**Response letter**

Dear Editor and Reviewers,

We are grateful for the comments and suggestions from the editors and the reviewers, which are crucial for improving our work. We have revised the manuscript to address the reviewers’ comments fully. Our point-by-point reply to the review comments is summarized below. In this document, the original reviewers’ comments are in **black**; our responses are in **blue**; the quotations in the revised manuscript are in **red**.

**Reviewers’ comments:**

**Reviewer #1:**

1. There are several small errors in the paper. For example, in the "1) Hyper association graph construction" part, the definition of "the set of hyperedges", s.t. i, j belong to {1,...,n₁} should be i, j, k belong to {1,...,n₁}. The word "regulizer" in 2) ℓ1-Norm constraint regularization should be "regularizer".

A: Thanks for pointing out the issues. We have gone through the manuscript carefully and made all necessary corrections.

1) *Hyper association graph construction*. Formally, the hyper association graph is denoted as , where

* represents the vertex set, and .
* and , indicating the set of hyperedges and .
* indicates the attribute vectors associated with each vertex.
* indicates the attribute vectors associated with each hyperedge.

In addition, we changed the term ‘regularization’ into ‘regularizer’, including the following three statements in section 3.2:

* *2) -Norm constraint regularizer*.
* where , , and are MLPs for vertex, hyperedges, row and column regularizer embeddings.
* For the row and column regularizer terms, the plain GCN is employed to update the features, as denoted in Eqs. 16 and 17.

Furthermore, we corrected multiple errors in the manuscript.

1) Original text:

*Hyper association graph construction*.

* and , indicating the set of hyperedges and .

Revised text:

* and , indicating the set of hyperedges and .

2) Original text:

If the nodes , and from are connected, and the nodes , and from are connected, then a hyperedge is constructed in , where

Revised text:

If the nodes , and from are connected, and the nodes , and from are connected, then a hyperedge is constructed in , where .

3) Original text:

Formally, the hyper association graph used in this study is converted into , where and indicate the row and column hyperedge; and represent features of row and column hyperedges; and they are initialized as zeros, respectively.

Revised text:

Formally, the hyper association graph used in this study is converted into , where and indicate the row and column hyperedge; and represent features of row and column hyperedges; and they are initialized as zeros, respectively.

4) Original text:

We employ the *edge graph transformer* (EGT) to aggregate features, which is denoted as for hyperedge .

Revised text:

We employ the *edge graph transformer* (EGT) to aggregate features, which is denoted as for hyperedge .

5) Original Eqs. 10 and 11

|  |  |
| --- | --- |
|  | (10) |
|  | (11) |

Revised Eqs. 10 and 11

|  |  |
| --- | --- |
|  | (10) |
|  | (11) |

6) Original text:

represents the set of the connected vertices for hyperedge in and is the number of the included vertices in the set.

Revised text:

represents the set of the connected vertices for hyperedge in and is the number of the included vertices in the set.

7) Original Eqs. 13 and 14:

|  |  |
| --- | --- |
|  | (13) |
|  | (14) |

where , and are three MLPs to perform feature embeddings for multi-head attention; represents the set of connected hyperedges for vertex in and is the number of hyperedges in the set. Then, we update the features for vertices, as shown in Eq. 15.

Revised text:

|  |  |
| --- | --- |
|  | (13) |
|  | (14) |

where , and are three MLPs to perform feature embeddings for multi-head attention; represents the set of connected hyperedges for vertex in and is the number of hyperedges in the set. Then, we update the features for vertices, as shown in Eq. 15.

2. In 2) l1-norm constraint regularization, the corresponding equations of the constraint regularization should be provided.

A: Thanks for pointing out the issue that we should provide -norm constraint in the loss function. In the revised manuscript, we added the corresponding equations and updated the loss function.

In detail, row and column hyperedges were added to enforce the​ *-Norm* constraint by normalizing the rows and columns of the assignment matrix. This ensures that the assignment probabilities sum to 1 across rows or columns. As mentioned in Section 3.2, the updated features of row/column hyperedges are denoted as and . Instead of adding the *-Norm* constraint on the predicted assignment matrix, i.e.  , we applied the *-Norm* constraint to the row and column hyperedges. The loss function for this part is denoted as

And the overall loss function is corrected as:

Our model adds row and column hyperedges implicitly by using normalization functions that enforce norm constraints on the assignment matrix. In the revised manuscript, we add the related information into section 3.2.

(Section 3.2 Hyper graph matching for coronary artery semantic labeling)

The confidence loss is defined as the mean squared error between the TCP and the approximated TCP in Eq. 22 and the overall loss function is defined in Eq. 23.

|  |  |
| --- | --- |
|  | (23) |

where represents the -norm constraint.

3. The language needs to be further polished.

A: Thanks for your suggestion. We have double checked the grammar issue and improved the English by a native speaker in the revised version. Here are the detailed revisions.

1) Original text: ICAs are instrumental to aid cardiologists in identifying blockages within the coronary arteries.

Revised text: ICAs are instrumental in aiding cardiologists in identifying blockages within the coronary arteries.

2) Original text: Automatically identifying the correct anatomical branches offers valuable insights for automatic generation of diagnostic reports and quantification of region of interests.

Revised text: Automatically identifying the correct anatomical branches offers valuable insights for the automatic generation of diagnostic reports and quantification of regions of interest.

3) Original text: In detail, our proposed approach includes a Hyper Association graph-based Graph-Matching Network with Uncertainty Quantification (HAGMN-UQ) to establish semantic correspondences between coronary arterial segments from ICAs.

Revised text: Specifically, our proposed approach includes a Hyper Association graph-based Graph-Matching Network with Uncertainty Quantification (HAGMN-UQ) to establish semantic correspondences between coronary arterial segments from ICAs.

4) Original text: However, both the slow inference speed and limited accuracy render it unsuitable for real-time scenarios.

Revised text: However, both the slow inference speed and limited accuracy render them unsuitable for real-time scenarios.

5) Original text: The objective of graph matching is to establish a meaningful correspondence between nodes and edges across different graphs.

Revised text: The objective of graph matching is to establish meaningful correspondences between nodes and edges across different graphs.

6) Original text: Graph neural network (GNN) has further enriched learning-based graph matching.

Revised text: Graph neural networks (GNNs) have further enriched learning-based graph matching.

7) Original text: Uncertainty quantification (UQ) is a crucial aspect in the optimization of the decision-making process, as it helps mitigate the effects of uncertainties and enhances the reliability of predictions.

Revised text: Uncertainty quantification (UQ) is a crucial aspect of optimizing of the decision-making process, as it helps mitigate the effects of uncertainties and enhances the reliability of predictions.

8) Original text: By leveraging the labeled data, the node correspondences between arterial segments are automatically identified, and a ground truth permutation matrix is generated.

Revised text: Leveraging the labeled data, node correspondences between arterial segments are automatically identified, and a ground truth permutation matrix is generated.

9) Original text: The designed algorithm for testing using our HAGMN-UQ with confidence loss is shown in Algorithm 1.

Revised text: The algorithm for testing using our HAGMN-UQ with confidence loss is presented in Algorithm 1.

10) Original text: During the testing, if the node has no structural loss, then we accept the prediction, as shown in Lines 7.

Revised text: During the testing, if the node has no structural loss, then the prediction is accepted, as shown in Lines 7.

11) Original text: In this study, 263 and 455 ICAs were acquired from Site 1 at The First Affiliated Hospital of Nanjing Medical University and Site 2 at the Medical University of South Carolina, respectively. The image size of ICA videos ranged from 512 × 512 to 864 × 864, and the pixel spacing ranged from 0.2 mm to 0.39 mm. The number of arterial segments for each view angle and each site is shown in Table S1 in the supplementary material.

Revised text: In this study, 263 ICAs were acquired from Site 1 at The First Affiliated Hospital of Nanjing Medical University, and 455 ICAs were acquired from Site 2 at the Medical University of South Carolina. The image sizes of ICA videos ranged from 512 × 512 to 864 × 864, and the pixel spacing ranged from 0.2 mm to 0.39 mm. Table 1 presents the number of images in each view angle, while Table S1 in the supplementary material displays the number of arterial segments for each view angle and each site.

12) Original text: For each baseline, we performed the 5-fold cross validation using the enrolled 718 ICAs. For the graph matching-based networks, we also split the enrolled subjects into a , and . However, during the testing, each tested ICA was compared with every ICA in . The performance comparisons are shown in Table 4.

Revised text: For each baseline, we conducted the 5-fold cross validation using the enrolled 718 ICAs. For the graph matching-based networks, subjects were divided into , and . During testing, each tested ICA was compared with every ICA in . The performance comparisons are shown in Table 4.

13) Original text: Additionally, to test the domain difference, we conducted experiments using images from one hospital and employed templates from another hospital during testing to evaluate the generalizability of the designed algorithms.

Revised text: To further test domain differences, experiments were conducted using images from one hospital and templates from another hospital during testing, evaluating the generalizability of the designed algorithms.

14) Original text: Further insight from Table 2 indicates that the proposed HAGMN-UQ achieved an accuracy of 0.9211 using ICAs from these two centers simultaneously.

Revised text: Further insights from Table 2 indicate that HAGMN-UQ achieved an accuracy of 0.9211 using ICAs from both centers simultaneously.

15) Original text: The difference may be caused by exposure time, dose of contrast agent (including operator experience and injection strength), and other potential sources of variability inherent to multi-center datasets. Examples of representative ICAs from Sites 1 and 2 are shown in Figure S6 in the supplementary materials.

Revised text: The difference may be attributed to factors such as exposure time, contrast agent dosage (including operator experience and injection strength), and other inherent sources of variability in multi-center datasets. Representative ICAs from Sites 1 and 2 are illustrated in Figure S6 in the supplementary materials.

16) Original text: However, we cannot guarantee that the binary segmentation model would generate satisfactory arterial contours for all ICAs due to variations in contrast dye degradation. To test the robustness of the designed model, we conducted three experiments to illustrate its robustness.

Revised text: However, we cannot guarantee that the binary segmentation model would generate satisfactory arterial contours for all ICAs due to the degradation of contrast dye. To test the robustness of the designed model, we conducted 3 experiments to illustrate the robustness of the proposed HAGMN-UQ.

17) Original text: The FLOPs is dependent on the number of vertices and hyperedges in the individual graph, and the number of the compared graphs.

Revised text: The FLOPs depend on the number of vertices and hyperedges in the individual graph, as well as the number of graphs being compared.

18) Original text: Additionally, we extended our investigation to validate the balancing factor between the permutation loss, as defined in Eq. 19, and the confidence loss, as defined in Eq. 22.

Revised text: Additionally, we extended our investigation to validate the balance factor between the permutation loss, as defined in Eq. 19, and the confidence loss, as defined in Eq. 22.

**Reviewer #2**: The authors addressed my questions well, and I particularly appreciated the detailed steps provided for implementing pre-processing, including vessel segmentation and centerline extraction. I have no more concerns.

A: Thank you for your positive feedback. We're glad to receive your feedback that the detailed steps for implementing pre-processing, including vessel segmentation and centerline extraction, were helpful. We appreciate your thorough review and are pleased that we have addressed all your concerns.

**Reviewer #3**: The author's answer partially solved my concerns. However, there are still several very serious problems:

1. The method provided a large number of assumptions or automatic and manual ways to obtain the complete coronary branches, which solved most of the problems of coronary artery semantic labeling. The remaining problem is simply how to classify them. This makes the problem much less challenging, meanwhile it changes the task of this paper from coronary semantic labeling to how to perform graph matching. Therefore, I do not think this paper can be called coronary artery semantic labeling.

A: Thanks for your questions and comments. Our primary objective is indeed to achieve accurate coronary artery semantic labeling, and we acknowledge that the successful extraction of individual arterial branches is a significant aspect of this task. However, classifying coronary arterial branches accurately remains challenging due to the morphological similarity among different types of arteries, the variability in anatomical structures and varying arterial anatomy under different projection view angles.

*1) Morphological similarity between different arterial branches*

Pixel-intensity-based models face challenges in distinguishing between individual arterial segments and generating semantic segmentation due to the morphological similarity among branches in the coronary vascular tree and the overlap of arteries in 2D (projection) ICAs, as shown in Figure RL1.

A collage of images of a vein

Description automatically generated

Figure RL1. Challenging examples for coronary artery semantic labeling.

Figure RL1 highlights these challenges, where yellow rectangles indicate arterial overlaps in 2D. In these three examples, the LAD branches from the first two ICAs exhibit similar morphological and pixelwise features to the D branch of the third ICA. This demonstrates the difficulties in coronary semantic segmentation caused by morphological similarity.

*2) Varying anatomy across different ICA projection view angles*

ICA is a key diagnostic procedure used to visualize the coronary arteries and detect any blockages or abnormalities. In clinical practice, several different angiographic views are utilized to obtain comprehensive images of the coronary anatomy during this procedure. The varying anatomy across different ICA projection view angles ensures a thorough assessment of the coronary vasculature from multiple perspectives; however, identifying arterial branch is challenging. The key views used in ICA include:

* Left Anterior Oblique (LAO): The x-ray beam is directed from the patient's left side. The patient's body is rotated to the right, and the heart is imaged from the left anterior aspect. This view helps in visualizing the LMA artery, LAD and LCX.
* Right Anterior Oblique (RAO): The x-ray beam is directed from the patient's right side, with the patient's body rotated to the left. This view provides detailed images of the right coronary artery (RCA) and parts of the left coronary system.
* Anteroposterior (AP): The x-ray beam is directed from the front of the patient to the back (anteroposterior direction). This straightforward projection is useful for assessing the coronary arteries and can be particularly helpful for evaluating stents and grafts.
* Caudal (CAU): The x-ray beam is angled towards the patient's feet. This view is particularly useful for visualizing the left main coronary artery bifurcation and the proximal segments of the LAD and LCX.
* Cranial (CRA): The x-ray beam is angled towards the patient's head. This view is useful for visualizing the distal segments of the coronary arteries, including the mid and distal LAD.

We add the description of ICA image scanning technique to the revised manuscript in the introduction section.

(Section 1. Introduction)

In clinical practice, by rotating the X-ray beam horizontally, three view angles are generated: Left Anterior Oblique (LAO), Right Anterior Oblique (RAO), and Anteroposterior (AP). By rotating the X-ray beam vertically, Caudal (CAU) and Cranial (CRA) projection angles are produced. Combining these projection views results in six commonly used view angles: LAO\_CAU, RAO\_CAU, AP\_CAU, LAO\_CRA, RAO\_CRA, and AP\_CRA. Each of these views provides unique and complementary information about coronary anatomy, which is essential for accurate diagnosis and treatment planning. The choice of views depends on the specific coronary segments of interest and the clinical scenario. However, the location, orientation, and position of different arterial branches vary significantly, as shown in Figure 1. This highlights the complexity of accurately extracting coronary arteries, given the significant anatomical variability between patients.

A collage of images of a human body

Description automatically generated

Figure 1. Representative examples of ICAs under different view angles.

Figure 1 showcases representative ICA images captured under various view angles, as depicted in the central scatter plot. Each dot in the scatter plot represents an ICA image from our dataset, with different colors indicating the specific view angles. The x-axis (Primary Angle) indicates the horizontal rotation of the x-ray beam, yielding RAO, AP, and LAO projection angles. The y-axis (Secondary Angle) represents the vertical rotation of the x-ray beam, resulting in CRA and CAU projection angles. The surrounding images illustrate that despite being captured at the same view angle, the coronary arteries exhibit significant variability in topology and position. This highlights the inherent anatomical differences between individuals and underscores the complexity of interpreting ICA images. Thus, even with the manual or semi-automatic generated ICA graphs, coronary artery semantic labeling is still challenging.

Regarding semantic segmentation and semantic labeling, here is the detailed explanation. Semantic segmentation is a pixel-level labeling task where the goal is to classify each pixel in an image into a specific class/category. For example, in the context of coronary artery segmentation, semantic segmentation would involve labeling each pixel in an angiogram as belonging to one of the coronary arteries (left anterior descending artery, left circumflex artery, right coronary artery, etc.) or background. Semantic labeling, on the other hand, is a broader task that involves assigning meaningful labels to individual objects or regions of interest within an image. It focuses more on the object-level or region-level labeling rather than pixel-level classification. In addition, the term of semantic labeling has been applied to coronary artery identification task in several papers, such as [1–6], published in high impact journals and conferences. Our approach initially extracts the entire coronary arterial tree and divides it into distinct branches. Subsequently, we utilize graph matching to classify each arterial branch into various categories. Therefore, based on the above description, the proposed method falls under coronary artery semantic labeling.

The manual or semi-automated annotation of coronary arteries is the first and most crucial step in artery identification. Coronary artery segmentation poses challenges at three distinct image levels: surrounding, local and semantic [7]. At the surrounding level, the low contrast between foreground vessels and the background is problematic. This is primarily caused by X-ray power and contrast agent limitations, hindering accurate foreground extraction and leading to incomplete vascular segment delineation. Additionally, difficulties in excluding non-vessel areas arise due to their similar appearance to vascular segments, such as catheter outlines, spines, and ribs, complicating stenosis severity assessment. At the local level, local ambiguity near coronary vessel boundaries poses a challenge. High-frequency detail loss during the 3D to 2D projection in ICA imaging results in smooth grayscale boundaries instead of distinct steps, adversely affecting vessel boundary delineation. This ambiguity is particularly critical near stenosis areas, impacting the accuracy of stenosis severity assessment. Furthermore, due to patient specific anatomy, the variations of projection angles, and contrast dye degradation, recognizing the type of individual coronary arteries remains challenging. According to training guidelines provided by the American College of Cardiology Core Cardiology Training Symposium, it is recommended that trainees spend 4 months and achieve a minimum cumulative number of 100 cases to understand coronary anatomy sufficiently to perform independent diagnostic cardiac catheterization [8]. Furthermore, the European Society of Cardiology recommends that a general cardiologist assist with or perform 300 coronary angiographies and interpret 1000 investigations [9]. Thus, manual annotation of coronary arteries is time-consuming and challenging. Therefore, automated identification of individual arteries is essential for improving efficiency and accuracy in cardiac diagnostics, especially for junior cardiologists and trainees.

Admittedly, building an end-to-end model for coronary artery semantic segmentation is more straightforward because it directly obtains pixel-to-pixel semantic mapping. However, as described in the introduction and discussion sections, the topology and anatomy vary significantly between patients, while the pixel intensities remain similar for different types of arteries. This makes it challenging for CNN-based end-to-end models to differentiate individual arteries. Most existing studies focus on major artery segmentation, as follows.

* Jun et al. proposed a T-Net [10], a nested encoder-decoder, to extract three major arteries, including LCX, LAD, and the right coronary artery (RCA) from the right ventricle. However, side branches such as OM1, OM2, D1, and D2 were not included.
* Zhang et al. proposed a Progressive Perception Learning framework [7] to extract the three main coronary arteries, including LCX, LAD, and RCA using ICAs.
* Xian et al. employed fully convolutional networks to extract LCX, LAD, and RCA using ICA videos [11].

Existing studies have primarily focused on the extraction of major arteries, often ignoring the side branches. Additionally, their models can only extract one type of artery at a time. In contrast, our model can extract multiple arteries simultaneously, providing a comprehensive understanding of the coronary artery vasculature and offering more information for clinical applications. We added this related paper into the related work section in the revised manuscript.

(Section 2.1 Coronary artery semantic labeling)

Pixel-to-pixel based approaches achieved impressive results in segmenting main arteries. Jun et al. proposed a T-Net [10], a nested encoder-decoder, to extract three major arteries, including LCX, LAD, and the right coronary artery (RCA) from the right ventricle. However, side branches such as OM1, OM2, D1, and D2 were not included. Zhang et al. proposed a Progressive Perception Learning framework [7] to extract the three main coronary arteries, including LCX, LAD, and RCA using ICAs. Similarly, Xian et al. employed fully convolutional networks to extract LCX, LAD, and RCA using ICA videos [11]. Existing studies have primarily focused on the extraction of major arteries, often ignoring the side branches. Additionally, their models can only extract one type of artery at a time. In contrast, our model can extract multiple arteries simultaneously, providing a comprehensive understanding of the coronary artery vasculature and offering more information for clinical applications.

In our previously published papers, AGMN [4] and EAGMN [5], we adapted the off-the-shelf graph neural network explainer, ZORRO [12], to elucidate the interpretability of graph matching in coronary artery semantic labeling tasks. Furthermore, topology—specifically the connectivity between coronary artery segments—and topology-based and positional features are the most informative in identifying arterial segment categories. In the revised manuscript, we applied ZORRO to explain the feature importance in HAGMN-UQ. The top-15 most informative features are shown in Figure S7.

(Supplementary materials, Section 9. Interpretation of HAGMN-UQ)

As discussed in the introduction section, the topology—specifically the connectivity between coronary artery segments—and topology-based and positional features are the most informative in identifying arterial segment categories. To verify the significance of the approach and explain the graph matching, we employ a perturbation-based method, ZORRO [12], to explain the feature importance of in HAGMN-UQ. ZORRO employs discrete feature masks and node masks to determine the significance of features and nodes in decision-making. Given an input graph, ZORRO iteratively and recursively adds important features and nodes based on the fidelity score, which measures the difference between the original predictions and the new predictions after masking out important features and nodes.

Specifically, ZORRO utilizes a hard mask, where a value of 0 indicates that a feature or node is not selected, and a value of 1 indicates that a feature or node is selected as important. Additionally, ZORRO explains the importance of features and nodes for each input graph at each iteration. As a hard mask selection algorithm, ZORRO aims to select the top- most frequently retrieved features and identify the most important nodes for graph representation learning.

To explain feature importance, we modified ZORRO, originally proposed to explain GNN for node classification tasks. Recall that we converted the graph matching problem into a vertex classification task using hyper-association graphs; identifying the concatenated features for each vertex does not explain the feature importance for the node in the individual graph. In our implementation, a unified feature mask is employed to mask out the selected features for and simultaneously. Then, the retained features from these two individual graphs are concatenated to form the features for each vertex in the association graph. In this paper, we applied ZORRO to explain the feature importance in HAGMN-UQ. The top-15 most informative features are shown in Figure S7. where the degree represents the number of connections a segment has in the vascular tree-generated ICA-graph, as defined in our paper in .

A colorful bar chart with black text

Description automatically generated

Figure S7. Feature importance ranking for classifying coronary arterial segments using HAGMN-UQ. Important features were surrogated by the frequency of the selected features during explaining all graph-matching pairs using the post-hoc GNN explaining algorithm, ZORRO [15]. The vertical axis indicates the feature names, while the horizontal axis indicates the frequencies of the features, which are also denoted as feature importance.

During our testing, we obtained 1658 association graphs, where is from testing set and is from template set. The frequency of the selected features is denoted as feature importance. For example, p2\_degree showed a frequency of 100%, indicating that each time we perform graph matching, this feature should be included, highlighting its importance. For the top 15 features with the highest frequency, as well as the most informative features, they all belong to topological features and position features. Specifically, p2\_degree and p1\_degree represent the degree of the arterial segments, where the degree represents the number of connections a segment has in the vascular tree-generated ICA-graph, as defined in our paper in . The y\_center and x\_center, including weighted centers, represent the position and coordinates of the arterial segments. The p1 and p2 related centers represent the position and coordinates of the endpoints of each arterial segment. Figure S7 further validated our hypothesis: recognizing the significance of topology in arterial identification, we are motivated to transform arteries and their interconnections into graph structures. Thus, incorporating both topology and positional features for coronary artery semantic labeling is practical and reasonable.

2. As a manual intervention is used, the operation of labeling the binary mask still needs to be performed by experienced experts. Therefore, they already know the semantic labeling for that coronary artery tree. This makes the motivation for this paper poor.

A: Thank you for your insightful feedback on our manuscript. We appreciate your concern regarding the reliance on manual intervention for semantic segmentation and the implication that experienced experts are required to label the binary masks, potentially diminishing the motivation for our work. As shown in previous question, identification of individual coronary arteries is challenging, as different ICAs show significant different anatomy and arteries vary between different view angles.

We would like to further address this concern as follows:

*Reduction of Manual Effort*: While our current method involves some degree of manual intervention, it significantly reduces the amount of manual effort compared to traditional approaches. The initial labeling by experts is utilized to train the model, but once trained, the model can automatically label coronary arteries in new images with high accuracy. This greatly minimizes the need for continuous expert involvement.

*Scalability and Practicality*: The ability of our model to simultaneously extract and label multiple arteries, including side branches, makes it highly scalable and practical for large datasets. This provides a comprehensive understanding of coronary artery vasculature, which is crucial for various clinical applications, such as planning interventions and monitoring disease progression.

*Future Work*: Further reducing the need for manual intervention by incorporating advanced techniques such as semi-supervised and unsupervised learning should be investigated. It is worth noting that for AI/deep learning-related algorithms and models, some level of human intervention is always required. Fully replacing human expertise is not feasible in clinical practice. The proposed method is intended to provide auxiliary support for junior cardiologists to enhance their understanding of coronary artery anatomy.

We describe the required training for coronary artery semantic labeling and state the challenges in the introduction section in the revised manuscript.

(Section 1. Introduction)

According to training guidelines provided by the American College of Cardiology Core Cardiology Training Symposium, it is recommended that trainees spend a minimum of 4 months of dedicated time in the cardiac catheterization laboratory interpreting coronary angiography and perform a minimum of 100 cases to obtain proficiency in performing and interpreting cases independently [8]. Furthermore, the European Society of Cardiology recommends that a general cardiologist assist with or perform 300 coronary angiograms and interpret 1000 investigations [9]. Thus, manual annotation of coronary arteries is time-consuming and challenging, requiring adjudication by experienced cardiologists with dedicated training in the field.

**Reference**

[1] Wu D, Wang X, Bai J, Xu X, Ouyang B, Li Y, Zhang H, Song Q, Cao K, Yin Y. Automated anatomical labeling of coronary arteries via bidirectional tree LSTMs. Int J CARS. 2019 Feb;14(2):271–280.

[2] Yang H, Zhen X, Chi Y, Zhang L, Hua X-S. CPR-GCN: Conditional Partial-Residual Graph Convolutional Network in Automated Anatomical Labeling of Coronary Arteries. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Seattle, WA, USA: IEEE; 2020. p. 3802–3810. Available from: https://ieeexplore.ieee.org/document/9156834/

[3] Cao Q, Broersen A, de Graaf MA, Kitslaar PH, Yang G, Scholte AJ, Lelieveldt BPF, Reiber JHC, Dijkstra J. Automatic identification of coronary tree anatomy in coronary computed tomography angiography. Int J Cardiovasc Imaging. 2017 Nov;33(11):1809–1819.

[4] Zhao C, Xu Z, Jiang J, Esposito M, Pienta D, Hung G-U, Zhou W. AGMN: Association graph-based graph matching network for coronary artery semantic labeling on invasive coronary angiograms. Pattern Recognition. 2023 Nov;143:109789.

[5] Zhao C, Xu Z, Hung G-U, Zhou W. EAGMN: Coronary artery semantic labeling using edge attention graph matching network. Computers in Biology and Medicine. 2023 Nov;166:107469.

[6] Zhang X, Cui Z, Feng J, Song Y, Wu D, Shen D. CorLab-Net: Anatomical Dependency-Aware Point-Cloud Learning for Automatic Labeling of Coronary Arteries. In: Lian C, Cao X, Rekik I, Xu X, Yan P, editors. Machine Learning in Medical Imaging [Internet]. Cham: Springer International Publishing; 2021 [cited 2024 Jul 8]. p. 576–585. Available from: https://link.springer.com/10.1007/978-3-030-87589-3\_59

[7] Zhang H, Gao Z, Zhang D, Hau WK, Zhang H. Progressive Perception Learning for Main Coronary Segmentation in X-ray Angiography. IEEE Transactions on Medical Imaging. 2022;1–1.

[8] Jacobs AK, Babb JD, Hirshfeld JW, Holmes DR, Society for Cardiovascular Angiography and Interventions. Task force 3: training in diagnostic and interventional cardiac catheterization endorsed by the Society for Cardiovascular Angiography and Interventions. J Am Coll Cardiol. 2008 Jan 22;51(3):355–361. PMID: 18206752

[9] Jensen UJ, Jensen J, Olivecrona G, Ahlberg G, Lagerquist B, Tornvall P. The role of a simulator-based course in coronary angiography on performance in real life cath lab. BMC Med Educ. 2014 Dec;14(1):49.

[10] Jun TJ, Kweon J, Kim Y-H, Kim D. T-Net: Nested encoder–decoder architecture for the main vessel segmentation in coronary angiography. Neural Networks. 2020 Aug;128:216–233.

[11] Xian Z, Wang X, Yan S, Yang D, Chen J, Peng C. Main Coronary Vessel Segmentation Using Deep Learning in Smart Medical. Huang C, editor. Mathematical Problems in Engineering. 2020 Oct 21;2020:1–9.

[12] Funke T, Khosla M, Rathee M, Anand A. ZORRO: Valid, Sparse, and Stable Explanations in Graph Neural Networks. IEEE Trans Knowl Data Eng. 2023;1–12.