Clinically Structured Chain-of-Thought Reasoning for Intelligent Guidance in Coronary Artery Disease

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

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

Abstract. Intelligent Coronary Artery Disease (CAD) guidance requires comprehensive reasoning over heterogeneous visual attributes and risk profiles aligned with the clinical diagnostic workflow. However, existing image-to-prediction methods lack systematic case-level modeling and evidence-grounded reasoning, limiting their clinical reliability. In this context, we introduce, for the first time, a Clinically Structured Chain-of-Thought (CS-CoT) framework for intelligent CAD guidance based on coronary CT angiography images. Leveraging the semantic understanding capabilities of Large Language Models (LLMs), the proposed vision-language model enables interpretable, clinically aligned, and interactive case-level diagnosis and management. Specifically, CS-CoT transforms heterogeneous anatomical, lesion, pericoronary, risk-related features, and diagnostic case prior into unified textual-space embeddings and organizes them into a structured, traceable chain-of-thought reasoning process aligned with the clinical diagnostic workflow. Extensive experiments on 2,067 clinical cases demonstrate that CS-CoT outperforms state-of-the-art LLMs, thereby providing reliable clinical decision support at the case level. The implementation details and prompt engineering are publicly available at the code-link (`https://github.com/PerceptionComputingLab/CS-CoT`).

Keywords: Coronary artery disease · Vision-language model · Chain-of-thought · Intelligent guidance

1 Introduction

Coronary Artery Disease (CAD) remains a leading cause of mortality worldwide, and its incidence continues to rise with the aging population [24]. Pathologically,

CAD involves the accumulation of atherosclerotic plaques in the coronary arteries, which leads to progressive luminal narrowing and impaired myocardial perfusion, potentially resulting in Acute Coronary Syndromes (ACS) [4]. Coronary CT Angiography (CCTA) is a noninvasive modality, that provides high-resolution imaging of coronary anatomy, stenosis severity, and plaque morphology [12]. With growing demands for precision cardiovascular care, intelligent CAD guidance frameworks have become essential for reducing inter-observer variability, promoting diagnostic standardization, and optimizing workflow efficiency [11].

Automated CAD diagnosis technologies have been increasingly reported [26], but most rely mainly on direct image-to-prediction mapping that focuses on visual feature learning, neglecting the understanding and inference of clinical contextual information, and thus offer limited clinical reliability. Recently, deep learning–based methods, often combined with visualization techniques such as Curved Planar Reformation (CPR), have achieved strong performance in lesion detection and classification. Zreik et al. [27], Denzinger et al. [6], and Tejero-de-Pablos et al. [15] applied equidistant point sampling along centerlines to perform stenosis classification at multiple levels of granularity, whereas Ma et al. [13], Zhang et al. [25], and Van et al. [21] employed causal interventions, object detection, and surface mesh inference to refine diagnostic strategies. Despite these advances, their clinical applicability remains limited due to the lack of systematic modeling and reasoning of case information and clinical workflows, as well as the absence of interactive capabilities that allow clinicians to interpret or adjust diagnostic guidance in a patient-specific context.

Current studies show that Large Language Models (LLMs) excel at semantic understanding, as demonstrated by their effectiveness in visual question answering [2], multimodal medical analysis [16], and clinical report generation [19]. However, the complexity of CCTA image analysis and the specialization of the CAD diagnostic workflow pose significant challenges to the clinically reliable deployment of LLMs in this domain. First, the heterogeneous visual representations and patient risk profiles required for patient-specific assessment are difficult to integrate in a standardized manner. Visual attributes, such as anatomical structures, lesion characteristics, and other relevant features that must be considered jointly in clinical diagnosis, are often analyzed selectively in previous vision-based models. Second, existing vision-language models are misaligned with the systematic workflow of semantic feature analysis and risk profile integration required in clinical practice. Structuring and organizing semantic features according to a clinically coherent reasoning process to generate case-specific diagnostic and management decisions conflicts with the inherently flat reasoning paradigm.

In this work, we proposed a vision-language model, named the Clinically Structured Chain-of-Thought (CS-CoT) framework, to provide intelligent CAD guidance based on CCTA images. To address the existing challenges, CS-CoT performs comprehensive case-level modeling and systematic CAD-oriented inference, achieving interpretability, clinical alignment, and interactivity. First, the multi-aspect coronary encoding module successively models anatomical structures, lesion characteristics, the pericoronary region, patient risk profiles and

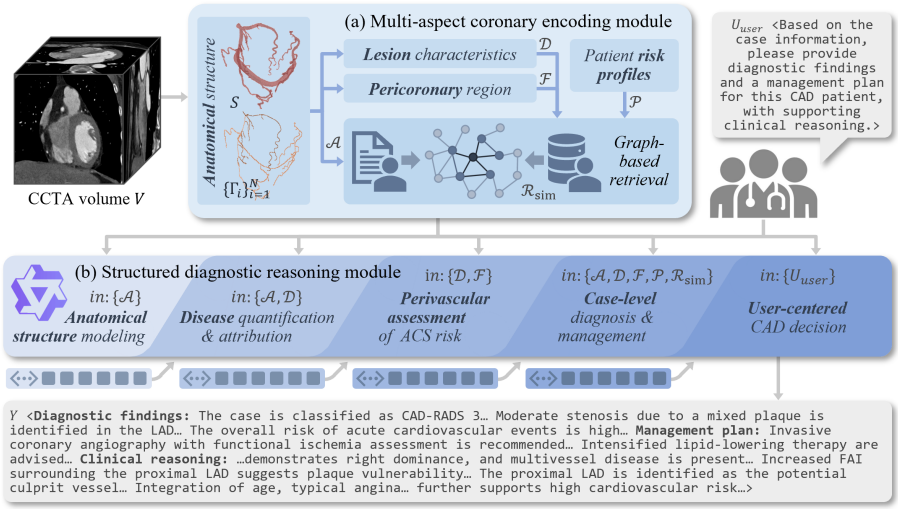

Fig. 1. CS-CoT performs multi-aspect coronary encoding (a) and structured diagnostic reasoning (b) to enable clinically reliable diagnostic and management decisions.

diagnostic case prior, and then standardizes these semantics before mapping them into the textual space. Second, the structured diagnostic reasoning module hierarchically organizes clinical features according to the diagnostic workflow, integrating them into a chain-of-thought process to achieve a comprehensive understanding of case-specific contexts and generate traceable CAD guidance. In summary, Our main contributions are as follows: (1) CS-CoT, as the first vision-language model for intelligent CAD guidance, enables interpretable, clinically aligned, and interactive case-specific diagnostic and management decisions. (2) Extensive evaluation on a dataset of 2,067 clinical cases demonstrates that our CAD-oriented framework achieves superior performance and enhanced clinical reliability over current State-Of-The-Art (SOTA) methods.

2 Method

The intelligent CAD guidance framework, CS-CoT, consists of two modules: the multi-aspect coronary encoding module and the structured diagnostic reasoning module, as shown in Fig.1. The former maps heterogeneous visual semantics and patient risk profiles into standardized embeddings in the textual space, while the latter performs structured chain-of-thought reasoning, guided by the user prompt, to generate case-specific guidance for the current diagnostic scenario. Formally, given a case consisting of a CCTA image V and risk profiles P, along with a user prompt U_{user}, CS-CoT generates guidance outputs Y conditioned on both the case information and the user focus:

$$Y = \mathcal{R}\big(\mathcal{M}(C, P), U_{\text{user}}\big) \tag{1}$$

where $\mathcal{M}$ and $\mathcal{R}$ denote the multi-aspect coronary encoding module and the structured diagnostic reasoning module, respectively.

2.1 Multi-aspect coronary encoding module

The module performs multi-aspect, case-level modeling and retrieves similar cases from the database, providing standardized diagnostic evidence to support the chain-of-thought reasoning process, as shown in Fig. 1(a). Specifically, heterogeneous clinical features are extracted from the CCTA volume across three aspects: anatomical structure, lesion characteristics, and the pericoronary region. Based on the extracted features and corresponding risk profiles, similar cases are retrieved using a graph-based approach to serve as prior diagnostic cases, offering broader clinical reference for the reasoning process.

To construct a complete representation of the coronary tree, the cascaded network architecture θ_{anat} performs coronary artery segmentation and branch labeling on the volume V, producing the coronary artery mask S and the set $\{\Gamma_i\}_{i=1}^{N}$ of branch centerlines, denoted as $S, \{\Gamma_i\}_{i=1}^{N} = \theta_{\text{anat}}(V)$. The network architecture θ_{anat} comprises a CorSegRec [17] for coronary segmentation and a CPR-GCN [23] for branch labeling, and incorporates an intelligent reconnection technique [17] to automatically complete any discontinuous segments, ensuring anatomical integrity. Next, the anatomical projection function g_{anat} maps each i-th branch to a textual-sapce embedding $\boldsymbol{a}_i$ based on anatomical features derived from the coronary mask S and its corresponding centerline Γ_i, denoted as $\boldsymbol{a}_i = g_{\text{anat}}(S, \Gamma_i)$. The anatomical features include *branch length*, *mean/maximum lumen diameter*, *start/end coordinate*, *curvature*, *parent branch name* (N/A for aortic origin), and *child branch count*. All $\boldsymbol{a}_i$ embeddings are aggregated to form the anatomical embeddings $\mathcal{A} = \{\boldsymbol{a}_i\}_{i=1}^{N}$ of the coronary tree, where each $\boldsymbol{a}_i$ consists of its corresponding *branch name* and a set of *<feature: value>* pairs. For the lesion characteristics, an artery-level architecture g_{les} reconstructs a CPR volume along each branch centerline Γ_i from the volume V and employs TR-Net [14] to localize lesions l_i and classify the stenosis severity s_i and plaque composition p_i, denoted as $\{l_i, s_i, p_i\}_{i=1}^{N} = \theta_{\text{les}}(V, \Gamma_i)$. Based on the lesion information $\{l_i, s_i, p_i\}_{i=1}^{N}$ and the anatomical context provided by S, the lesion projection function g_{les} extracts branch-level lesion features and maps them to lesion embeddings $\mathcal{D} = \{\boldsymbol{d}_i\}_{i=1}^{N}$, where $\boldsymbol{d}_i = g_{\text{les}}(l_i, s_i, p_i)$. The lesion features include *lesion count*, *minimal lumen diameter*, *lesion location*, *stenosis severity*, *plaque composition*, and *remodeling index*. For each branch, a perivascular region with a diameter comparable to the lumen is selected, fat tissue is segmented within $[-190, -30]$ HU [18], and the Fat Attenuation Index (FAI) h_i and the number of connected components c_i are computed. The perivascular projection function g_{per} then maps these features to perivascular embeddings $\mathcal{F} = \{\boldsymbol{f}_i\}_{i=1}^{N}$, where $\boldsymbol{f}_i = g_{\text{per}}(h_i, c_i)$. At the case level, the image-derived embeddings $\mathcal{A}$, $\mathcal{D}$, and $\mathcal{F}$, together with the patient risk profile P-based embeddings $\mathcal{P}$, collectively form the textual space of the current patient, denoted as $\boldsymbol{z}_{\text{case}} = [\mathcal{A}, \mathcal{D}, \mathcal{F}, \mathcal{P}]$. The patient risk profiles include age, sex, smoking status, hypertension, diabetes

mellitus, hyperlipidemia, family history, body mass index, symptom profile, and other relevant clinical risk indicators(e.g., troponin).

To enhance clinical reliability by integrating diagnostic case prior, the case graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed using BioClinicalBERT-derived similarity, where each node represents a case and edge weights are defined as $w_{ij} = \mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{z}_j)$ for $(i, j) \in \mathcal{E}$ [10]. For the diagnostic case, similarity scores are propagated through the graph via diffusion, and the top-k most relevant cases are retrieved to provide clinically and structurally analogous references for downstream reasoning:

$$\begin{aligned} &\mathcal{R}_{\mathrm{sim}} = \{\boldsymbol{r}_i\}_{i=1}^{K} = \mathrm{TopK}\left(\{s_j\}_{v_j \in \mathcal{V}}\right) \\ &\text{s.t.,} \quad s_j = \boldsymbol{z}_{\mathrm{case}}^{\top} (\mathbf{I} - \alpha \mathbf{W})^{-1} \boldsymbol{z}_j \end{aligned} \tag{2}$$

where $\mathbf{W}$ is the row-normalized adjacency matrix and $\alpha \in (0, 1)$ controls the diffusion propagation. Overall, the resulting textual space is defined as:

$$\mathcal{S} = (\mathcal{A}, \mathcal{D}, \mathcal{F}, \mathcal{P}, \mathcal{R}_{\mathrm{sim}}) \tag{3}$$

where serves as the case-level input for downstream chain-of-thought reasoning and patient-specific diagnosis and treatment guidance.

2.2 Structured diagnostic reasoning module

The module hierarchically integrates the textual space S and the user prompt U into the chain-of-thought aligned with the clinical diagnostic workflow, enabling evidence-grounded and traceable intelligent diagnosis. Built upon a Qwen-based LLM as the foundational model ϕ for inference and interaction, the proposed reasoning framework organizes the case-level clinical context into a sequence of ordered stages of understanding and cognitive reasoning, replacing blind feature aggregation with structured scheduling of clinical evidence. Specifically, ϕ first performs global modeling of coronary anatomical structure, followed by patient-level disease quantification and attribution based on branch-level characteristics. It then assesses the ACS risk based on perivascular features. Finally, ϕ integrates imaging findings, patient risk profiles, and prior diagnostic cases to generate a complete clinical decision tailored to the current diagnostic scenario.

For global anatomical understanding derived from imaging, ϕ leverages the anatomical embeddings $\mathcal{A}$, under the support of the CoT prompt U_{cot}^{1}, evaluates the overall configuration of the coronary tree, including segment count, coronary dominance pattern, tree completeness, and a prioritized analysis plan for major vessels. Next, guided by the anatomical context and the CoT prompt U_{cot}^{2}, ϕ integrates branch-level lesion descriptions $\mathcal{D}$ to derive patient-level disease conclusions. This stage determines the presence of obstructive disease, assess left main involvement, identify multivessel disease, characterize potential high-risk plaques, and localize the culprit vessel. Perivascular contextual signals are then incorporated to refine the ACS risk assessment. Integrating the CoT prompt U_{cot}^{3}, the perivascular embeddings $\mathcal{F}$, and the intermediate reasoning state, ϕ evaluates whether inflammatory markers in the perivascular region indicate increased plaque vulnerability or elevated ACS risk. Subsequently, by comparing

Table 1. Expert evaluation rubric for clinical credibility and logical consistency.

Score	Clinical Credibility ($C.C.$)	Logical Consistency ($L.C.$)
81∼100	Clinically complete outputs suitable for direct clinical reference;	Fully coherent reasoning that strictly follows the clinical reasoning workflow;
61∼80	Generally accurate outputs with only minor imprecision;	Generally coherent reasoning with minor logical jumps or weak transitions;
41∼60	Partially accurate outputs requiring clinician revision due to insufficient assessment;	Partially coherent reasoning with broken reasoning chains or unclear stage transitions.;
21∼40	Outputs with major diagnostic errors or missing critical information;	Reasoning containing major inconsistencies or disorganized logic;
0∼20	Outputs with severe diagnostic inaccuracies inconsistent with medical knowledge.	Reasoning lacking an explicit structure or exhibiting largely unstructured generation.

the current case with the retrieved similar cases, the Cot prompt U^4_{cot} guides ϕ to generate comprehensive diagnostic findings and management plan, accompanied by clinically grounded rationale. Finally, the reasoning output is further refined according to user prompt U_{user}, yielding guideline-compliant and user-centered CAD decision support. The structured reasoning process is defined as:

$$\begin{aligned} &Y = \phi(\boldsymbol{h}_4, U_{\text{user}}) \quad \text{s.t.,} \quad \boldsymbol{h}_i = \phi^{(i)}\big(\boldsymbol{h}_{i-1}, U^i_{cot}, \mathcal{X}_i\big),\ i \in [1,4] \\ &\mathcal{X} = [\{\mathcal{A}\}, \{\mathcal{A}, \mathcal{D}\}, \{\mathcal{D}, \mathcal{F}\}, \{\mathcal{A}, \mathcal{D}, \mathcal{F}, \mathcal{P}, \mathcal{R}_{\text{sim}}\}] \end{aligned} \tag{4}$$

where $\mathcal{X}$ denotes the subset of the textual space $\mathcal{S}$ required for the multiple stages of the chain-of-thought reasoning.

3 Experiment

3.1 Dataset and configuration

The experimental dataset comprised 2,067 retrospectively collected cases (age: 55.9±7.5 years; 1,383 males; 28.6% with acute chest pain) acquired from three clinical centers using five scanner models between 2020 and 2024. Diagnostic reports were prepared following the Coronary Artery Disease–Reporting and Data System (CAD-RADS) 2.0 [5], with standardized documentation of CCTA findings, patient risk profiles, prognosis, treatment strategies, and follow-up outcomes. Five experienced radiologists performed detailed annotations of the coronary tree, branch labels, and lesion characteristics on CCTA images to support visual feature learning. Seventy percent of the cases were allocated to the training set, while the remaining 30% were evenly split between the validation and test sets.

During diagnosis, disease grading followed CAD-RADS 2.0 [5], with perivascular features and their association with ACS risk referenced from [18]. For visual

Table 2. Quantitative comparison with SOTA models for intelligent CAD guidance.

Method	*Anat.*	*Attr.*	*Risk*	*Case*	*F*1	*E.S.*	*C.C.*	*L.C.*
Qwen-2-7B-Instruct [22]	0.822	0.862	0.892	0.803	0.802	0.851	84.6	86.6
Med-Alpaca [8]	0.798	0.829	0.864	0.776	0.783	0.821	80.4	84.3
Claude 3.7 Sonnet [3]	0.844	0.878	0.906	0.816	0.832	0.884	88.3	90.2
Clinical Camel [20]	0.805	0.860	0.871	0.811	0.786	0.843	85.7	85.3
OpenAI GPT-4o [1]	0.861	0.886	0.920	0.837	0.847	0.895	92.4	93.1
CS-CoT (Ours)	**0.865**	**0.902**	**0.928**	**0.859**	**0.852**	**0.916**	**94.1**	**93.5**

encoding, the networks in θ_{anat} and θ_{les} were trained per the original configuration [17] [23] [14]. The diffusion parameter α was set to 0.2, and the top $K = 5$ most similar cases were selected. For diagnostic reasoning, ϕ utilizes Qwen-2-7B-Instruct [22], fine-tuned using LoRA [9], with low-rank adapters injected into the query and value projection layers. The model was trained for 5 epochs with a batch size of 4, learning rate 2e-5, gradient accumulation over 8 steps, and mixed-precision, using a maximum sequence length of 4,096 tokens. Chain-of-thought supervision was generated by decomposing structured CAD-RADS reports into four staged reasoning trajectories. During reasoning, temperature was set to 0.2 and top-p sampling to 0.9 to ensure stable, clinically aligned outputs.

3.2 Comparison with SOTAs

To evaluate the performance of the proposed CS-CoT framework in intelligent CAD guidance, we compared it against the baseline and SOTAs, including Qwen-2-7B-Instruct [22], Med-Alpaca [8], Claude 3.7 Sonnet [3], Clinical Camel [20], and OpenAI GPT-4o [1]. All fine-tunable models were trained on the same dataset to ensure a fair comparison. As these models are inherently incapable of processing CCTA images, they were provided with the standardized textual space generated by module $\mathcal{M}$ as input to complete the diagnostic task. Following the diagnostic workflow and clinical standards, we automatically evaluated the experimental models using accuracy and average F1-score for anatomy analysis (*Anat.*), disease attribution (*Attr.*), risk assessment (*Risk*), and case-level diagnosis (*Case*). In addition, cosine similarity between generated and target reports was computed based on BERT-encoded token embeddings [7]. Five radiologists independently evaluated the intelligent guidance results for 200 cases across balanced disease severity levels using a 100-point rubric, assessing both clinical credibility (*C.C.*) and logical consistency (*L.C.*), as shown in Tab.1.

The experimental results indicate that CS-CoT demonstrates high clinical feasibility across automated metrics, embedding similarity, and expert scoring, as shown in Tab.2. In analyses related to anatomical structures, lesions, risks, and cases, CS-CoT outperformed SOTA models, demonstrating its ability to interpret image- and text-derived information across multiple aspects in accordance with patient-specific clinical contexts. The high consistency between generated and target reports in the BERT-encoded space further confirms that CS-CoT

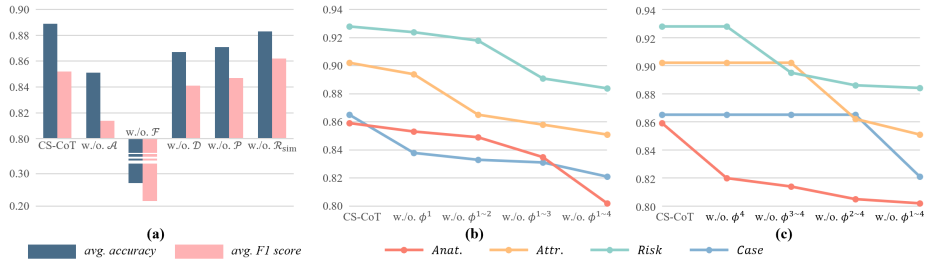

Fig. 2. Ablation study results on the effectiveness of the CS-CoT framework design.

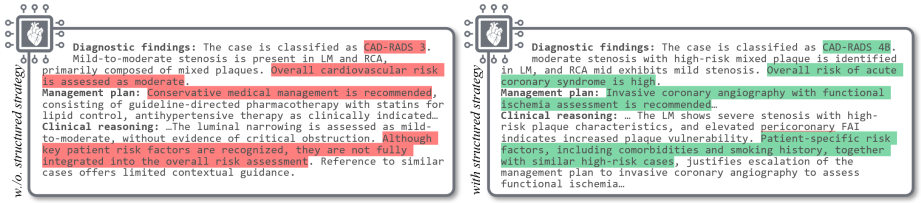

Fig. 3. Case excerpt of CS-CoT with and without clinically structured reasoning.

produces outputs with clinically appropriate organization and articulation, enhancing the precise capture of disease and risk by our model. Moreover, achieving the highest scores in clinical credibility and logical completeness demonstrates that the guidance generated by CS-CoT is more aligned with radiologists' reasoning, providing evidence-grounded and logically coherent recommendations during interactive usage.

3.3 Ablation study

To validate the effectiveness of module $\mathcal{M}$, we conducted ablation studies by sequentially removing individual embeddings from $\mathcal{S}$ during the reasoning process. When the lesion embeddings $\mathcal{D}$ were removed, the model failed to provide effective diagnostic guidance, whereas removing other embeddings led to varying degrees of performance degradation, as shown in Fig.2a. This indicates that lesion characteristics play a decisive role in CAD guidance, while multi-aspect analysis helps provide a more comprehensive perspective for accurate clinical judgment. To evaluate the effectiveness of module $\mathcal{R}$, we disabled the chain-of-thought reasoning across different forward and backward intervals. Compared with forward-interval disabling (Fig.2b), backward-interval disabling (Fig.2c) impacts case-level diagnosis earlier, even with fewer steps disabled. Across both disabling patterns, the disabling of any specific stage immediately causes a significant drop in the corresponding performance metrics. Compared with disabling forward reasoning, suppressing reverse reasoning resulted in a more pronounced decline in diagnostic performance. These findings reveal that the integration of complete case information is crucial for case-level diagnosis, and that each step

in the structured reasoning process supports the robustness of the corresponding CAD feature analysis. Furthermore, as shown in Fig.3, risk-related features with fewer tokens may be potentially overlooked in the absence of clinically structured reasoning, which can lead to errors in clinical decision-making.

4 Conclusion

This work proposes the first clinically structured vision–language model, CS-CoT, for intelligent CAD guidance. The model comprehensively analyzes anatomical structures, lesion attributions, and related risk features through a chain-of-thought reasoning process aligned with the clinical workflow. Extensive experiments demonstrate that CS-CoT achieves strong performance while maintaining interpretability, clinical alignment, and interactivity.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Nos. 82527807, 62372135, and 62501195), and by the King Abdullah University of Science and Technology (KAUST) Office of Research Administration (ORA) under Award Nos. RGC/3/6655-01-01, RGC/3/6208-01-01, RGC/3/5679-01-01, REI/1/5289-01-01, REI/1/5992-01-01, URF/1/6713-01-01, FCC/1/5932-12-11, URF/1/6599-01-02, RGC/3/6295-01-01, and URF/1/6993-01-01, as well as the Center of Excellence for Smart Health (KCSH; Award No. 5932) and the Center of Excellence on Generative AI (Award No. 5940).

Disclosure of Interests. The authors declare no competing interests.

References

1. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: Gpt-4 technical report. arXiv preprint arXiv:2303.08774 (2023)
2. Alayrac, J.B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., Lenc, K., Mensch, A., Millican, K., Reynolds, M., et al.: Flamingo: a visual language model for few-shot learning. Advances in neural information processing systems **35**, 23716–23736 (2022)
3. Anthropic: Claude 3 model card. Tech. rep., Anthropic (2024), `https://www.anthropic.com/news/claude-3-family`
4. Björkegren, J.L., Lusis, A.J.: Atherosclerosis: recent developments. Cell **185**(10), 1630–1645 (2022)
5. 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)

6. Denzinger, F., Wels, M., Ravikumar, N., Breininger, K., et al.: 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)
7. Gardazi, N.M., Daud, A., Malik, M.K., Bukhari, A., Alsahfi, T., Alshemaimri, B.: Bert applications in natural language processing: a review. Artificial Intelligence Review **58**(6), 166 (2025)
8. Han, T., Adams, L.C., Papaioannou, J.M., Grundmann, P., Oberhauser, T., Löser, A., Truhn, D., Bressem, K.K.: Medalpaca–an open-source collection of medical conversational ai models and training data. arXiv preprint arXiv:2304.08247 (2023)
9. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al.: Lora: Low-rank adaptation of large language models. Iclr **1**(2), 3 (2022)
10. Karpukhin, V., Oguz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., Yih, W.t.: Dense passage retrieval for open-domain question answering. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP). pp. 6769–6781 (2020)
11. Krittanawong, C., Johnson, K.W., Rosenson, R.S., Wang, Z., Aydar, M., Baber, U., Min, J.K., Tang, W.W., Halperin, J.L., Narayan, S.M.: Deep learning for cardiovascular medicine: a practical primer. European heart journal **40**(25), 2058–2073 (2019)
12. Kyriakoulis, I., Kumar, S.S., Palaiodimos, L., Kokkinidis, D.G.: Non-invasive imaging modalities for the evaluation of coronary artery disease among patients with high body mass index: A focus on coronary computed tomography angiography (ccta). The American Journal of Cardiology (2026)
13. Ma, X., Fang, X., Luo, G., Qiu, X., Liu, X., Huang, C., Wang, K., Qiu, Z., Zhang, T., Li, Y., et al.: Interference-free causality learning promotes cross-level, fine-grained diagnosis of coronary artery disease in coronary ct angiography. IEEE Transactions on Medical Imaging (2025)
14. Ma, X., Luo, G., Wang, W., Wang, K.: Transformer network for significant stenosis detection in ccta of coronary arteries. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 516–525. Springer (2021)
15. Tejero-de Pablos, A., Huang, K., Yamane, H., Kurose, Y., et al.: 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. pp. 732–740. Springer (2019)
16. Panagoulias, D.P., Palamidas, A., Virvou, M., Tsihrintzis, G.A.: A framework for evaluation and requirement extraction for fine-tuning of large language models in multimodal medical diagnosis. Knowledge-Based Systems **326**, 113975 (2025)
17. Qiu, Y., Shan, D., Wang, Y., Dong, P., Wu, D., Yang, X., Hong, Q., Shen, D.: A topology-preserving three-stage framework for fully-connected coronary artery extraction. Medical Image Analysis **103**, 103578 (2025)
18. Sagris, M., Antonopoulos, A.S., Simantiris, S., Oikonomou, E., Siasos, G., Tsioufis, K., Tousoulis, D.: Pericoronary fat attenuation index—a new imaging biomarker and its diagnostic and prognostic utility: a systematic review and meta-analysis (2022)
19. Tanno, R., Barrett, D.G., Sellergren, A., Ghaisas, S., Dathathri, S., See, A., Welbl, J., Lau, C., Tu, T., Azizi, S., et al.: Collaboration between clinicians and vision–language models in radiology report generation. Nature Medicine **31**(2), 599–608 (2025)

20. Toma, A., Lawler, P.R., Ba, J., Krishnan, R.G., Rubin, B.B., Wang, B.: Clinical camel: An open expert-level medical language model with dialogue-based knowledge encoding. arXiv preprint arXiv:2305.12031 (2023)
21. 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)
22. Yang, A., Yang, B., Hui, B., Zheng, B., Yu, B., Zhou, C., Li, C., Li, C., Liu, D., Huang, F., et al.: Qwen2 technical report. arXiv preprint arXiv:2407.10671 (2024)
23. 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. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 3803–3811 (2020)
24. Yang, L., Wang, Z., Zhao, S., Liu, M., Zhu, Y., Hu, F., Gao, X., Wu, Y.: Effectiveness of digital healthcare to improve clinical outcomes in discharged patients with coronary artery disease. NPJ Digital Medicine **8**(1), 473 (2025)
25. 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)
26. Zreik, M., Van Hamersvelt, R.W., Khalili, N., Wolterink, J.M., Voskuil, M., Viergever, M.A., Leiner, T., Išgum, I.: Deep learning analysis of coronary arteries in cardiac ct angiography for detection of patients requiring invasive coronary angiography. IEEE transactions on medical imaging **39**(5), 1545–1557 (2019)
27. 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)