing fold across five folds are reported as the final results. All the experiments are excuted on a machine equipped with an NVIDIA A100 GPU with 80 GB internal memory and implemented in Pytorch with the following parameters: Adam optimizer and batch size as 50.

3.1. Dataset

FAME2 dataset [17] includes 563 patients with stable coronary artery diseases. Each patient has at least one lesion in their main arteries. Both the arteries and lesions have been annotated by the cardiology team from the Lausanne University Hospital (CHUV), Switzerland. Two rotation projections with various angles are employed to multiple 2D images to image the complete coronary lesions. The primary end point of events in this datast is a composite of cardiac death, MI, urgent and non-urgent revascularization. The patients underwent re-evaluation after a follow-up period of two years, and their lesions are categorized into events or not. In this dataset of 1748 2D images, 1551 images have been detected with lesions. Furthermore, 414 images are labeled as events. Only the images with lesion are used to train and test the models. FFR measurement in the area of the lesion, diameter of the stenosis (DS), the minimum lumen diameter (MLD), the length of the lesion, age and gender are also collected as the clinical information. The data is divided into five folds with 113, 113, 113, 112, 112 patients (306, 310, 314, 305, 316 2D images). The folds are stratified, i.e, the same ratio of events/no-events lesions is maintained across folds.

3.2. Baseline models

Clinical baselines: In clinic routine practice, the cutoff thresholds of $\geq$0.5 and $\leq$0.75 for DS and FFR [3] respectively, are used as guidelines to categorize a lesion as potentially contributing to future events.

Traditional machine learning baselines: XGBoost, SVM, Logistic Regression, and Decision Tree are employed as machine learning baseline models. 46 radiomics features extracted from the entire artery and lesion using PyRadiomics [18] are combined with the clinical information as the inputs. The type of radiomics features are related to the shapes of the lesion, such as the 2D mesh surface and Sphericity, and the texture including the Grey Level Co-occurrence Matrix (GLCM).

ResNet based baseline: Fig. 3 shows the architecture of ResNet [9] based model which takes 200x200 image patch containing the whole lesion and clinical information as input. The cropped patch and corresponding lesion annotation are concatenated along the channel dimension as the input. The image features extracted using the ResNet are projected to a lower dimension via a FC layer. Then those features are concatenated with the clinical information, followed by another FC layer.

3.3. Choice of graph construction strategy

Table 1 reports the prediction results on the FAME2 dataset derived from GSV model, using the graph generated by KNN and DT. For KNN, we set the number of neighbors to 5 and 10, respectively. For DT, the values of 5 and 10 represent the

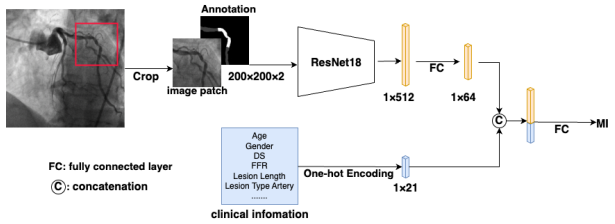

Fig. 3. ResNet based baseline model. The cropped image patch, lesion annotation and the clinical information are combined as the input.

Table 1. Prediction performance on FAME2 derived from GSV using different graph generation approaches for single view. The best results are highlighted in bold.

Input	Accuracy	Recall	Precision	F1	Spec
KNN-5	0.71±0.04	***0.70±0.06***	0.47±0.03	0.56±0.04	0.72±0.06
KNN-10	0.69±0.06	0.66±0.06	0.47±0.07	0.55±0.03	0.69±0.10
DT-5	0.68±0.02	0.69±0.05	***0.48±0.05***	***0.57±0.04***	0.67±0.02
DT-10	***0.73±0.05***	0.68±0.04	0.48±0.07	0.56±0.03	***0.74±0.08***

maximum allowed Euclidean distance for a connection to be preserved. The results reveal that a DT construction outperforms KNN overall, yielding better results in term of ACC, Pre, F1 and Spec. Although DT-10 achieved higher ACC and Spec, DT-5 show the best performance with Prec of 0.48 and F1 of 0.57, respectively. For the rest of the experiments, we construct the graph using DT-5.

3.4. Prediction results

Table 2 summarizes the averaged prediction results across all testing folds in the five-fold cross-validation derived from different models using data with either single view or exclusively with two views. The results demonstrate that when considering each view as a single input, the proposed GSV achieves superior performance, particularly in terms of Prec, F1-score and Spec with 0.48, 0.57 and 0.67, respectively. Although machine learning classifiers may achieve a better Spec, the Recall is significantly lower, leading to a subpar F1 score. ResNet shows reasonable values across all metrics, but still lower than GSV. In the case of two-view data, the GSV model treats each view as two separate cases, whereas the GTV model treats both views as a single subject. When evaluating the performance obtained from GTV and GSV using data with two views, GTV exhibits a slight improvement with a Recall of 0.69, Precision of 0.51 and an F1-score of 0.59. This demonstrates that integrating two views can lead to an enhanced prediction. Additionally, only the GNN based models outperform the FFR method in terms of F1-score, even though both CNN and GNN integrate lesion features.

3.5. Ablation study on clinically-relevant features

Table 3 presents the results from an ablation study conducted on clinically-relevant information and its impact on the events prediction performance. When comparing Table 2 to Table 3, we observe that the F1-score obtained from GSV using single-view input decreases from 0.57 to 0.49, while the F1-score of GTV decreases from 0.59 to 0.51. As expected, when

Table 2. Prediction performance derived from different models. DS: diameter stenosis. FFR: fractional flow reserve. LR: Logistic Regression. DTr: Decision Tree. GSV: GNN model using a single view. GTV: GNN model using two views.

Data	Model	Accuracy	Recall	Precision	F1	Spec
All images (Single view)	DS	0.56	0.69	0.34	0.45	0.51
	FFR	0.59	***0.85***	0.38	0.52	0.49
	XGBoost	0.69±0.02	0.28±0.04	0.32±0.04	0.30±0.04	0.81±0.02
	SVM	0.75±0.02	0.09±0.05	0.37±0.10	0.14±0.06	***0.95±0.03***
	LR	***0.76±0.01***	0.05±0.04	0.29±0.25	0.28±0.40	0.92±0.11
	DTr	0.65±0.01	0.39±0.05	0.30±0.04	0.34±0.04	0.72±0.02
	ResNet	0.61±0.05	0.59±0.09	0.38±0.02	0.46±0.03	0.62±0.09
	GSV	0.68±0.02	***0.69±0.05***	***0.48±0.05***	***0.57±0.04***	0.67±0.02
Images with two views	GTV	0.72±0.04	***0.69±0.07***	***0.51±0.05***	***0.59±0.04***	0.73±0.05
	GSV	***0.72±0.03***	0.68±0.08	0.50±0.04	0.58±0.05	***0.74±0.01***

Table 3. Prediction performance without clinically-relevant features on FAME2 dataset. LR: Logistic Regression. DTr: Decision Tree. GSV: GNN model using a single view. GTV: GNN model using two views.

Data	Model	Accuracy	Recall	Precision	F1	Spec
All images (Single view)	XGBoost	0.67±0.01	0.19±0.05	0.24±0.05	0.21±0.04	0.82±0.02
	SVM	***0.76±0.02***	0.01±0.01	0.11±0.15	0.02±0.03	***0.99±0.01***
	LR	***0.76±0.02***	0.06±0.04	0.38±0.23	0.11±0.07	0.97±0.004
	DTr	0.61±0.04	0.26±0.06	0.22±0.05	0.24±0.05	0.72±0.04
	ResNet	0.32±0.06	***0.88±0.12***	0.28±0.01	0.42±0.01	0.10±0.11
	GSV	0.53±0.02	0.73±0.06	***0.37±0.02***	***0.49±0.02***	0.44±0.06
Images with two views	GTV	0.64±0.06	***0.65±0.08***	***0.43±0.08***	***0.51±0.04***	***0.64±0.12***
	GSV	***0.64±0.05***	0.64±0.07	0.41±0.04	0.50±0.04	0.63±0.09

clinically relevant information is absent, not only GNN-based models underperform, but all baseline models perform worse as well, demonstrating the critical importance of incorporating clinical data for accurate events prediction. Moreover, it is noteworthy that even without the utilization of clinical information, the GNN-based models still consistently outperform the other models.

4. DISCUSSION AND CONCLUSION

In this work, we developed and evaluated a novel GNN based framework which takes into account the geometry of the stenosis combined with clinical information for the automated prediction of future events from coronary angiography images. The main findings of our study are the following: (1) Geometry feature is a key factor for events prediction, which integrated with clinical relevant features such as FFR and DS, can boost the prediction performance. (2) The proposed GNN-based approach exhibits a promising performance in predicting future events, outperforming the clinical measurements of FFR, DS and the other baseline methods. (3) Although the improvement from using two views may not be substantial, view fusion can contribute to improved predictions.

Future events prediction is a highly complex task, as angiography alone may not encompass all the necessary information. Nevertheless, our study demonstrates that through proper exploitation of these data, a substantial performance improvement can be achieved, surpassing the metrics commonly employed in clinical practice. In this study, we consider only two views from the same lesion for view fusion. However, an efficient fusion module for more than two views should be investigated in future research.

5. COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available on reasonable request. This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the institution's human research committee.

Conflicts of Interest. The authors have no relevant financial or non-financial interests to disclose.

6. REFERENCES

[1] M. AB Khan, M. J. Hashim, H. Mustafa, M. Y. Baniyas, Shaikha K. Al S., and et al., "Global epidemiology of ischemic heart disease: results from the global burden of disease study," *Cureus*, vol. 12, no. 7, 2020.

[2] H. Samady, W. Lepper, E.R. Powers, K. Wei, M. Ragosta, GG. Bishop, IJ. Sarembock, and et al., "Fractional flow reserve of infarct-related arteries identifies reversible defects on noninvasive myocardial perfusion imaging early after myocardial infarction," *Journal of the American College of Cardiology*, vol. 47, no. 11, pp. 2187–2193, 2006.

[3] B. De Bruyne, J. Pijls, N.H. andBartunek, K. Kulecki, J.W. Bech, H. De Winter, and et al., "Fractional flow reserve in patients with prior myocardial infarction," *Circulation*, vol. 104, no. 2, pp. 157–162, 2001.

[4] G. Ciccarelli, E. Barbato, G.G. Toth, B. Gahl, P. Xaplanteris, S. Fournier, A. Milkas, and et al., "Angiography versus hemodynamics to predict the natural history of coronary stenoses: fractional flow reserve versus angiography in multivessel evaluation 2 substudy," *Circulation*, vol. 137, no. 14, pp. 1475–1485, 2018.

[5] S. Gupta, D.T. Ko, P. Azizi, M.R. Bouadjenek, M. Koh, and et al., "Evaluation of machine learning algorithms for predicting readmission after acute myocardial infarction using routinely collected clinical data," *Canadian Journal of Cardiology*, vol. 36, no. 6, pp. 878–885, 2020.

[6] D. Mandair, P. Tiwari, S. Simon, K.L. Colborn, and M.A. Rosenberg, "Prediction of incident myocardial infarction using machine learning applied to harmonized electronic health record data," *BMC medical informatics and decision making*, vol. 20, no. 1, pp. 1–10, 2020.

[7] Y. Li, Y. Wu, J. He, W. Jiang, J. Wang, Y. Peng, Y. Jia, and et al., "Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography," *European Radiology*, vol. 32, no. 9, pp. 6037–6045, 2022.

[8] R. Avram, J.E. Olgin, Z. Ahmed, L. Verreault-Julien, A. Wan, J. Barrios, S. Abreau, D. Wan, J.E. Gonzalez, J.C. Tardif, and D.Y. So, "Cathai: fully automated coronary angiography interpretation and stenosis estimation," *npj Digital Medicine*, vol. 6, no. 1, pp. 142, 2023.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[10] T. Mahendiran, D. Thanou, O. Senouf, D. Meier, N. Dayer, F. Aminfar, and et al., "Deep learning-based prediction of future myocardial infarction using invasive coronary angiography: a feasibility study," *Open Heart*, vol. 10, no. 1, pp. e002237, 2023.

[11] I.D. Sievering, O. Senouf, T. Mahendiran, D. Nanchen, S. Fournier, O. Muller, and et al., "Anatomy-informed multimodal learning for myocardial infarction prediction," *medRxiv*, pp. 2023–07, 2023.

[12] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," *AI open*, vol. 1, pp. 57–81, 2020.

[13] J.M. Wolterink, T. Leiner, and I. Išgum, "Graph convolutional networks for coronary artery segmentation in cardiac ct angiography," in *Graph Learning in Medical Imaging: First International Workshop, GLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 1*. Springer, 2019, pp. 62–69.

[14] K.M. Bransby, V. Tufaro, M. Cap, G. Slabaugh, C. Bourantas, and Q. Zhang, "3d coronary vessel reconstruction from bi-plane angiography using graph convolutional networks," *IEEE International Symposium on Biomedical Imaging*, 2023.

[15] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?," *International Conference on Learning Representations*, 2019.

[16] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, pp. 2623–2631.

[17] B. De Bruyne, N.H. Pijls, B. Kalesan, E. Barbato, P.A. Tonino, and et al., "Fractional flow reserve–guided pci versus medical therapy in stable coronary disease," *New England Journal of Medicine*, vol. 367, no. 11, pp. 991–1001, 2012.

[18] J.J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, and et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer research*, vol. 77, no. 21, pp. e104–e107, 2017.