A Trusted Lesion-assessment Network for Interpretable Diagnosis of Coronary Artery Disease in Coronary CT Angiography

Xinghua Ma[1,2], Xinyan Fang[1], Mingye Zou[1], Gongning Luo[1,2*], Wei Wang[3], Kuanquan Wang[1], Zhaowen Qiu[4*], Xin Gao[2*], Shuo Li[5,6]

[1]Faculty of Computing, Harbin Institute of Technology, Harbin, China
[2]Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[3]Faculty of Computing, Harbin Institute of Technology, Shenzhen, China
[4]College of Computer and Control Engineering, Northeast Forestry University, Harbin, China
[5]Department of Computer and Data Science, Case Western Reserve University, Cleveland, United States
[6]Department of Biomedical Engineering, Case Western Reserve University, Cleveland, United States
luogongning@hit.edu.cn; qiuzw@nefu.edu.cn; xin.gao@kaust.edu.sa

Abstract

Coronary Artery Disease (CAD) poses a significant threat to cardiovascular patients worldwide, underscoring the critical importance of automated CAD diagnostic technologies in clinical practice. Previous technologies for lesion assessment in Coronary CT Angiography (CCTA) images have been insufficient in terms of interpretability, resulting in solutions that lack clinical reliability in both network architecture and prediction outcomes, even when diagnoses are accurate. To address the limitation of interpretability, we introduce the Trusted Lesion-Assessment Network (TLA-Net), which provides a clinically reliable solution for multi-view CAD diagnosis: (1) The causality-informed evidence collection constructs a causal graph for the diagnostic process and implements causal interventions, preventing confounders' interference and enhancing the transparency of the network architecture. (2) The clinically-aligned uncertainty integration hierarchically combines Dirichlet distributions from various views based on clinical priors, offering confidence coefficients for prediction outcomes that align with physicians' image analysis procedures. Experimental results on a dataset of $2,618$ lesions demonstrate that TLA-Net, supported by its interpretable methodological design, exhibits superior performance with outstanding generalization, domain adaptability, and robustness. The source code is publicly available at the following link (https://github.com/PerceptionComputingLab/CAD_Diagnosis).

Introduction

The interpretability of methodologies is crucial for computer-aided diagnosis, especially when addressing severe conditions like coronary artery disease (CAD) (Kagiyama et al. 2019; Luo et al. 2023). CAD, a prevalent cardiovascular condition, arises from arterial stenosis caused by atheromatous plaques of varying compositions. This condition can lead to life-threatening events such as myocardial infarction and heart failure (Pagliaro et al. 2020). Adhering to the Society of Cardiovascular Computed Tomography (SCCT) guidelines (Leipsic et al. 2014), arte-

rial stenoses are categorized as significant (Sig.) or non-significant (Non-sig.) based on whether the luminal narrowing $> 50\%$. Atheromatous plaques are categorized as Calcified (Calc.), Non-calcified (Non-calc.), or Mixed, based on the presence of calcium deposits. Coronary CT angiography (CCTA) is widely used for the screening, diagnosis, and prognosis of CAD (Abdelrahman et al. 2020). Automated CCTA-based diagnosis offers a more efficient and intelligent diagnostic mode for radiologists. Nevertheless, the level of interpretability significantly affects the clinical reliability of automated diagnosis, which in turn impacts its practicality in clinical settings.

Despite considerable progress in CAD automated diagnosis, researchers often overlook the significance of interpretability, which affects clinical application. Early computer-aided diagnosis for CAD relied on semi-automated technologies (Xu et al. 2012; Zuluaga et al. 2011). While these methods required substantial manual input from healthcare professionals, they inherently offered some interpretability. The rise of deep learning in Computer Vision (CV) has led to many automated CAD diagnostic methods (Singh et al. 2020). (Zreik et al. 2018) combined Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for multi-class classification of stenoses and plaques. (Denzinger et al. 2019) and (Ma et al. 2021) further advanced these C&RNN architectures, making notable progress in detecting Sig. stenosis. (Zhang, Ma, and Li 2022) analyzed Curved Planar Reformation (CPR) volumes and employed Faster R-CNN (Girshick 2015) for CAD-related object detection, while (Van Herten et al. 2023) inferred surface meshes of arteries and plaques for CAD Reporting & Data System (CAD-RADS) scoring. However, the lack of attention to interpretability during network parameter optimization hinders their suitability for real clinical settings, diminishing radiologists' trust.

The limitations in interpretability of previous technologies throughout the diagnostic process can be attributed to two main factors: network architecture and prediction outcomes. (1) Network architecture: The causality in the diagnostic process based on imaging signals can be disrupted by spurious associations caused by confounders. In the causal
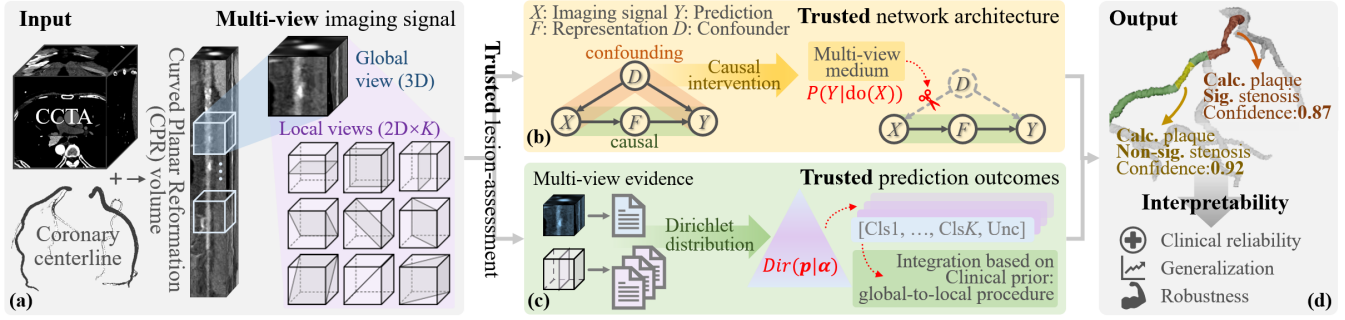

Figure 1: **Trusted CAD Diagnostic Solution.** Our interpretable methodological design: Causal interventions based on the diagnostic process within the network architecture (b) and the integration of clinical-prior-based uncertainty in prediction outcomes (c) effectively enhance the clinical reliability, generalization, and robustness (d) of automated diagnosis.

graph (Fig.1b), The imaging signal X undergoes feature extraction to obtain a representation F, which is then classified to yield the lesion prediction Y, forming a causal relationship $X \rightarrow F \rightarrow Y$. The confounder D negatively impacts both the imaging signal X and lesion prediction Y through $D \rightarrow X$ and $D \rightarrow Y$, forming a confounding relationship $X \leftarrow D \rightarrow Y$. In clinical practice, factors such as variations in imaging quality, blurred boundaries between tissues & plaques, and unobserved confounding variables can act as confounders that disrupt the causal relationship, leading to uninformed generalization by the network architecture. (2) Prediction Outcomes: Providing diagnostic results without including the confidence interval or probability distribution undermines trust and prevents the assessment of the prediction's reliability. In clinical practice, the lack of uncertainty estimation is particularly problematic in scenarios where both radiologists and patients need to rely on trust. Additionally, it can obstruct valuable feedback and make it challenging to detect and correct inaccuracies when faced with previously unseen situations. As a result, this can lead to misguided clinical decisions and potentially harmful outcomes for patients.

In this work, we introduce causal intervention and uncertainty estimation in a manner highly relevant to clinical scenarios, thereby providing a trusted network architecture and prediction outcomes for interpretable diagnosis. The trusted methodological design closely aligns the entire diagnostic process with clinical practice, not only reducing the risk of misdiagnosis and missed diagnoses through high generalization but also enhancing the domain adaptability and robustness of lesion assessment (Fig.1d). The causality perception of network architecture enhances the representation's ability to capture more general semantics, effectively addressing potential gaps between the source & target domains, such as data variations from different clinical centers, equipment models, or imaging parameters. The uncertainty estimation of prediction outcomes provides radiologists with confidence coefficients, which helps manage low-confidence results in extreme or previously unseen situations, such as unclear signals in low-quality imaging.

Specifically, we propose a Trusted Lesion-Assessment Network (TLA-Net) to analyze global and local views fo-

cused on coronary arteries in CCTA imaging for interpretable CAD diagnosis. For the network architecture, the causality-informed evidence collection aggregates multi-view semantic representations to build an intervention medium and uses the intervened expectation $P(Y|\mathrm{do}(X))$ instead of the likelihood $P(Y|X)$ to block the confounding effects on evidence prediction (Fig.1b). For the prediction outcomes, the clinical-aligned uncertainty integration estimates the uncertainty based on the evidence provided by different views using a Dirichlet distribution, and adaptively integrates it into the final lesion assessment according to the clinical prior of global-to-local procedure(Fig.1c). During this process, the representation and confounders described by multiple views complement each other, providing not only a robust medium for causal intervention but also a more comprehensive analysis for lesion prediction. Our main contributions are summarized as follows: (1) TLA-Net, as a pioneering framework for interpretable CAD diagnosis, offers critical value in advancing the automation of clinical diagnostic processes. (2) Adhering to clinical practice, the introduction of causal intervention and uncertainty estimation significantly improves the clinical reliability of automated diagnosis. (3) A thorough evaluation on a dataset of 2618 lesions demonstrates that TLA-Net outperforms State-of-the-Art (SOTA) automated CAD diagnostic methods.

Related Work

Automated CAD Diagnosis Automatic CAD diagnosis technologies are continuously being developed to address the high labor intensity and inter-observer variability associated with manual diagnosis (Gudigar et al. 2021). Traditional methods, such as thresholding, were initially used to detect plaques (Kristanto et al. 2013; Schepis et al. 2010), but these were often affected by artifacts and had limited effectiveness in detecting non-calcified plaques. Recently, deep learning-based methods have significantly improved CAD diagnosis, particularly with the use of CNNs and Transformers (Fischer et al. 2020; Ma et al. 2021; Zhai et al. 2022). These methods enhance plaque classification and stenosis detection by capturing richer spatial information and global context. Given the high correlation between plaques & stenosis, some studies (Zreik et al. 2018; Zhang,

Ma, and Li 2022; Van Herten et al. 2023; Ma et al. 2024) utilized a multi-task learning model, detecting stenosis while simultaneously classifying plaques.

Vision Causal Reasoning Causal reasoning, as an effective means to enhance the credibility of deep learning, has attracted the attention of many researchers in the CV field, such as image classification (Lopez-Paz et al. 2017; Rao et al. 2021), semantic segmentation (Zhang et al. 2020; Chen et al. 2022), and visual language (Yang et al. 2021; Niu et al. 2021) tasks. Research has demonstrated that spurious associations can severely impact model performance (Niu et al. 2021; Yang et al. 2021; Chen et al. 2022), and causality can effectively identify and eliminate these confounding effects through constructing structural causal models and adopting intervention measures. For example, (Zhang et al. 2020; Chen et al. 2022) employed backdoor adjustment methods to address the ever-elusive confounding effect in semantic segmentation tasks. (Yang et al. 2021; Zhang et al. 2022) utilized front-door adjustment for causal intervention without explicitly defining confounders. Additionally, some studies (Chen et al. 2020; Niu et al. 2021; Yue et al. 2021) have leveraged counterfactual frameworks to eliminate spurious associations in their respective tasks.

Uncertainty Estimation Uncertainty estimation is used to improve the reliability and interpretability of neural networks in deep learning. Standard neural network models are typically deterministic, providing only prediction results without any measure of prediction uncertainty. This limitation affects the model's reliability and robustness in complex scenarios (Sensoy, Kaplan, and Kandemir 2018). Researchers have developed various effective tools to quantify model uncertainty, including Bayesian methods (Gal and Ghahramani 2015; Molchanov, Ashukha, and Vetrov 2017), Monte-Carlo dropout (Gal and Ghahramani 2016; Tian et al. 2020), and ensemble methods (Durasov et al. 2021; Lakshminarayanan, Pritzel, and Blundell 2017). In visual classification tasks, the application of Dempster-Shafer evidence theory in multiple predictions can fully utilize multi-view information, thereby enhancing classification reliability (Han et al. 2020; Gao et al. 2023).

While causal reasoning and uncertainty estimation are widely used in CV tasks, their application to CAD diagnosis for improved interpretability and broader clinical use remains limited. Our method utilizes multiple CCTA views and incorporates causal intervention and uncertainty integration specifically designed for CAD diagnosis, ensuring reliable lesion assessment with enhanced interpretability.

Method

The TLA-Net, an interpretable automated CAD diagnostic solution (Fig.2), consists of two trusted modules: Causality-informed Evidence Collection (CEC) and Clinical-aligned Uncertainty Integration (CUI). Within the CPR volume, a set of L 3D cubes is systematically selected at uniform intervals. Each cube serves as input to TLA-Net with multiple views, including both global and local views (Fig.1a). Based on these multi-view image signals, the CEC collects evidence with causal intervention to control for confounders, while the CUI hierarchically integrates multiple uncertainties for lesion assessment. For each of the L cubes, TLA-Net provides predictions for lesion attributes and uncertainty.

Causality-informed Evidence Collection (CEC) for Trusted Network Architecture

The CEC performs causal interventions in semantic representation and evidence collection through a multi-view aggregated medium, achieving a trusted network architecture (Fig.2a). Based on the causal effect flow between X and F, as well as the causal and confounding relationship between F and Y, all possible outcomes of the random variable Y is defined as:

$$P\left(\boldsymbol{y}|\mathrm{do}\left(\boldsymbol{x}\right)\right) = \sum_f P\left(\boldsymbol{f}|\boldsymbol{x}\right)\sum_d P\left(\boldsymbol{y}|\boldsymbol{f},\boldsymbol{d}\right)P\left(\boldsymbol{d}\right) \quad (1)$$

Given the representation $\boldsymbol{f}$ corresponding to imaging signal $\boldsymbol{x}$ and the lesion category c, the process of evidence collection through causal intervention is expressed as $\sum_c (P(\boldsymbol{y}_c|\boldsymbol{f},\boldsymbol{d}))p(\boldsymbol{d})$. When using Softmax to predict the probabilities of different lesion categories, the above expression be transformed into $\mathbb{E}_d[\mathrm{Softmax}(\varphi_d(\boldsymbol{f},\boldsymbol{d}))]$, where φ_d denotes the operation of causal intervention. Following the Normalized Weighted Geometric Mean (NWGM), the intervened expectation is approximated as:

$$P\left(\boldsymbol{y}|\mathrm{do}\left(\boldsymbol{x}\right)\right) = \mathbb{E}_d\left[\mathrm{Softmax}\left(\varphi_d\left(\boldsymbol{f},\boldsymbol{d}\right)\right)\right]$$
$$\overset{\mathrm{NWGM}}{\approx} \mathrm{Softmax}\left(\mathbb{E}_d\left[\varphi_d\left(\boldsymbol{f},\boldsymbol{d}\right)\right]\right) \quad (2)$$

The input 3D-cube (global view) and its 2D-slices (local views) are referred to as $\boldsymbol{x}_g \in \mathbb{R}^{N\times N\times N}$ and $\boldsymbol{x}_{li} \in \mathbb{R}^{N\times N}(i \in [1,K])$, respectively. In the configuration with K views, a total of 9 2D-views are acquired, all centered on the center point of the 3D-cube. For global representation, the 3D-cube undergoes processing through the 3D Shift Transformer (Tang et al. 2022), establishing high-quality correlations of long-range image information. The view $\boldsymbol{x}_g$ is divided into $4\times 4\times 4$ sub-cubes, each of which is transformed into an embedding through a Multi-Layer Perceptron (MLP). These embeddings are then processed sequentially by Transformers with a shifted window mechanism and patch merging. For local representation, the K 2D-views are analyzed individually through 2D-CNN to comprehensively capture semantic features that clearly describe detailed information at each position. The view $\boldsymbol{x}_{li}$ is processed through a 4-layer convolutional encoder (Ronneberger, Fischer, and Brox 2015), where a 2×2 max-pooling is inserted between two layers. The feature maps are transformed into 1-dimensional embeddings through MLP layers, producing global semantics $\boldsymbol{f}_g \in \mathbb{R}^{512}$ and local semantics $\boldsymbol{f}_{li} \in \mathbb{R}^{512}$ $(i \in [1,K])$, ensuring comprehensive evidence collection.

In the causal graph, the unobservable variable $\boldsymbol{d}$ can be transformed into $\boldsymbol{x}$ through Bayesian network decomposition and conditional independence. Therefore, multiple views of $\boldsymbol{x}$-related representations are associated and aggregated to act as the intervention medium. The intervention
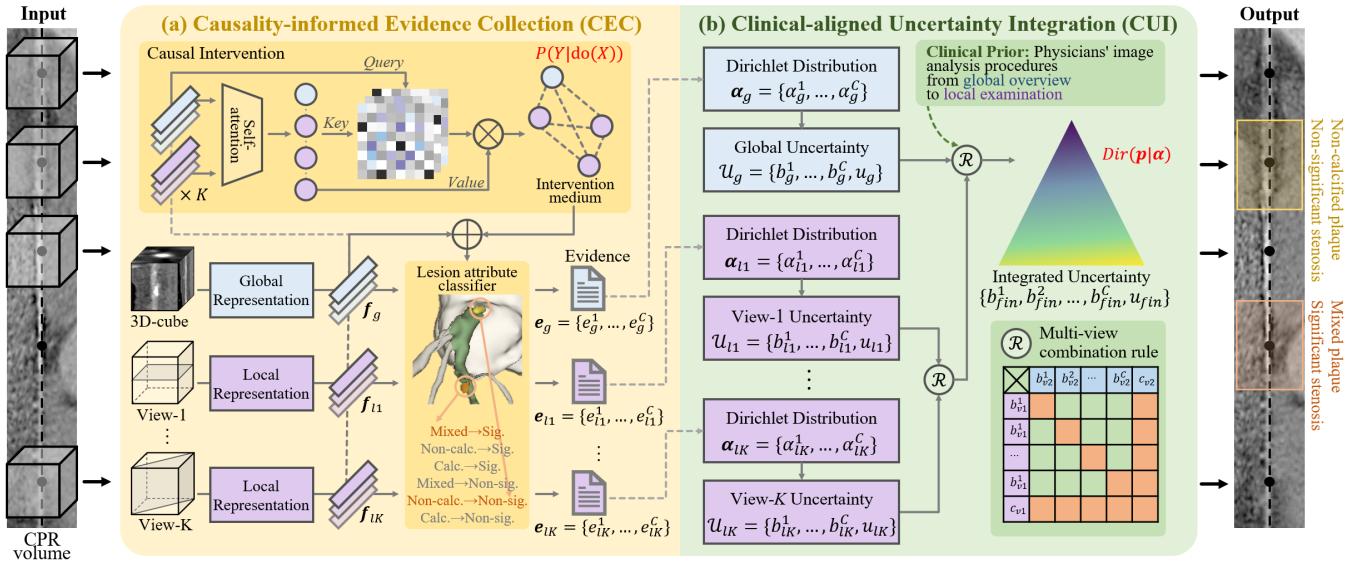
Figure 2: **Overview of TLA-Net.** The CEC relies on a multi-view medium to perform causal interventions, obtaining the expectation $P(Y|\text{do}(X))$ for trusted network architecture (a), and the CUI integrates multi-view distributions $Dir(\boldsymbol{p}|\boldsymbol{\alpha})$ align with clinical prior for trusted prediction outcomes (b).

operation of the view v is defined as:

$$\varphi_d(\boldsymbol{f}_v, \boldsymbol{d}) = \left[\text{Softmax} \left(\frac{Query_f^T Key_d}{\sqrt{dim}} \right) \cdot Value_d \right] \oplus \boldsymbol{f}_v$$

$$s.t., \hat{\boldsymbol{d}} = \text{Self\_attention} \left([\boldsymbol{f}_g, \boldsymbol{f}_{l1}, \ldots, \boldsymbol{f}_{lK}] \right)$$

(3)

where $Query_f$, Key_d and $Value_d$ are derived from $\boldsymbol{f}$ and $\hat{\boldsymbol{d}}$ through linear projections, with dim representing the dimension of representation.

Adhering to the SCCT guideline (Leipsic et al. 2014) for CAD reporting, the two stenosis degree categories (*i.e.*, Sig. & Non-Sig.) and the three plaque component categories (*i.e.*, Calc., Non-calc, & Mixed) are unified into a total of six lesion categories. The global and local semantics $\boldsymbol{f}'_g$ and $\boldsymbol{f}'_{li}$ ($i \in [1, K]$) after causal intervention, are inputted into a C-class softmax classifier with an MLP to provide evidence for lesion assessment. Here, $C = 2 \times 3 + 1$ denotes the number of classification heads, which includes six lesion categories plus an additional category indicating the absence of a lesion. Following the lesion attribute classifier, evidence for uncertainty estimation is collected, including global evidence $\boldsymbol{e}_g = \{\{e^c_g\}^C_{c=1}\}$ and local evidence $\boldsymbol{e}_{li} = \{\{e^c_{li}\}^C_{c=1}\}$ ($i \in [1, K]$).

Clinical-aligned Uncertainty Integration (CUI) for Trusted Prediction Outcomes

The CUI utilizes Bayesian-based subjective logic to integrate evidence from multiple views, in alignment with clinical prior knowledge, achieving trusted prediction outcomes (Fig.2b). Dirichlet distributions are calculated to control class probabilities based on collected global g and local

li ($i \in [1, K]$) evidence. For view v, the Dirichlet distribution (Han et al. 2020) is defined as $Dir(\boldsymbol{p}_v|\boldsymbol{\alpha}_v)$.

For the single-view v (*i.e.*, $g, l_1, \ldots, l_K$), according to the theory proposed by (Sensoy, Kaplan, and Kandemir 2018), the parameters of the Dirichlet distribution are determined as $\boldsymbol{\alpha}_v = \boldsymbol{e}_v + 1$. Subsequently, the uncertainty $\mathcal{U}_v$ for the single-view undergoes transformation based on the Dirichlet distribution:

$$\mathcal{U}_v = \left\{ \{b^c_v\}^C_{c=1}, u_v \right\}$$

$$s.t. \quad 1 = u_v + \sum^C_{c=1} b^c_v$$

(4)

where the belief mass b^c_v of category c is calculated as $(\alpha^c_v - 1)/S_v$, the Dirichlet strength S_v equals $\sum^C_{c=1} \alpha^c_v$, and overall uncertainty u_v is defined as C/S_v. Using the evidence $\boldsymbol{e}_g$ and $\boldsymbol{e}_{li}$ ($i \in [1, K]$) from the CEC, the global uncertainty $\mathcal{U}_g$ for global overview and the local uncertainty $\mathcal{U}_{li}$($i \in [1, K]$) for local examination can be derived from Eq.4, respectively.

For multi-view integration, based on the "Clinical prior — Physicians' image analysis procedures from global overview to local examination (Kallergi 2005)", the global view is integrated with K local views to hierarchically combine lesion assessments from each view. Dempster's combination rule (Jøsang and Hankin 2012) is applied to produce an overall lesion assessment with associated uncertainty, which is then transformed into the final predictive distribution:

$$\mathcal{U}_{fin} = \mathcal{R}(\mathcal{U}_g, \mathcal{U}_l)$$

$$s.t. \quad \mathcal{U}_l = \mathcal{R}(\mathcal{U}_{l1}, \mathcal{U}_{l2}, \ldots, \mathcal{U}_{lK})$$

(5)

where $\mathcal{R}(\cdot)$ denotes Dempster's combination rule. For this combination rule, given view-1 $v1$ uncertainty $\mathcal{U}_{v1} = \{\{b^c_{v1}\}^C_{c=1}, u_1\}$ and view-2 $v2$ uncertainty $\mathcal{U}_{v2} = \{\{b^c_{v2}\}^C_{c=1}, u_2\}$, the combined uncertainty $\mathcal{U} =$

$\{\{b^c\}_{c=1}^C, u\} = \mathcal{R}(\mathcal{U}_{v1}, \mathcal{U}_{v2})$ is defined as:

$$b^c = 1/(1-M) \cdot (b_{v1}^c b_{v2}^c + b_{v1}^c u_{v2} + b_{v2}^c u_{v1})$$
$$u = 1/(1-M) \cdot u_{v1} u_{v2} \tag{6}$$

where $1/(1-M)$ is the scale factor for normalization, and $M = \sum_{i \neq j} b_{v1}^i b_{v2}^j$ denotes a measure of the amount of conflict between the two mass sets. Referring to Eq.5 and Eq.6, this rule initially integrates the local uncertainties $\mathcal{U}_{li}(i \in [1, K])$ associated with each local assessment, followed by merging these integrated local uncertainties with the global uncertainty $\mathcal{U}_g$ to yield the uncertainty $\mathcal{U}_{fin}$ for the final lesion assessment, thereby accomplishing a trusted automated CAD diagnosis.

To achieve the optimization objective of TLA-Net, the loss function $\mathcal{L}_{tla}$ incorporates all sources of uncertainty, including global overview, local examination across all 2D-views, and final evaluation, allowing for end-to-end training of the solution:

$$\mathcal{L}_{tla} = \mathcal{L}_g + \sum_{i=1}^K \mathcal{L}_{li} + \mathcal{L}_{fin} \tag{7}$$

where $\mathcal{L}_g$, $\mathcal{L}_{li}$, and $\mathcal{L}_{fin}$ denote the loss for the global uncertainty, the local uncertainty, and the final uncertainty, respectively. For assessing each-view uncertainty v (*i.e.*, g, li ($i \in [1, K]$), and fin) assessment, the single-uncertainty loss $\mathcal{L}_v$ consists of an adjusted cross-entropy loss $\mathcal{L}_v^{ace}$, which ensures that more evidence is provided for the correct label compared to other categories, and a regularization term that enhances the uncertainty for misclassified samples:

$$\mathcal{L}_v = \mathcal{L}_v^{ace} + \lambda KL \left[Dir\left(\boldsymbol{p}_v | \tilde{\boldsymbol{\alpha}}_v\right) \| Dir\left(\boldsymbol{p}_v | \mathbf{1}\right) \right] \tag{8}$$

where $\lambda \in (0, 1)$ is the balancing factor, which gradually increases during training, and $\tilde{\boldsymbol{\alpha}}_v = \boldsymbol{y}_v + (1 - \boldsymbol{y}_v) \odot \boldsymbol{\alpha}_v$. The adjusted cross-entropy loss is defined as:

$$\begin{aligned}\mathcal{L}_v^{ace} &= \mathbb{E}_{\boldsymbol{p}_v \sim Dir(\boldsymbol{p}_v | \tilde{\boldsymbol{\alpha}}_v)} \left[\mathcal{L}_{CE}\left(\boldsymbol{p}_v, \boldsymbol{y}_v\right) \right] \\ &= \sum_{c=1}^C y_v^c \left(\psi(S_v) - \psi\left(\alpha_v^c\right)\right)\end{aligned} \tag{9}$$

where ψ is the digamma function and $\boldsymbol{y}_v$ is the one-hot label.

Experiment

Dataset

Experimental datasets from two Clinical Centers (CCs), each using distinct equipment models and imaging parameters, were utilized for performance evaluation: CC1: CCTA scans from 218 patients (average age: 57.4 ± 6.2 years, 163 males, 2019~2022), performed using a 320-row CT scanner. CC2: CCTA scans from 352 patients (average age: 54.7 ± 8.1 years, 274 males, 2022~2024), performed using a dual-source CT scanner. Following SCCT guidelines (Leipsic et al. 2014) for CAD reporting, five experienced physicians annotated coronary segments. Inter-observer variability was assessed through voting, and intra-observer variability was evaluated via sampling reviews. Using the marching cubes algorithm (Rajon and Bolch 2003), $3,090$ CPR volumes of main coronary branches were reconstructed, identifying $2,618$ lesions: $1,621$ Non-sig. stenoses (690 Calc. / 265 Non-calc. / 666 mixed plaques) and 997 Sig. stenoses (320 Calc. / 214 Non-calc / 463 mixed plaques).
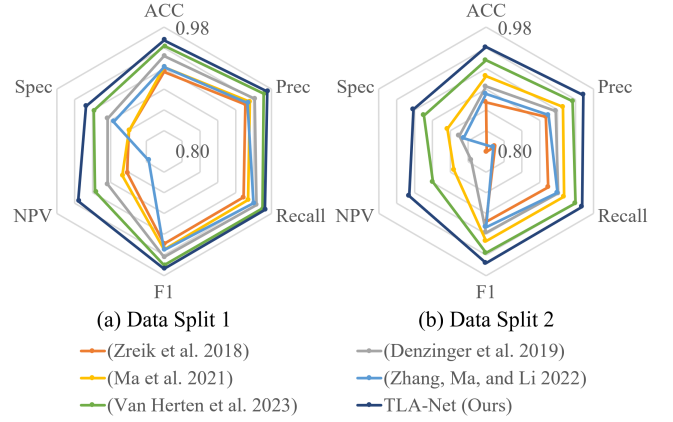

Figure 3: Radar chart for patient-level quantitative comparison with SOTA methods: Results under proportional Data Split (DS) (a) and under center-based DS (b).

Implementation Details

For TLA-Net input, 30 3D-cubes are sampled with a 5-interval sampling within CPR volume, and the size parameter N for the input 3D-cube and 2D-views is set to 64. To address potential inaccuracies in coronary centerline extraction and rotational uncertainty of MPR, the center voxel is randomly shifted by 3 voxels along any of the six neighboring directions (*i.e.*, anterior, posterior, left, right, superior, inferior) and rotated at a random angle perpendicular to the centerline. During training, the Adam optimizer is used to adjust the network parameters, with an initial learning rate set to $1e-4$, and Polynomial Decay is applied to adjust the learning rate. The balancing factor λ linearly increases from 0 to 1 during training, which reduces the early reliance on KL divergence and avoids insufficient exploration of the parameter space. After 500 epochs, the weights of the model that performed best on the validation set are saved for lesion assessment on the test set.

Evaluation Metrics

Experiments are conducted on two Data Splits (DSs) to evaluate generalization and domain adaptability. DS1: Images are divided proportionally, with 70% from two CCs used for training and the remaining 30% split evenly between validation and test sets. DS2: Images are divided by center, using images from CC1 for training and validation and images from CC2 for testing. Evaluations are performed at both the artery and patient levels. Metrics including average ACCuracy (ACC), Precision (Prec), recall, F1 score, Negative Predictive Value (NPV), and Specificity (Spec) are calculated from the confusion matrix and subjected to appropriate t-tests. At the artery-level, a lesion is considered correctly predicted if at least 50% of the lesion area is accurately identified. At the patient level, patients are classified based on the severity of the most severe stenosis or the presence of luminal stenosis. Additionally, the Area Under the receiver operating characteristic Curve (AUC) and Expected Calibration Error (ECE) are assessed to evaluate uncertainty estimation.

Method	ACC	Prec	Recall	F1	NPV	Spec
(Tejero-de Pablos et al. 2019)	0.830 (<0.001)	0.866 (±0.001)	0.866 (±0.001)	0.866 (±0.001)	0.687 (±0.018)	0.685 (±0.019)
(Zreik et al. 2018)	0.874 (<0.001)	0.907 (<0.001)	0.902 (<0.001)	0.905 (<0.001)	0.758 (±0.017)	0.748 (±0.027)
(Denzinger et al. 2019)	0.866 (<0.001)	0.896 (±0.001)	0.894 (±0.001)	0.894 (±0.001)	0.748 (±0.017)	0.749 (±0.014)
(Ma et al. 2021)	0.907 (<0.001)	0.926 (±0.001)	0.924 (±0.001)	0.925 (±0.001)	0.838 (±0.003)	0.829 (±0.008)
(Zhang, Ma, and Li 2022)	0.934 (<0.001)	0.945 (±0.001)	0.949 (<0.001)	0.947 (<0.001)	0.865 (±0.008)	0.887 (±0.001)
(Van Herten et al. 2023)	0.926 (<0.001)	0.943 (<0.001)	0.940 (<0.001)	0.941 (<0.001)	0.872 (±0.002)	0.856 (±0.009)
TLA-Net (Ours)	**0.953 (<0.001)**	**0.962 (<0.001)**	**0.963 (<0.001)**	**0.962 (<0.001)**	**0.907 (±0.004)**	**0.921 (±0.001)**
(Zreik et al. 2018)	0.809 (±0.001)	0.874 (±0.002)	0.876 (±0.001)	0.873 (<0.001)	0.537 (±0.099)	0.484 (±0.105)
(Zhang, Ma, and Li 2022)	0.849 (<0.001)	0.896 (±0.001)	0.901 (±0.001)	0.897 (<0.001)	0.642 (±0.047)	0.643 (±0.018)
(Van Herten et al. 2023)	0.875 (<0.001)	0.915 (<0.001)	0.918 (<0.001)	0.916 (<0.001)	0.683 (±0.032)	0.697 (±0.018)
TLA-Net (Ours)	**0.913 (<0.001)**	**0.939 (±0.001)**	**0.945 (<0.001)**	**0.942 (<0.001)**	**0.771 (±0.022)**	**0.806 (±0.003)**

Table 1: Artery-level quantitative comparison ($mean \pm std$) with SOTA methods under proportional data split for generalization evaluation (Top: Stenosis Degree; Bottom: Plaque Component).

Method	ACC	Prec	Recall	F1	NPV	Spec
(Tejero-de Pablos et al. 2019)	0.794 (<0.001)	0.846 (<0.001)	0.839 (±0.002)	0.842 (<0.001)	0.630 (±0.023)	0.625 (±0.048)
(Zreik et al. 2018)	0.853 (<0.001)	0.883 (±0.001)	0.885 (±0.001)	0.883 (±0.001)	0.713 (±0.032)	0.720 (±0.018)
(Denzinger et al. 2019)	0.837 (<0.001)	0.88 (<0.001)	0.873 (±0.001)	0.876 (<0.001)	0.717 (±0.009)	0.7 (±0.033)
(Ma et al. 2021)	0.878 (<0.001)	0.900 (±0.002)	0.904 (<0.001)	0.902 (±0.001)	0.770 (±0.013)	0.783 (±0.007)
(Zhang, Ma, and Li 2022)	0.905 (<0.001)	0.925 (<0.001)	0.925 (±0.001)	0.925 (<0.001)	0.834 (±0.003)	0.827 (±0.006)
(Van Herten et al. 2023)	0.912 (<0.001)	0.929 (±0.001)	0.933 (<0.001)	0.931 (<0.001)	0.822 (±0.013)	0.838 (±0.005)
TLA-Net (Ours)	**0.944 (<0.001)**	**0.956 (<0.001)**	**0.955 (<0.001)**	**0.955 (<0.001)**	**0.904 (±0.001)**	**0.895 (±0.003)**
(Zreik et al. 2018)	0.774 (<0.001)	0.848 (<0.001)	0.847 (±0.001)	0.847 (±0.001)	0.483 (±0.032)	0.482 (±0.042)
(Zhang, Ma, and Li 2022)	0.845 (±0.001)	0.893 (±0.001)	0.897 (<0.001)	0.894 (±0.001)	0.607 (±0.045)	0.618 (±0.024)
(Van Herten et al. 2023)	0.851 (<0.001)	0.896 (±0.001)	0.9 (<0.001)	0.898 (±0.001)	0.633 (±0.032)	0.654 (±0.014)
TLA-Net (Ours)	**0.909 (<0.001)**	**0.938 (<0.001)**	**0.940 (<0.001)**	**0.939 (<0.001)**	**0.762 (±0.021)**	**0.775 (±0.009)**

Table 2: Artery-level quantitative comparison ($mean \pm std$) with SOTA methods under center-based data split for domain adaptability evaluation (Top: Stenosis Degree; Bottom: Plaque Component).

Comparison with SOTAs

The quantitative comparison results demonstrate that TLA-Net consistently outperforms SOTAs, as evidenced by significant improvements in all metrics and at two DSs. In the DS1 evaluation (Tab.1 & Fig.3a), TLA-Net showed superior performance, indicating its strong generalization capability for detecting various lesion attributes. In the DS2 evaluation (Tab.2 & Fig.3b), TLA-Net's advantages were even more pronounced, highlighting its effectiveness in leveraging source domain data to accurately address target domain testing. At the artery-level, TLA-Net excelled in detecting stenosis severity and plaque composition, providing a reliable basis for identifying and characterizing coronary artery lesions. At the patient-level, its high reliability further underscores TLA-Net's potential to assist healthcare professionals in accurately assessing the severity of CAD, thereby supporting CAD-RADS timely treatment decisions. Compared to the best-performing SOTA methods (Zhang, Ma, and Li 2022; Van Herten et al. 2023), the p-values from the t-test were all less than $1e-4$, indicating that TLA-Net's prediction distribution is statistically significantly different. This confirms that the performance improvement is genuine and not due to random chance.

The qualitative comparison also shows that TLA-Net can more accurately identify areas of coronary artery lesions (Fig.4). Specifically, when dealing with lesions caused by a series of small plaques, TLA-Net can precisely detect each plaque and the degree of stenosis it causes, effectively avoiding misdiagnoses and missed diagnoses. Moreover, even with lower-quality images (Fig.4-Bottom), TLA-Net's enhanced lesion localization and qualitative analysis, facilitated by trusted methodological design, ensure that the results meet clinical needs, demonstrating its robustness.

Ablation Analysis

Effectiveness of Causal Intervention To assess the impact of the designed causal intervention on the evidence collection of the trusted network architecture in CEC, we replaced the $P(Y|\mathrm{do}(X))$ operation with self-attention (Vaswani et al. 2017) computation for ablation analysis without altering the model's parameter count (Fig.5). After removing $P(Y|\mathrm{do}(X))$ (ABL-CI), the performance declined across both DSs, highlighting that the designed causal intervention effectively prevents confounders' interference, thereby better preserving the causal relationships within the diagnostic process.

Effectiveness of Uncertainty Estimation We assessed the impact of uncertainty estimation on prediction trustworthiness by comparing CUI with previous uncertainty-related methods (Tab.3 & Fig.5) and evaluating results based solely on class probabilities (ABL-UE). Variations in performance with different uncertainty estimation approaches show that CUI's clinically aligned design is well-suited to the needs of automated diagnosis. The superior uncertainty estimation results of CUI, in terms of ACC and AUC, suggest that the diverse imaging signal from each view offers more reliable guidance for predictions when integrated with clinical pri-
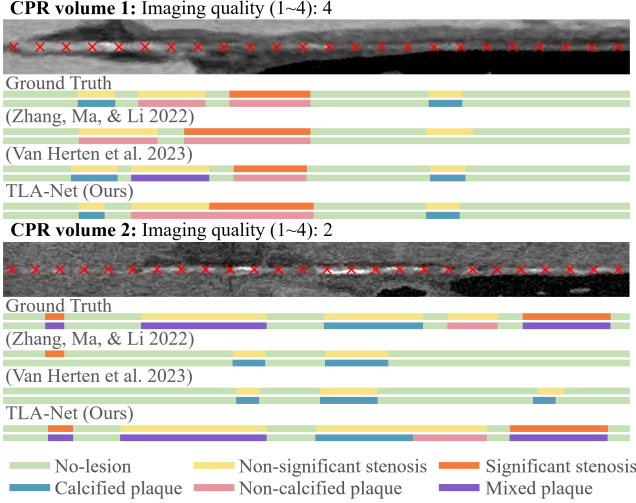
CPR volume 1: Imaging quality (1~4): 4

Ground Truth

(Zhang, Ma, & Li 2022)

(Van Herten et al. 2023)

TLA-Net (Ours)

CPR volume 2: Imaging quality (1~4): 2

Ground Truth

(Zhang, Ma, & Li 2022)

(Van Herten et al. 2023)

TLA-Net (Ours)

- No-lesion
- Non-significant stenosis
- Significant stenosis
- Calcified plaque
- Non-calcified plaque
- Mixed plaque

Figure 4: Visual examples for qualitative comparison with the best-performing SOTA methods on the coronary artery segment (The red × indicates the sampling points).
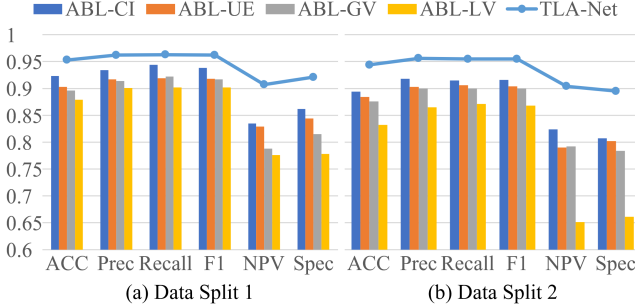

Figure 5: Artery-level ablation analysis for the effectiveness of the trusted methodological design: Results under proportional Data Split (DS) (a) and under center-based DS (b).

ors. The lower ECE indicates a more accurate correspondence between the confidence coefficient and prediction results with CUI. Additionally, the more pronounced performance drop observed when averaging classification probabilities in DS2 underscores the role of uncertainty estimation in enhancing domain adaptability.

Selection and Number of Views To understand how the selection and number of views affect TLA-Net's performance, we evaluated the network using only global or local views. Comparing the performance (Tab.3) between TLA-Net and these variants (ABL-GV & ABL-LV) revealed that both global & local semantic extraction and evidence collection contribute positively to lesion assessment. We also investigated the impact of the number of views by randomly selecting subsets from the local views (Fig.6a). The ablation results showed that satisfactory performance can be achieved with 6 to 8 selected views, with optimal performance reached at 9 views. Furthermore, a comparison of the performance of variants that retain only the three views (Fig.6b) containing the central axis of each cube (a total of

Method	ACC	AUC	ECE
Softmax	0.902 / 0.887	0.921 / 0.907	0.217 / 0.247
Dropout	0.911 / 0.889	0.938 / 0.916	0.188 / 0.201
Ensemble	0.926 / 0.906	0.935 / 0.924	0.192 / 0.198
Ours	**0.953 / 0.944**	**0.963 / 0.950**	**0.159 / 0.173**

Table 3: Artery-level quantitative comparison (Data Split (DS)1 / DS2) of various uncertainty estimation methods.
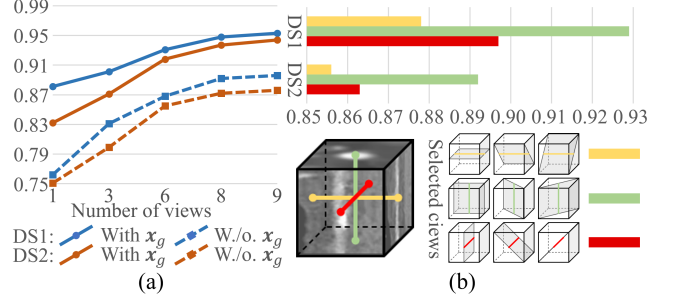

Figure 6: Combination Chart for view-related analysis: Results of various number of views (a), and results of various view selections (b).

three groups) indicates that views centered on the coronary centerline are crucial for TLA-Net's performance.

Discussion

This work enhances the interpretability of automated CAD through two main strategies. Firstly, abstracting lesion assessment into a causal graph and implementing causal interventions reduce the black-box nature of deep learning methods and control confounders during generalization. Secondly, incorporating uncertainty estimation alongside clinical priors provides confidence coefficients for lesion prediction, aiding in the identification of potential anomalies and errors in clinical practice.

These improvements in interpretability significantly enhance architectural transparency and outcome credibility, thereby increasing physician trust and acceptance while supporting personalized treatment plans. Moreover, TLA-Net's interpretability minimizes the risks of misdiagnosis and missed diagnoses while enhancing generalization, domain adaptability, and robustness. Its ability to generalize across different lesion categories and adapt to diverse clinical settings is reflected in quantitative comparisons across various centers. The solution's robustness to low-quality imaging and the effectiveness of each methodological design are also confirmed by qualitative examples and ablation studies.

Overall, TLA-Net addresses critical challenges in CAD-RADS and meets the demands of clinical practice. Future work will explore the impact of multiple modalities on comprehensive disease monitoring, aiming for multimodal joint diagnosis and further interpretability improvement.

Conclusion

In this work, we propose TLA-Net, a novel trusted CAD diagnostic solution that addresses the interpretability limi-

tations of existing technologies. The trusted network architecture, based on causal intervention, manages confounders and concentrates semantics and evidence on lesions. Additionally, the integration of multi-view uncertainty, informed by clinical prior knowledge, ensures high confidence in prediction outcomes and helps prevent missed or misdiagnoses. Experimental results demonstrate that TLA-Net outperforms SOTA methods in lesion assessment, exhibiting superior generalization. Supported by an interpretable network architecture and reliable prediction outcomes, our solution offers automated CAD diagnosis with clinical reliability, domain adaptability, and robustness.

References

Abdelrahman, K. M.; Chen, M. Y.; Dey, A. K.; Virmani, R.; Finn, A. V.; Khamis, R. Y.; Choi, A. D.; Min, J. K.; Williams, M. C.; Buckler, A. J.; et al. 2020. Coronary computed tomography angiography from clinical uses to emerging technologies: JACC state-of-the-art review. *Journal of the American College of Cardiology*, 76(10): 1226–1243.

Chen, L.; Yan, X.; Xiao, J.; Zhang, H.; Pu, S.; and Zhuang, Y. 2020. Counterfactual samples synthesizing for robust visual question answering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10800–10809.

Chen, Z.; Tian, Z.; Zhu, J.; Li, C.; and Du, S. 2022. C-cam: Causal cam for weakly supervised semantic segmentation on medical image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11676–11685.

Denzinger, F.; Wels, M.; Ravikumar, N.; Breininger, K.; Reidelshöfer, A.; Eckert, J.; Sühling, M.; Schmermund, A.; and Maier, A. 2019. 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*, 593–601. Springer.

Durasov, N.; Bagautdinov, T.; Baque, P.; and Fua, P. 2021. Masksembles for uncertainty estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13539–13548.

Fischer, A. M.; Eid, M.; De Cecco, C. N.; Gulsun, M. A.; Van Assen, M.; Nance, J. W.; Sahbaee, P.; De Santis, D.; Bauer, M. J.; Jacobs, B. E.; et al. 2020. Accuracy of an artificial intelligence deep learning algorithm implementing a recurrent neural network with long short-term memory for the automated detection of calcified plaques from coronary computed tomography angiography. *Journal of thoracic imaging*, 35: S49–S57.

Gal, Y.; and Ghahramani, Z. 2015. Bayesian convolutional neural networks with Bernoulli approximate variational inference. *arXiv preprint arXiv:1506.02158*.

Gal, Y.; and Ghahramani, Z. 2016. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, 1050–1059. PMLR.

Gao, Z.; Liu, Y.; Wu, F.; Shi, N.; Shi, Y.; and Zhuang, X. 2023. A reliable and interpretable framework of multi-view learning for liver fibrosis staging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 178–188. Springer.

Girshick, R. 2015. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, 1440–1448.

Gudigar, A.; Nayak, S.; Samanth, J.; Raghavendra, U.; AJ, A.; Barua, P. D.; Hasan, M. N.; Ciaccio, E. J.; Tan, R.-S.; and Rajendra Acharya, U. 2021. Recent trends in artificial intelligence-assisted coronary atherosclerotic plaque characterization. *International journal of environmental research and public health*, 18(19): 10003.

Han, Z.; Zhang, C.; Fu, H.; and Zhou, J. T. 2020. Trusted multi-view classification. In *International Conference on Learning Representations*.

Jøsang, A.; and Hankin, R. 2012. Interpretation and fusion of hyper opinions in subjective logic. In *2012 15th International Conference on Information Fusion*, 1225–1232. IEEE.

Kagiyama, N.; Shrestha, S.; Farjo, P. D.; and Sengupta, P. P. 2019. Artificial intelligence: practical primer for clinical research in cardiovascular disease. *Journal of the American Heart Association*, 8(17): e012788.

Kallergi, M. 2005. 12 Evaluation Strategies for Medical-Image Analysis and Processing Methodologies. *Medical image analysis methods*, 433.

Kristanto, W.; van Ooijen, P. M.; Jansen-van der Weide, M. C.; Vliegenthart, R.; and Oudkerk, M. 2013. A meta analysis and hierarchical classification of HU-based atherosclerotic plaque characterization criteria. *PLoS One*, 8(9): e73460.

Lakshminarayanan, B.; Pritzel, A.; and Blundell, C. 2017. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30.

Leipsic, J.; Abbara, S.; Achenbach, S.; Cury, R.; Earls, J. P.; Mancini, G. J.; Nieman, K.; Pontone, G.; and Raff, G. L. 2014. SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee.

Journal of cardiovascular computed tomography, 8(5): 342–358.

Lopez-Paz, D.; Nishihara, R.; Chintala, S.; Scholkopf, B.; and Bottou, L. 2017. Discovering causal signals in images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6979–6987.

Luo, G.; Ma, X.; Guo, J.; Zou, M.; Wang, W.; Cao, Y.; Wang, K.; and Li, S. 2023. Trajectory-aware Adaptive Imaging Clue Analysis for Guidewire Artifact Removal in Intravascular Optical Coherence Tomography. *IEEE Journal of Biomedical and Health Informatics*.

Ma, X.; Luo, G.; Wang, W.; and Wang, K. 2021. 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*, 516–525. Springer.

Ma, X.; Zou, M.; Fang, X.; Liu, Y.; Luo, G.; Wang, W.; Wang, K.; Qiu, Z.; Gao, X.; and Li, S. 2024. 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*, 645–655. Springer.

Molchanov, D.; Ashukha, A.; and Vetrov, D. 2017. Variational dropout sparsifies deep neural networks. In *International conference on machine learning*, 2498–2507. PMLR.

Niu, Y.; Tang, K.; Zhang, H.; Lu, Z.; Hua, X.-S.; and Wen, J.-R. 2021. Counterfactual vqa: A cause-effect look at language bias. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12700–12710.

Pagliaro, B. R.; Cannata, F.; Stefanini, G. G.; and Bolognese, L. 2020. Myocardial ischemia and coronary disease in heart failure. *Heart Failure Reviews*, 25(1): 53–65.

Rajon, D.; and Bolch, W. E. 2003. Marching cube algorithm: review and trilinear interpolation adaptation for image-based dosimetric models. *Computerized Medical Imaging and Graphics*, 27(5): 411–435.

Rao, Y.; Chen, G.; Lu, J.; and Zhou, J. 2021. Counterfactual attention learning for fine-grained visual categorization and re-identification. In *Proceedings of the IEEE/CVF international conference on computer vision*, 1025–1034.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, 234–241. Springer.

Schepis, T.; Marwan, M.; Pflederer, T.; Seltmann, M.; Ropers, D.; Daniel, W. G.; and Achenbach, S. 2010. Quantification of non-calcified coronary atherosclerotic plaques with dual-source computed tomography: comparison with intravascular ultrasound. *Heart*, 96(8): 610–615.

Sensoy, M.; Kaplan, L.; and Kandemir, M. 2018. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems*, 31.

Singh, S. P.; Wang, L.; Gupta, S.; Goli, H.; Padmanabhan, P.; and Gulyás, B. 2020. 3D deep learning on medical images: a review. *Sensors*, 20(18): 5097.

Tang, Y.; Yang, D.; Li, W.; Roth, H. R.; Landman, B.; Xu, D.; Nath, V.; and Hatamizadeh, A. 2022. 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*, 20730–20740.

Tejero-de Pablos, A.; Huang, K.; Yamane, H.; Kurose, Y.; Mukuta, Y.; Iho, J.; Tokunaga, Y.; Horie, M.; Nishizawa, K.; Hayashi, Y.; et al. 2019. Texture-based classification of significant stenosis in CCTA multi-view images of coronary arteries. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*, 732–740. Springer.

Tian, J.; Cheung, W.; Glaser, N.; Liu, Y.-C.; and Kira, Z. 2020. UNO: Uncertainty-aware noisy-or multimodal fusion for unanticipated input degradation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 5716–5723. IEEE.

Van Herten, R. L.; Hampe, N.; Takx, R. A.; Franssen, K. J.; Wang, Y.; Suchá, D.; Henriques, J. P.; Leiner, T.; Planken, R. N.; and Išgum, I. 2023. Automatic Coronary Artery Plaque Quantification and CAD-RADS Prediction using Mesh Priors. *IEEE transactions on medical imaging*.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention Is All You Need.(Nips), 2017. *arXiv preprint arXiv:1706.03762*, 10: S0140525X16001837.

Xu, Y.; Liang, G.; Hu, G.; Yang, Y.; Geng, J.; and Saha, P. K. 2012. Quantification of coronary arterial stenoses in CTA using fuzzy distance transform. *Computerized Medical Imaging and Graphics*, 36(1): 11–24.

Yang, X.; Zhang, H.; Qi, G.; and Cai, J. 2021. Causal attention for vision-language tasks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9847–9857.

Yue, Z.; Wang, T.; Sun, Q.; Hua, X.-S.; and Zhang, H. 2021. Counterfactual zero-shot and open-set visual recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 15404–15414.

Zhai, Z.; van Velzen, S. G.; Lessmann, N.; Planken, N.; Leiner, T.; and Išgum, I. 2022. Learning coronary artery calcium scoring in coronary CTA from non-contrast CT using unsupervised domain adaptation. *Frontiers in cardiovascular medicine*, 9: 981901.

Zhang, D.; Zhang, H.; Tang, J.; Hua, X.-S.; and Sun, Q. 2020. Causal intervention for weakly-supervised semantic segmentation. *Advances in Neural Information Processing Systems*, 33: 655–666.

Zhang, H.; Xiao, L.; Cao, X.; and Foroosh, H. 2022. Multiple adverse weather conditions adaptation for object detection via causal intervention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(3): 1742–1756.

Zhang, Y.; Ma, J.; and Li, J. 2022. Coronary R-CNN: Vessel-Wise Method for Coronary Artery Lesion Detection and

Analysis in Coronary CT Angiography. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part III*, 207–216. Springer.

Zreik, M.; Van Hamersvelt, R. W.; Wolterink, J. M.; Leiner, T.; Viergever, M. A.; and Išgum, I. 2018. 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.

Zuluaga, M. A.; Magnin, I. E.; Hernández Hoyos, M.; Delgado Leyton, E. J.; Lozano, F.; and Orkisz, M. 2011. Automatic detection of abnormal vascular cross-sections based on density level detection and support vector machines. *International journal of computer assisted radiology and surgery*, 6: 163–174.