A Causal-holistic Adaptive Intervention Network for Tailoring Automated Coronary Artery Disease Diagnosis to Individual Patients

Xinghua Ma[1], Xingyu Qiu[1], Yuetan Chu[3,4,5], Kuanquan Wang[1], Zhaowen Qiu[2], Gongning Luo[1✉], and Xin Gao[3,4,5]

[1] Faculty of Computing, Harbin Institute of Technology, Harbin, China
luogongning@hit.edu.cn
[2] College of Computer and Control Engineering, Northeast Forestry University, Harbin, China
[3] Computer Science Program, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[4] Center of Excellence for Smart Health, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[5] Center of Excellence on Generative AI, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Abstract. Given the global prevalence and high mortality of coronary artery disease (CAD), automated CAD diagnosis should evolve toward personalized methods to maximize its clinical value. However, existing techniques have been limited to artery-level prediction, lacking patient-level causality and failing to effectively account for individual patient confounders. In this work, for the first time, we introduce a Causal-Holistic Adaptive Intervention Network (CAIN) that tailors personalized CAD diagnosis for individual patients. CAIN generates semantic representations at both the patient and artery dual-levels for each case, constructing a holistic causal graph that captures individual-specific characteristics. It then implements adaptive causal intervention based on the patient's specific condition, using dynamically updated and differentiated intervention variables. Experimental results on CCTA scans from 602 patients and 6,830 coronary branches across three clinical centers show that CAIN outperforms state-of-the-art methods, offering personalized clinical guidance. The source code is available at (`https://github.com/PerceptionComputingLab/CAIN`).

Keywords: Coronary CT Angiography (CCTA) · Coronary artery disease · Causal intervention · Personalized diagnosis.

1 Introduction

Coronary Artery Disease (CAD) has long been a leading cause of morbidity and mortality worldwide, as atherosclerosis narrows the coronary artery lumen, leading to reduced blood flow to the myocardium. [12] Recent clinical studies indicate

a rising prevalence of CAD, which has led to a growing demand for automated diagnostics utilizing Coronary CT Angiography (CCTA). [14] To maximize clinical applicability, automated CAD diagnosis should evolve from generalized, rigid assessments to personalized analyses [11], ensuring that clinical guidance aligns with each individual's unique condition.

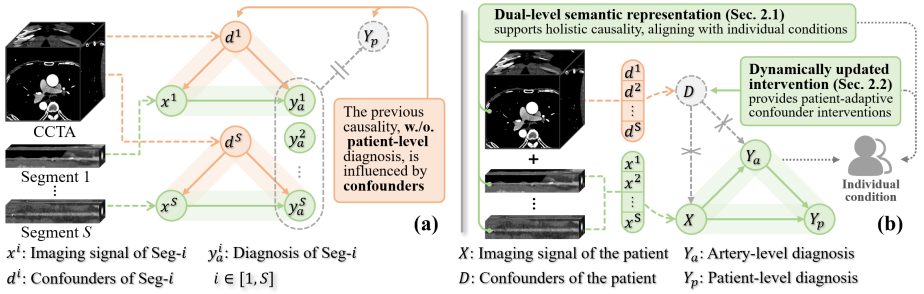

Fig. 1. Our CAIN applies adaptive causal intervention (b) within each patient's holistic causal graph to overcome the limitations (a) of previous artery-level methods.

Current automated CAD diagnosis technology has made significant progress [1],[8], largely driven by deep learning. However, existing methods focus primarily on artery-level predictions and lack personalized analysis tailored to individual patients. Zreik et al. [20] introduced a method using Curved Planar Reformation (CPR), integrating convolutional and recurrent neural networks to classify image cubes from coronary segments. Denzinger et al. [5] and Ma et al. [9] refined this method by integrating gated recurrent units and Transformer models to analyze spatial correlations within the segments. More recently, Zhang et al. [19], Van et al. [16], and Ma et al. [10] have proposed advanced techniques with diverse research focuses, improving the lesion analysis through Faster R-CNN, surface mesh inference, and data-efficient learning. Despite these advancements, all the above-mentioned methods use coronary segments as the largest input unit, limiting predictions to the artery-level. Only a few studies [20],[16] have attempted simple statistical aggregation of artery-level results to approximate patient-level verification. As a result, relying on branch segments as input prevents these methods from capturing patient-level characteristics and providing truly personalized diagnoses.

As illustrated in the causal graph (Fig.1a) of the automated CAD diagnosis process, the limitations of previous artery-level methods in the context of personalized diagnosis are reflected in the following two aspects: (1) Ignoring the imaging signals of the patient's coronary tree and relying on CPR-based artery-level predictions for each coronary segment leads to the absence of the patient-level prediction $\boldsymbol{Y}_p$ and its corresponding causal link in the causal graph. Given that clinical standards [4] explicitly define the crucial role of the Coronary Artery Disease Reporting & Data System (CAD-RADS) assessment in guiding

patient treatment plans, the absence of such causal relationships significantly reduces the clinical applicability of related technologies. (2) In artery-level predictions $\boldsymbol{y}_a^i$ based on imaging signals $\boldsymbol{x}^i$, false associations, such as $\boldsymbol{d}^i \rightarrow \boldsymbol{x}^i$ and $\boldsymbol{d}^i \rightarrow \boldsymbol{y}_a^i$, are introduced by confounders $\boldsymbol{d}^i$ disrupt the causality $\boldsymbol{x}^i \rightarrow \boldsymbol{y}_a^i$, potentially resulting in misdiagnosis or missed diagnosis. Moreover, this error is further amplified during the statistical aggregation process used to approximate the patient-level prediction $\boldsymbol{Y}_p$. Recent clinical studies have identified both external and internal factors as potential confounders. External factors include data heterogeneity caused by variations in scanning instruments and imaging parameters [17], while internal factors include lesion-irrelevant information such as blurred boundaries or imaging artifacts [2]. Given these limitations, the causality in the previous coronary artery-level CAD diagnostic process can be expressed as:

$$P(\boldsymbol{y}_a^i|\boldsymbol{x}^i) = P(\boldsymbol{y}_a^i|\boldsymbol{x}^i, \boldsymbol{d}^i)P(\boldsymbol{d}^i|\boldsymbol{x}^i) \quad i \in \{1, 2, \ldots, S\} \tag{1}$$

where S denotes the number of coronary branches in the patient's coronary tree.

In this work, we propose a Causal-holistic Adaptive Intervention Network (CAIN) for tailoring personalized CAD diagnosis to individual patients. To address the limitations of artery-level methods, CAIN constructs a holistic causal graph aligning with individual conditions that includes $\boldsymbol{Y}_p$, and describes the whole coronary tree to facilitate adaptive causal intervention for the current patient, as shown in Fig.1b. Specifically, the dual-level semantic representation, which integrates patient-level CCTA scans and artery-level CPR volumes, supports the construction of a holistic causal graph. Next, the confounders for adaptive causal intervention are retrieved from a dynamically updated bank, enabling personalized diagnoses at both the patient and artery levels. Our main contributions are outlined as follows: (1) CAIN introduces, for the first time, a personalized solution for CAD diagnosis in individual patients through the construction of a holistic causal graph and adaptive causal intervention. (2) Comprehensive experiments based on a dataset of 602 patients, 6,830 branches from three clinical centers demonstrate that CAIN outperforms State-Of-The-Art (SOTA) methods, particularly at the patient level.

2 Method

The personalized CAD diagnostic network, CAIN, consists of two modules: dual-level semantic representation and dynamically updated intervention, as shown in Fig.2. These modules apply the intervention to the confounders $\boldsymbol{D}$ within the causal graph (Fig.1b) which incorporates patient-level prediction $\boldsymbol{Y}_p$:

$$\begin{aligned} P(\boldsymbol{Y}_p|\mathrm{do}(\boldsymbol{X})) = &\textstyle\sum_{\boldsymbol{x}^i}\sum_{\boldsymbol{y}_a^i}\sum_{\boldsymbol{d}^i} P(\boldsymbol{Y}_p|\boldsymbol{X}=\boldsymbol{x}^i, \boldsymbol{Y}_a=\boldsymbol{y}_a^i, \boldsymbol{D}=\boldsymbol{d}^i) \\ &P(\boldsymbol{Y}_a|\boldsymbol{X}=\boldsymbol{x}^i)P(\boldsymbol{D}=\boldsymbol{d}^i) \quad i \in \{1, 2, \ldots, S\} \end{aligned} \tag{2}$$

2.1 Dual-level semantic representation

The module comprehensively represents the imaging signals of the individual patient at both the artery and patient levels, providing a personalized semantic

foundation for constructing the holistic causal graph. Specifically, the patient-level CCTA scans $\boldsymbol{x}_{cta}$ and the artery-level CPR volumes $\boldsymbol{x}_{cpr}^{i}$ $i \in [1, S]$ are based on the coronary tree as input, where S denotes the number of branches. According to clinical research statistics [6], representing the 16 major coronary branches most strongly associated with CAD, including RCA-p/m/d, L-PDA, LM, LAD-p/m/d, D-1/2, LCX-p/d, OM, L-PLB, R-PDA, and R-PLB.

The patient-level representation captures the overall condition of the coronary tree and provides global features for characterizing individual patients. To obtain this representation, the patient-level feature extractor, $\mathcal{F}_{cta}$, partitions the 3D input $\boldsymbol{x}_{cta}$ into a set of cubic patches. These patches are then processed through a series of Transformer encoders employing the shift window [7] mechanism. Each encoder performs cascaded multi-head attention calculations, with cyclic patch shifting applied or not. Following each encoder, a patch merging operation is applied to downsample features. The extracted features are then transformed into a patient embedding, $\boldsymbol{e}_p \in \mathbb{R}^{32\times512}$, using pointwise convolution and projection. The artery-level representation details the vascular lumen of each branch, extracting local features from all branches to represent the coronary tree. To obtain this representation, the artery-level feature extractor, $\mathcal{F}_{cpr}$, extracts artery embeddings for different branches from their corresponding CPR volumes. $\boldsymbol{x}_{cpr}^{i}$ $i \in [1, S]$ corresponds to each artery branch and undergoes multi-scale lumen information extraction through multiple convolutional blocks with 3D max-pooling layers, enabling comprehensive analysis of branch-specific conditions and providing an artery embedding $\boldsymbol{e}_a^i$ $i \in [1, S]$ for each coronary branch.

To establish associations among multiple artery branches in the latent variable space and capture inter-branch relationships, artery embeddings from different branches, $[\boldsymbol{e}_a^1, \cdots, \boldsymbol{e}_a^S]$, are processed using a series of Transformer encoders $\mathcal{E}_{tr}^{i}$ $(i \in [1, k_1])$ for branch association computation. Next, the query embeddings $\boldsymbol{e}_{qry}^{i} \in \mathbb{R}^{Q\times512}$ $i \in [1, S]$, which are randomly initialized and correspond to the branches, interact with the respective embeddings through a series of Transformer decoders $\mathcal{D}_{tr}^{i}$ $(i \in [1, k_2])$, to adaptively integrate multi-branch information into $\boldsymbol{e}_{qry}^{i}$. This branch latent association reduces information redundancy while providing prior signals for causal interventions.

2.2 Dynamically updated intervention

The module continuously updates the confounder bank based on differentiated samples from individual patients, enabling adaptive causal interventions for personalized diagnoses. To ensure intervention variables in the confounder bank $\mathcal{B} \in \mathbb{R}^{M\times512}$ accurately represent each patient's physiological characteristics, it is constructed through a dynamic update procedure based on diverse patient-level samples during training. Here, M denotes the maximum capacity of the confounder bank for storing vectors. Specifically, the feature vectors $\boldsymbol{v}_F$s, obtained through dimensionality reduction from the dual-level embeddings $\boldsymbol{e}_p$ and $\boldsymbol{e}_a^i$ $(i \in [1, S])$ of the patient samples, are sequentially enqueued into the bank $\mathcal{B}$. If the number of confounder variables in $\mathcal{B}$ has not yet reached the storage limit M, $\boldsymbol{v}_F$ is pushed into $\mathcal{B}$. Otherwise, the intervention variable $\boldsymbol{v}_I$ in $\mathcal{B}$ that

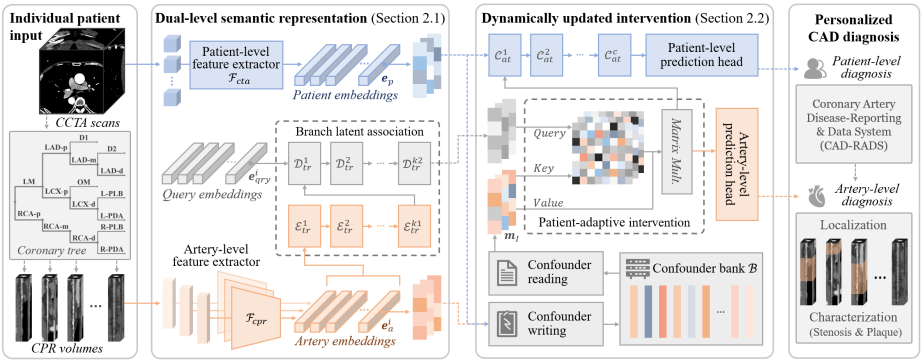

Fig. 2. The overview of CAIN: Dual-level semantic representation and dynamically updated intervention implement the adaptive causal intervention on the holistic causality.

has the highest cosine similarity to $\boldsymbol{v}_F$ is identified. The two are then aggregated and stored in the place of $\boldsymbol{v}_I$ in $\mathcal{B}$, which is defined as:

$$\tilde{\boldsymbol{v}}_I = \beta \boldsymbol{v}_I + (1-\beta)\boldsymbol{v}_F \tag{3}$$

where β is a hyperparameter regulating the aggregation, and $\tilde{\boldsymbol{v}}_I$ is the aggregated variable that is subsequently written back into $\mathcal{B}$. This updating strategy allows the confounder bank to capture sample diversity within limited storage space, thereby ensuring adaptive causal interventions across patient variations.

To design adaptive intervention based on the imaging features of the entire coronary tree, we select intervention variables from all S branches to ensure a comprehensive description of the patient's confounders. Given the causal model $P(\boldsymbol{Y}|\text{do}(\boldsymbol{X}))$, the modeling process $\sum_{\boldsymbol{d}}(P(\boldsymbol{y}_l|\boldsymbol{x},\boldsymbol{d}))p(\boldsymbol{d})$ for predicting the l-label $\boldsymbol{y}_c$ based on information $\boldsymbol{x}$ can be rewritten as $\mathbb{E}_{\boldsymbol{d}}[\text{Softmax}(f_{\boldsymbol{d}}(\boldsymbol{x},\boldsymbol{d}))]$ under Softmax-based probability prediction. According to the Normalized Weighted Geometric Mean (NWGM) [18], the expected value after causal modeling is given by:

$$\mathbb{E}_{\boldsymbol{d}}[\text{Softmax}(f_{\boldsymbol{d}}(\boldsymbol{x},\boldsymbol{d}))] \overset{\text{NWGM}}{\approx} \text{Softmax}(\mathbb{E}_{\boldsymbol{d}}[f_{\boldsymbol{d}}(\boldsymbol{x},\boldsymbol{d})]) \tag{4}$$

Thus, causal intervention can be transformed into a probabilistic prediction within a context that accounts for confounders. In the process of CAD diagnosis, the intervention medium $\boldsymbol{m}_I \in \mathbb{R}^{32\times512}$, which represents the confounders, is obtained through confounder reading. This medium is formed by concatenating the intervention variables that exhibit the highest cosine similarity to $\boldsymbol{e}_p$ and $\boldsymbol{e}_a^i$ ($i \in [1,S]$) from $\mathcal{B}$, followed by projection. The query embedding $\boldsymbol{e}_q^i$ ($i \in [1,S]$) for each coronary branch, along with the medium $\boldsymbol{m}_I$, is used to compute the $Query$, Key, and $Value$ through linear mapping. Attention mechanism-based distributional computations capture the discrepancy between image representation and confounders, thereby mitigating the interference of confounders in the CAD diagnosis. Following causal intervention, $\boldsymbol{e}_q^i$ of each branch is fed into the artery-level prediction head with parallel regression and classification heads

for lesion localization, and stenosis and plaque characterization. All $\boldsymbol{e}_q^i$ are concatenated and integrated with the patient embedding $\boldsymbol{e}_p$ through a series of cross-attention calculations $\mathcal{C}_{at}^i$ $(i \in [1, c])$, integrating the dual-level semantics to enable the patient-level CAD-RADS classification.

The objective function of CAIN consists of a lesion localization loss $\mathcal{L}_{loc}$ and a lesion characterization loss $\mathcal{L}_{char}$, which corresponds to each i-branch, as well as a CAD-RADS classification loss:

$$\mathcal{L}_{\text{CAIN}} = \textstyle\sum_{i=1}^{S}(\mathcal{L}_{loc}^i + \mathcal{L}_{char}^i)/S - \sum_j \boldsymbol{y}_{rads}^j \log(\boldsymbol{p}_{rads}^j) \tag{5}$$

where $\boldsymbol{y}_{rads}$ and $\boldsymbol{p}_{rads}$ are one-hot encoded vectors for the CAD-RADS labels and the class probability of CAD-RADS predictions, respectively. The definitions of $\mathcal{L}_{loc}$ and $\mathcal{L}_{char}$ are defined as follows:

$$\begin{aligned} \mathcal{L}_{loc} &= \textstyle\sum_{j=1}^{Q}[-\log \hat{p}_{\hat{\varrho}(j)}(c_j) + \mathbb{1}_{\{c_j \neq \varnothing\}}\mathcal{L}_{roi}] \\ s.t.,\ \hat{\varrho} &= \underset{\varrho \in \mathfrak{S}_Q}{\arg\min} \textstyle\sum_j^Q [-\mathbb{1}_{\{c_j \neq \varnothing\}} \hat{p}_{\varrho(j)}(c_j) + \mathbb{1}_{\{c_j \neq \varnothing\}}\mathcal{L}_{roi}] \\ \mathcal{L}_{char} &= -1/2(\textstyle\sum_j \boldsymbol{y}_{sten}^j \log(\boldsymbol{p}_{sten}^j) + \sum_k \boldsymbol{y}_{plq}^k \log(\boldsymbol{p}_{plq}^k)) \end{aligned} \tag{6}$$

where $\varrho \in \mathfrak{S}Q$ is the permutation that minimizes the cost using Hungarian matching [3], $\hat{p}_{\varrho(j)}(c_j)$ denotes the probability of category label c_j, and $\mathcal{L}_{roi}$ is the RoI loss [13]; $\boldsymbol{y}_{sten,plq}$ and $\boldsymbol{p}_{sten,plq}$ are one-hot encoded vectors for the plaque and stenosis labels and the class probability for their predictions, respectively.

3 Experiment

Dataset. A total of 602 CCTA scans from different patients were retrospectively collected from three clinical centers (352/32/218). 6,830 coronary branch CPR volumes (4,039/320/2,471) were reconstructed, covering the LM, LAD, LCX, and RCA branches for all patients. To account for individual patient variability, any missing branches were replaced with volumes set to -1024 HU. To enhance experimental stability, CCTA scans underwent geometric augmentation, and CPR reconstruction included random centerline shifts of up to three voxels.

Implementation details. The inputs to CAIN, $\boldsymbol{x}_{cta}$ and CPR volumes $\boldsymbol{x}_{cpr}$, were initialized to 128×128×128 and 256×64×64, respectively. $\mathcal{F}_{cta}$ and $\mathcal{F}_{cpr}$ adopt an identical encoding architecture to that of [15] and [10], respectively. The quantities k_1, k_2, and c, corresponding to $\mathcal{E}_{tr}$, $\mathcal{D}_{tr}$, and $\mathcal{C}_{at}$, were each set to 4. The query number Q in $\boldsymbol{e}_{qry}$ was set to 16, and the maximum capacity M of $\mathcal{B}$ was set to 1000. The hyperparameter β of the confounding writing was set to 0.65. The model weights achieving the best validation performance after 200 epochs were used for testing.

Evaluation metrics. In accordance with clinical standards and previous studies, evaluations were performed at both the artery and patient levels, incorporating the individual CAD-RADS level as well as the stenosis and plaque characteristics of each artery segment. The evaluation metrics adopted from the

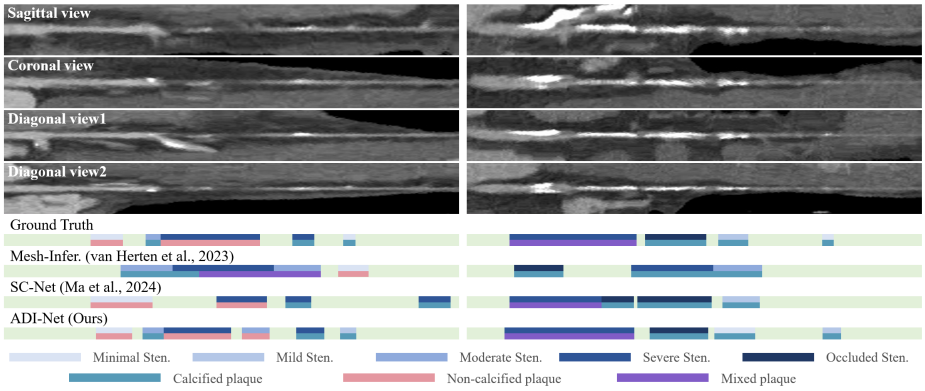

Fig. 3. Qualitative comparison with SOTAs in lesion localization and characterization.

SOTA methods included accuracy (ACC), precision (Prec), recall, F1 score, negative predictive value (NPV), and specificity (Spec), calculated based on the confusion matrix. To assess the cross-center diagnostic capability, data from Center 1 were used for training, Center 2 for validation, and Center 3 for testing.

3.1 Experimental results

Comparison with SOTAs. To evaluate the performance of CAIN for personalized CAD diagnosis, we compared it with SOTAs for automated CAD diagnosis. In the study by Denzinger et al. [5], plaque was not considered, so only stenosis-related metrics were evaluated for 2D-RCNN+PT. As shown in Tab.1, the quantitative results show that CAIN significantly outperforms SOTAs across all tasks at both the artery and patient levels, with t-test p-values all below 0.05. At the artery-level, CAIN accurately localizes lesions and characterizes stenosis and plaque in each coronary branch. At the patient level, it reliably predicts the CAD-RADS level, thus assessing the risk of disease for each individual patient. In the evaluations above, experiments with dataset splits based on clinical centers also showed that CAIN can handle imaging variations caused by different imaging parameters, further proving its cross-center generalization capability and clinical practicality. Additionally, CAIN shows superior performance in Precision, Recall, and F1 scores, emphasizing its ability to effectively reduce both misdiagnosis and missed diagnoses, thereby optimizing the allocation of medical resources through early screening. As shown in Fig.3, qualitative comparisons with the latest SOTAs indicate that CAIN can accurately localize lesions within coronary segments and provide precise assessments of stenosis and plaque composition. These representative cases illustrate the diversity of the vessels and demonstrate CAIN's robustness in handling individual patient variations.

Ablation analysis. To validate the effectiveness of patient-level configurations, we removed $\mathcal{F}_{cpr}$, $\mathcal{C}_{at}^{i}$ $(i \in [1, c])$, and patient-level prediction, replacing them with the branch statistical aggregation from previous research, applied to

Table 1. Quantitative comparison with SOTAs at the artery level and patient level.

Stenosis characterization at the artery-level						
Method	ACC	Prec	Recall	F1	NPV	Spec
3D-RCNN [20]	0.785 (<0.001)	0.682 (±0.058)	0.711 (±0.018)	0.685 (±0.044)	0.955 (<0.001)	0.956 (<0.001)
2D-RCNN+PT. [5]	0.797 (<0.001)	0.702 (±0.039)	0.728 (±0.019)	0.712 (±0.029)	0.957 (<0.001)	0.958 (<0.001)
Coro. RCNN [19]	0.824 (<0.001)	0.732 (±0.047)	0.761 (±0.019)	0.741 (±0.035)	0.962 (<0.001)	0.963 (<0.001)
Mesh-Infer. [16]	0.852 (<0.001)	0.780 (±0.026)	0.806 (±0.012)	0.791 (±0.019)	0.968 (<0.001)	0.969 (<0.001)
SC-Net [10]	0.872 (<0.001)	0.798 (±0.025)	0.777 (±0.039)	0.786 (±0.032)	0.974 (<0.001)	0.973 (<0.001)
CAIN (Ours)	**0.897 (<0.001)**	**0.846 (±0.014)**	**0.867 (±0.006)**	**0.855 (±0.009)**	**0.978 (<0.001)**	**0.978 (<0.001)**
Plaque characterization at the artery-level						
Method	ACC	Prec	Recall	F1	NPV	Spec
3D-RCNN [20]	0.810 (<0.001)	0.790 (±0.005)	0.803 (±0.005)	0.795 (±0.004)	0.934 (±0.001)	0.935 (<0.001)
Coro. RCNN [19]	0.844 (<0.001)	0.829 (±0.006)	0.817 (±0.006)	0.820 (±0.004)	0.948 (<0.001)	0.947 (±0.001)
Mesh-Infer. [16]	0.883 (<0.001)	0.873 (±0.002)	0.868 (±0.004)	0.870 (±0.003)	0.960 (<0.001)	0.959 (<0.001)
SC-Net [10]	0.876 (<0.001)	0.866 (±0.003)	0.866 (±0.001)	0.866 (±0.002)	0.957 (<0.001)	0.957 (<0.001)
ADI-Net (Ours)	**0.912 (<0.001)**	**0.904 (±0.001)**	**0.900 (±0.001)**	**0.902 (±0.001)**	**0.970 (<0.001)**	**0.969 (<0.001)**
CAD-RADS classification at the patient-level						
Method	ACC	Prec	Recall	F1	NPV	Spec
3D-RCNN [20]	0.821 (<0.001)	0.738 (±0.040)	0.719 (±0.063)	0.724 (±0.050)	0.971 (<0.001)	0.970 (<0.001)
2D-RCNN+PT. [5]	0.835 (<0.001)	0.804 (±0.019)	0.790 (±0.018)	0.787 (±0.013)	0.973 (<0.001)	0.973 (<0.001)
Coro. RCNN [19]	0.858 (<0.001)	0.849 (±0.014)	0.833 (±0.005)	0.835 (±0.006)	0.976 (<0.001)	0.976 (<0.001)
Mesh-Infer. [16]	0.913 (<0.001)	0.896 (±0.011)	0.898 (±0.003)	0.893 (±0.005)	0.985 (<0.001)	0.986 (<0.001)
SC-Net [10]	0.885 (<0.001)	0.852 (±0.014)	0.834 (±0.017)	0.839 (±0.013)	0.981 (<0.001)	0.981 (<0.001)
CAIN (Ours)	**0.959 (<0.001)**	**0.951 (±0.003)**	**0.951 (±0.003)**	**0.951 (±0.003)**	**0.993 (<0.001)**	**0.993 (<0.001)**

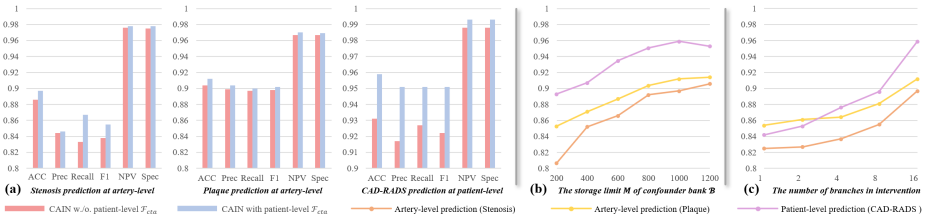

Fig. 4. Experimental results of the ablation analysis for CAIN architecture design.

the CAD-RADS level. As shown in Fig.4a, after removing patient-related configurations, there was a significant decline in CAD-RADS classification performance compared to artery-level predictions, which aligns with the trend observed in SOTA methods. This suggests that the constructed holistic causal graph, which more accurately reflects the specific conditions of individual patients, mitigates the negative impact of confounders on individual patient prediction. Next, we explored how the maximum capacity M of $\mathcal{B}$ influences CAIN performance. As shown in Fig.4b, with $M = 1000$, the intervention variable in $\mathcal{B}$ effectively captured individual patient differences while balancing memory usage. To validate the effectiveness of the patient-adaptive intervention, we controlled the number of coronary branches involved, randomly selecting 1, 2, 4, or 8 branches to obtain the intervention variables, instead of using all 16 branches that comprehensively represent the patient's coronary tree. As shown in Fig.4c, as the number of involved branches decreased, performance declined significantly, underscoring the importance of fully utilizing coronary tree information to support targeted causal interventions for individual patients.

4 Conclusion

In this work, we propose a novel causal-holistic adaptive intervention network, which, for the first time, enables personalized CAD diagnosis based on the unique conditions of each individual. Compared to methods limited to artery-level predictions, our architecture integrates patient-level holistic causal graphs, providing patient-adaptive causal interventions, leveraging dual-level semantic representation and dynamically updated intervention. Comprehensive experiments across three clinical centers demonstrate that CAIN outperforms SOTAs in dual-level predictions for individual patients, further advancing personalized CAD diagnosis.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Nos. 62372135 and 62272135), the King Abdullah University of Science and Technology (KAUST) Office of Research Administration (ORA) under Award No. REI/1/5234-01-01, REI/1/5414-01-01, REI/1/5289-01-01, REI/1/5404-01-01, REI/1/5992-01-01, URF/1/4663-01-01, Center of Excellence for Smart Health (KCSH), under award number 5932, Center of Excellence on Generative AI, under award number 5940, Heilongjiang Provincial Key Research and Development Plan 2024ZX12C23, 2023ZX02C10, 2022ZX01A30, and GA23C007, Hunan Provincial Key Research and Development Plan 2023SK2060, Jiangsu Provincial Key Research and Development Plan BE2023081, and the Natural Science Foundation of Heilongjiang Province under Grant LH2024F019.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Anwer, A.M., Karacan, H., Enver, L., Cabuk, G.: Machine learning applications for vascular stenosis detection in computed tomography angiography: a systematic review and meta-analysis. Neural Computing and Applications **36**(29), 17767–17786 (2024)
2. Cao, J.J., Shen, L., Nguyen, J., Rapelje, K., Porter, C., Shlofmitz, E., Jeremias, A., Cohen, D.J., Ali, Z.A., Shlofmitz, R.: Accuracy and limitation of plaque detection by coronary cta: a section-to-section comparison with optical coherence tomography. Scientific Reports **13**(1), 11845 (2023)
3. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers. In: European conference on computer vision. pp. 213–229. Springer (2020)
4. Cury, R.C., Leipsic, J., Abbara, S., Achenbach, S., Berman, D., Bittencourt, M., Budoff, M., Chinnaiyan, K., Choi, A.D., Ghoshhajra, B., et al.: Cad-rads™ 2.0–2022 coronary artery disease-reporting and data system: an expert consensus document of the society of cardiovascular computed tomography (scct), the american college of cardiology (acc), the american college of radiology (acr), and the north america society of cardiovascular imaging (nasci). Cardiovascular Imaging **15**(11), 1974–2001 (2022)

5. Denzinger, F., Wels, M., Ravikumar, N., Breininger, K., Reidelshöfer, A., Eckert, J., Sühling, M., Schmermund, A., Maier, A.: Coronary artery plaque characterization from ccta scans using deep learning and radiomics. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22. pp. 593–601. Springer (2019)
6. Kuettner, A., Kopp, A.F., Schroeder, S., Rieger, T., Brunn, J., Meisner, C., Heuschmid, M., Trabold, T., Burgstahler, C., Martensen, J., et al.: Diagnostic accuracy of multidetector computed tomography coronary angiography in patients with angiographically proven coronary artery disease. Journal of the American College of Cardiology **43**(5), 831–839 (2004)
7. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)
8. Ma, X., Fang, X., Zou, M., Luo, G., Wang, W., Wang, K., Qiu, Z., Gao, X., Li, S.: A trusted lesion-assessment network for interpretable diagnosis of coronary artery disease in coronary ct angiography. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 6009–6017 (2025)
9. Ma, X., Luo, G., Wang, W., Wang, K.: Transformer network for significant stenosis detection in ccta of coronary arteries. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VI 24. pp. 516–525. Springer (2021)
10. Ma, X., Zou, M., Fang, X., Liu, Y., Luo, G., Wang, W., Wang, K., Qiu, Z., Gao, X., Li, S.: Spatio-temporal contrast network for data-efficient learning of coronary artery disease in coronary ct angiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 645–655. Springer (2024)
11. Nielsen, R.V., Fuster, V., Bundgaard, H., Fuster, J.J., Johri, A.M., Kofoed, K.F., Douglas, P.S., Diederichsen, A., Shapiro, M.D., Nicholls, S.J., et al.: Personalized intervention based on early detection of atherosclerosis: Jacc state-of-the-art review. Journal of the American College of Cardiology **83**(21), 2112–2127 (2024)
12. Nowbar, A.N., Gitto, M., Howard, J.P., Francis, D.P., Al-Lamee, R.: Mortality from ischemic heart disease: Analysis of data from the world health organization and coronary artery disease risk factors from ncd risk factor collaboration. Circulation: cardiovascular quality and outcomes **12**(6), e005375 (2019)
13. Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., Savarese, S.: Generalized intersection over union: A metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 658–666 (2019)
14. Serruys, P.W., Hara, H., Garg, S., Kawashima, H., Nørgaard, B.L., Dweck, M.R., Bax, J.J., Knuuti, J., Nieman, K., Leipsic, J.A., et al.: Coronary computed tomographic angiography for complete assessment of coronary artery disease: Jacc state-of-the-art review. Journal of the American College of Cardiology **78**(7), 713–736 (2021)
15. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 20730–20740 (2022)

16. Van Herten, R.L., Hampe, N., Takx, R.A., Franssen, K.J., Wang, Y., Suchá, D., Henriques, J.P., Leiner, T., Planken, R.N., Išgum, I.: Automatic coronary artery plaque quantification and cad-rads prediction using mesh priors. IEEE transactions on medical imaging (2023)
17. Weichsel, L., Giesen, A., André, F., Renker, M., Baumann, S., Breitbart, P., Beer, M., Maurovitch-Horvat, P., Szilveszter, B., Vattay, B., et al.: Comparison of two contemporary quantitative atherosclerotic plaque assessment tools for coronary computed tomography angiography: single-center analysis and multi-center patient cohort validation. Diagnostics **14**(2), 154 (2024)
18. Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., Bengio, Y.: Show, attend and tell: Neural image caption generation with visual attention. In: International conference on machine learning. pp. 2048–2057. PMLR (2015)
19. Zhang, Y., Ma, J., Li, J.: Coronary r-cnn: vessel-wise method for coronary artery lesion detection and analysis in coronary ct angiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 207–216. Springer (2022)
20. Zreik, M., Van Hamersvelt, R.W., Wolterink, J.M., Leiner, T., Viergever, M.A., Išgum, I.: A recurrent cnn for automatic detection and classification of coronary artery plaque and stenosis in coronary ct angiography. IEEE transactions on medical imaging **38**(7), 1588–1598 (2018)