Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sensitivity-oriented fusion of resting ECGs and structured CCTA reports for identifying angiography-defined severe right coronary artery stenosis: a single-centre retrospective study

Yue He
Nanjing University of Posts and Telecommunications
Qiuyang Yang
Nanjing University of Posts and Telecommunications
Yong Li
Jiangsu Province Hospital
Shenqi Jing
Jiangsu Province Hospital
Yun Liu
liuyun@njmu.edu.cn
Jiangsu Province Hospital
Tao Shan
Jiangsu Province Hospital

Research Article

Keywords: Coronary computed tomography angiography, 12-lead electrocardiogram, Multimodal deep learning, Severe right coronary artery stenosis, Sensitivity-oriented screening, Clinical decision support

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10387984/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Sensitivity-oriented fusion of resting ECGs and structured CCTA reports for identifying angiography-defined severe right coronary artery stenosis: a single-centre retrospective study

Yue He[1], Qiuyang Yang[1], Yong Li[1,2] , Shenqi Jing[1,2], Yun Liu[1,2*] and Tao Shan[1,2*]

1 School of Communications and Information Engineering, Nanjing University of Post and Telecommunications, Nanjing 210023, P. R. China

2 The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, P. R. China

* presents the co-response authors

Corresponding Authors

Yun Liu, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, P. R. China. E-mail: liuyun@njmu.edu.cn

Tao Shan, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, P. R. China. E-mail: nanjingtaotao@126.com

Abstract

Background: Failure to identify severe right coronary artery (RCA) stenosis may delay invasive assessment and treatment planning. Structured coronary computed tomography angiography (CCTA) reports provide anatomical screening information, but report-derived stenosis summaries do not capture electrophysiological consequences. We fused resting 12-lead electrocardiograms (ECGs) with structured CCTA report features in a multimodal model. We then evaluated sensitivity-oriented identification of RCA diameter stenosis of at least 70% in coronary angiography (CAG) reports.

Results: This single-centre retrospective study included 1,904 ECG records from 760 patients who underwent ECG, CCTA, and CAG. CAG reports provided the reference labels. The model combined a one-dimensional residual network, axial multi-head self-attention, lead attention, and feature-wise linear modulation by report-derived stenosis features from four coronary arteries. Performance was estimated using internal patient-grouped fivefold cross-validation.

At the patient level, the multimodal model achieved sensitivity of 0.849±0.037 and specificity of 0.690±0.037. Its area under the receiver operating characteristic curve (AUC) was 0.831±0.020, with accuracy of 0.764±0.028 and F1 score of 0.761±0.051. The structured-report model achieved an AUC of 0.794, sensitivity of 0.391, and specificity of 0.837. Thus, multimodal fusion increased sensitivity while reducing specificity. The ECG-only model achieved an AUC of 0.574. Ablation analyses associated CCTA fusion, temporal-decay weighting, and attention modules with internal performance. An ancillary post hoc analysis reclassified 38 of 66 CCTA false-positive patients as negative, but it did not establish clinical net benefit.

Conclusions: In this selected single-centre cohort, ECG and structured CCTA report fusion produced a high-sensitivity operating profile for angiography-defined severe RCA stenosis, with a corresponding reduction in specificity. The model may warrant evaluation as a sensitivity-oriented secondary screening aid, but it cannot yet be used to rule out disease or claim prevention of missed diagnoses. Independent multicentre validation, prespecified threshold locking, calibration, and prospective clinical-utility assessment are required before implementation.

Keywords

Coronary computed tomography angiography; 12-lead electrocardiogram; Multimodal deep learning; Severe right coronary artery stenosis; Sensitivity-oriented screening; Clinical decision support

Background

Atherosclerotic coronary artery disease remains a major cause of death and disability worldwide[1]. Progressive plaque formation and calcification may narrow the coronary lumen and impair myocardial perfusion[2]. Reliable identification of severe coronary stenosis is therefore a central task in cardiovascular imaging and clinical triage[3].

The right coronary artery (RCA) supplies the right ventricle and the inferior wall of the left ventricle. It also frequently supplies the sinoatrial and atrioventricular nodes. Severe RCA stenosis may therefore be associated with ischaemia, conduction abnormalities, and arrhythmia[4]. Missing such disease can delay definitive anatomical assessment and treatment planning.

Invasive coronary angiography (CAG) remains an important reference method for evaluating coronary luminal stenosis[5]. Coronary computed tomography angiography (CCTA) is widely used as a non-invasive test because CAG entails procedural risk, cost, and patient burden[6]. Clinical decision support should preserve the high case-finding value of this non-invasive pathway while identifying patients who may require further assessment.

CCTA interpretation also has recognised limitations. In heavily calcified vessels, blooming artefacts may obscure the luminal boundary and overestimate stenosis[7]. A small set of report-derived stenosis values is only a coarse anatomical summary. It should not be equated with the information available from full CCTA images.

A conventional resting electrocardiogram (ECG) has limited sensitivity for stable coronary disease[8]. Deep-learning models can nevertheless detect subtle waveform information that may not be apparent to human readers[9]. AI-ECG studies across cardiovascular disorders[10,11] have also suggested value for identifying obstructive coronary disease and related ischaemic risk[12].

ECG-only models remain vulnerable to interindividual variation, non-specific abnormalities, and noise[13]. Multimodal learning can combine complementary functional and anatomical information[14] and is increasingly being investigated in cardiovascular applications[15]. For a screening-oriented task, the principal question is whether this fusion improves sensitivity without concealing the associated false-positive burden.

Structured CCTA reports describe lesion location and stenosis severity, whereas ECGs provide information on myocardial electrical activity[16]. Their fusion could support case finding when either modality alone is incomplete. However, the value of this strategy for sensitivity-oriented identification of severe RCA stenosis remains uncertain.

We developed a multimodal model in which structured CCTA report features condition an ECG backbone. The target was CAG-report-defined RCA stenosis of at least 70%. We compared the fusion model with ECG-only and structured-report models, focusing on sensitivity and its specificity trade-off. Ablation, post hoc decision simulation, and interpretability analyses were used as supporting evaluations.

Results

Sensitivity-oriented patient-level performance

We compared three input configurations to assess whether multimodal fusion improved case finding: ECG-only, structured-report only, and ECG plus structured CCTA report features. All configurations used the same patient-grouped cross-validation partitions. Tables 1 and 2 report the primary patient-level results, with sensitivity interpreted together with specificity.

The multimodal model achieved a mean patient-level sensitivity of 0.849 ± 0.037 and AUC of 0.831 ± 0.020. Its accuracy was 0.764 ± 0.028, and its F1 score was 0.761 ± 0.051. These values describe internal cross-validation performance in the selected cohort.

The structured-report model achieved an AUC of 0.794, sensitivity of 0.391, and specificity of 0.837. Its inputs were four report-derived stenosis values rather than original CCTA images. This comparator therefore does not represent the diagnostic performance of a complete clinical CCTA

examination.

The ECG-only model achieved an AUC of 0.574 and sensitivity of 0.296. Both values were lower than those of the structured-report and multimodal models under the same data partitions.

Compared with the structured-report model, multimodal fusion increased sensitivity from 0.391 to 0.849. Specificity decreased from 0.837 to 0.690, while AUC increased from 0.794 to 0.831. The intended sensitivity gain therefore carried a measurable false-positive trade-off.

Table 1 Patient-level performance of the three input configurations

Model	AUC	F1	Sensitivity	Specificity	Accuracy
ECG+CCTA report	0.831	0.761	0.849	0.690	0.764
CCTA report-only	0.794	0.369	0.391	0.837	0.629
ECG-only	0.574	0.354	0.296	0.779	0.575

Table 2 Patient-level fivefold validation results for the multimodal model

Fold	AUC	Accuracy	Sensitivity	Specificity	F1 score	Brier score
1	0.830	0.744	0.852	0.651	0.756	0.169
2	0.861	0.799	0.862	0.734	0.814	0.152
3	0.834	0.797	0.902	0.693	0.816	0.160
4	0.798	0.735	0.842	0.646	0.742	0.195
5	0.832	0.746	0.787	0.726	0.677	0.182
Mean	0.831	0.764	0.849	0.690	0.761	0.172
Standard deviation	0.020	0.028	0.037	0.037	0.051	0.015

Discrimination and precision – recall performance

Figure 1 presents patient-level receiver operating characteristic (ROC) and precision – recall (PR) curves for the multimodal model across the five validation folds. Curves were calculated after patient-level probability aggregation. ROC curves display the true-positive and false-positive rates, whereas PR curves display precision and recall. Curves were interpolated to a common grid before calculating fold means, and shaded regions show the standard deviation.

The PR curve further characterised sensitivity-oriented performance under class imbalance. The multimodal model achieved a mean average precision of 0.764±0.054.

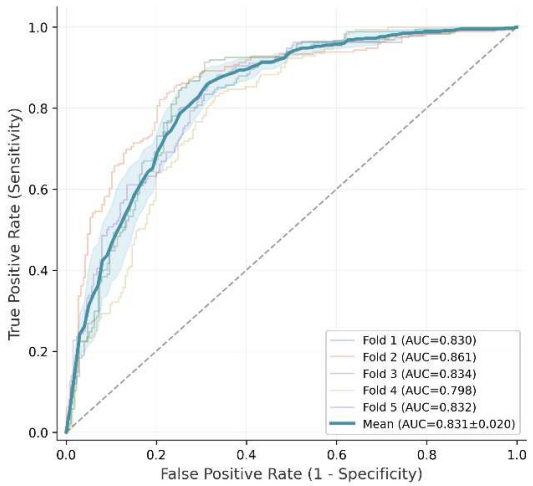


Fig. 1(a) Patient-level receiver operating characteristic curves across validation folds

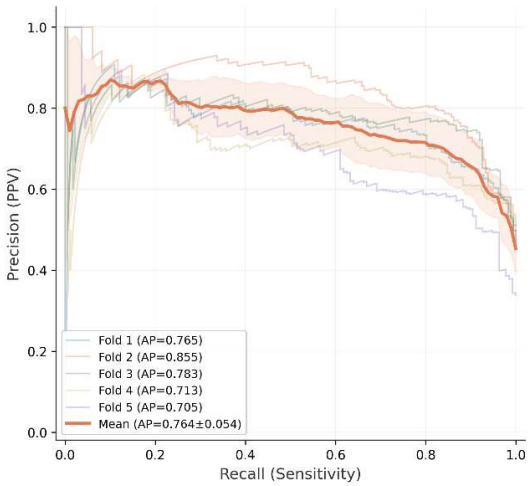


Fig. 1(b) Patient-level precision – recall curves across validation folds

Exploratory ECG-record-level threshold analysis

As a separate post hoc analysis, ECG-record-level predictions from the five validation folds were pooled. A grid search selected a threshold of 0.4 by maximising F1 score on these same predictions. The resulting confusion matrix is exploratory because threshold selection and evaluation used the same pooled validation data.

Figure 2 therefore illustrates threshold-dependent record-level classification only. It does not replace the prespecified patient-level analysis or provide an independently validated screening

threshold.

At the post hoc threshold, the 1,904 ECG records comprised 702 true positives, 765 true negatives, 157 false negatives, and 280 false positives.

The negative predictive value of this ECG-record-level confusion matrix was 0.830. This value should not be interpreted as patient-level rule-out performance.

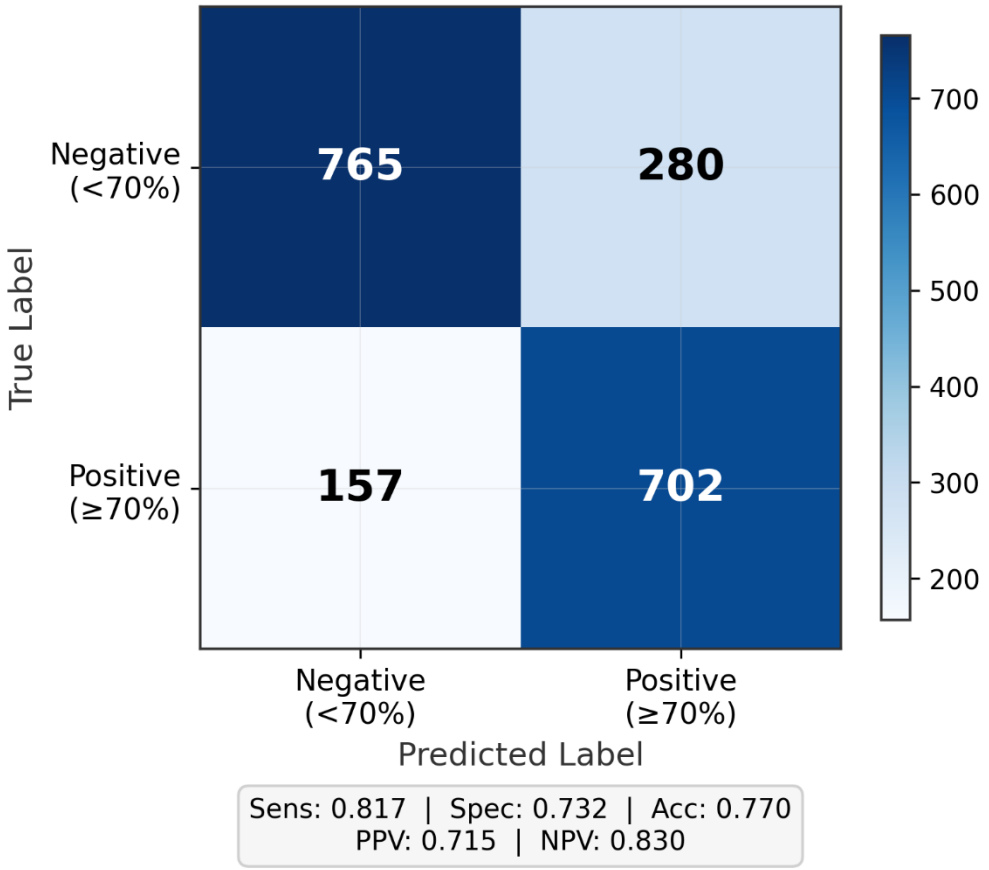


Fig. 2 Exploratory pooled ECG-record-level confusion matrix

Ancillary analysis in CCTA-positive patients

Because calcification can produce false-positive CCTA findings[7], we performed an ancillary analysis among patients whose CCTA reports indicated RCA stenosis of at least 70% (Fig. 3). This analysis was not the basis of the sensitivity-oriented primary claim.

Among 254 CCTA-positive patients, CAG reports confirmed severe RCA stenosis in 188 (74.0%) and placed 66 (26.0%) below the 70% threshold.

The multimodal model reclassified 38 of the 66 CCTA false-positive patients as negative. These patients represented 15.0% (38/254) of the CCTA-positive subgroup.

Calcium-related artefacts can cause CCTA to overestimate anatomical stenosis[7,17]. The observed reclassification is compatible with improved specificity in some CCTA-positive cases. However, true-positive losses were not fully characterised in this subgroup, and the analysis does not establish reduced CAG use or clinical net benefit.

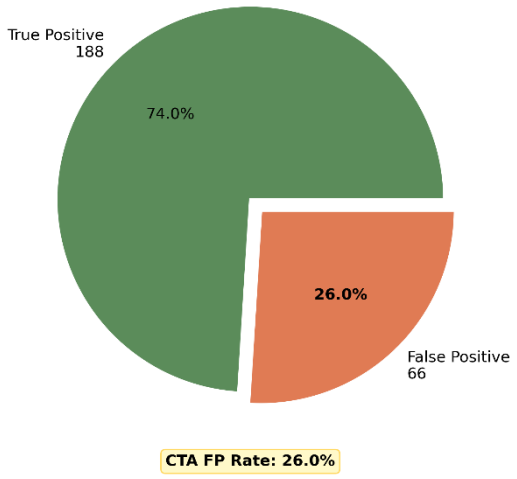


Fig. 3(a) CAG-defined stenosis in the CCTA-positive subgroup

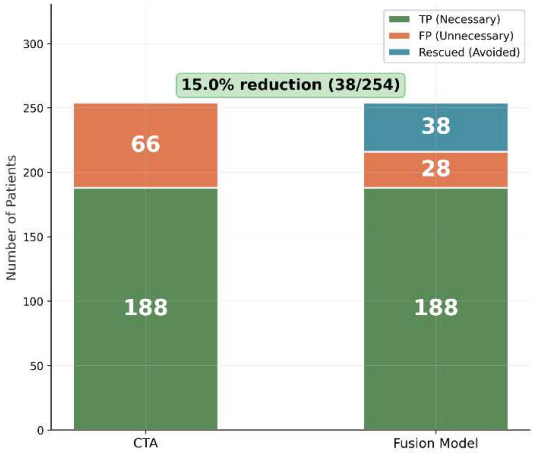


Fig. 3(b) Exploratory multimodal reclassification in the CCTA-positive subgroup

Ablation analyses of sensitivity-oriented model components

Ablation analyses examined the contributions of structured CCTA report fusion, temporal-decay weighting, axial attention, and lead attention (Fig. 4).

Removing CCTA fusion produced the ECG-only configuration, whose AUC was below 0.60 and whose F1 score was lower. Structured CCTA report features therefore contributed to internal performance in this task.

Removing temporal-decay weighting reduced sensitivity to approximately 0.38, near the structured-report model value. This association does not determine whether the reduction arose from

temporal physiological mismatch, the distribution of sample weights, or other optimisation effects.

ECG manifestations associated with myocardial ischaemia may vary with disease progression, collateral circulation, and treatment[18]. The temporal-decay result is consistent with a time-dependent contribution, but this interpretation requires independent validation.

Removing either axial attention or lead attention also reduced internal performance, supporting their inclusion within the evaluated model configuration.

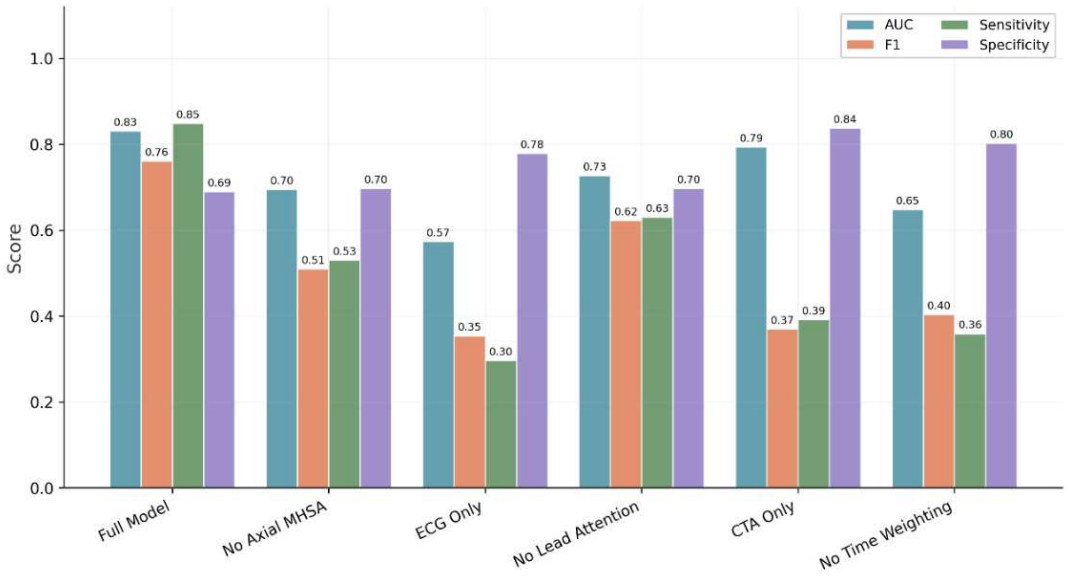


Fig. 4(a) Performance comparison in ablation analyses

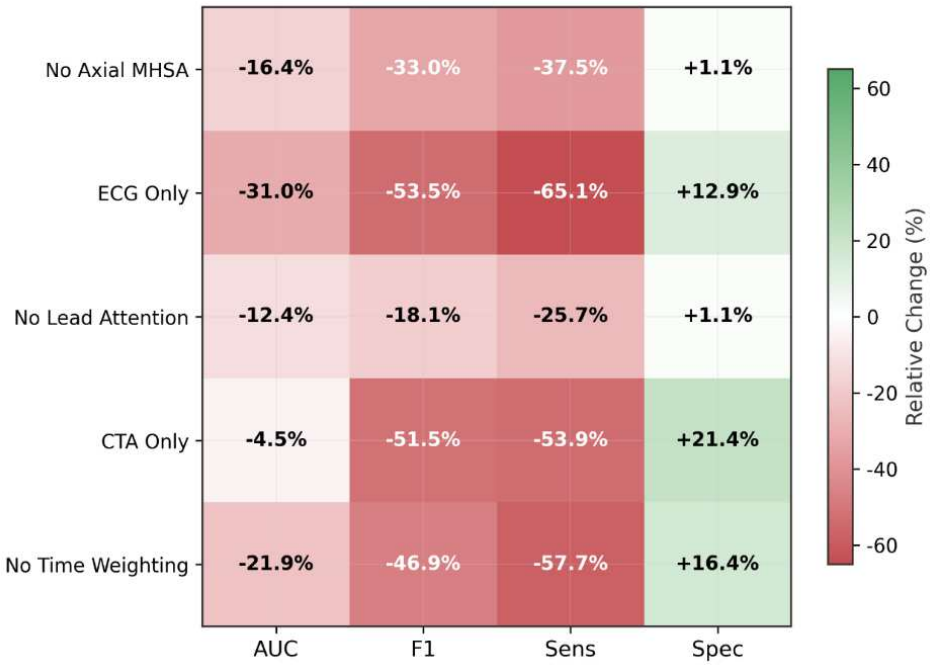


Fig. 4(b) Overview of ablation metrics

Model-focused interpretability analyses

Gradient-weighted class activation mapping (Grad-CAM)[19] and lead-attention weights were used to describe ECG regions that influenced model outputs (Figs. 5 and 6).

Lead-attention analysis assigned relatively high weights to leads I, II, and aVR. These weights describe model allocation only and do not establish the clinical or causal importance of individual leads.

Grad-CAM activation occurred mainly near the terminal QRS complex and subsequent ST–T segment, which overlap regions used in some clinical assessments of ischaemia[20]. Activation maps depend on the model, sample, and visualisation method. They neither establish an RCA-specific mechanism nor replace clinical interpretation.

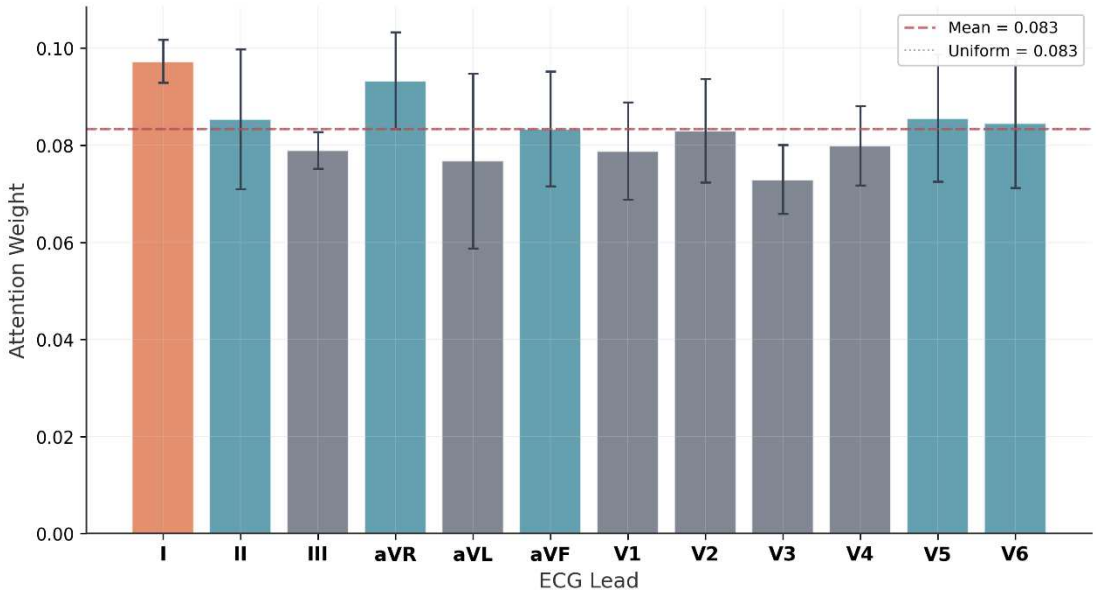


Fig. 5 Distribution of lead-attention weights

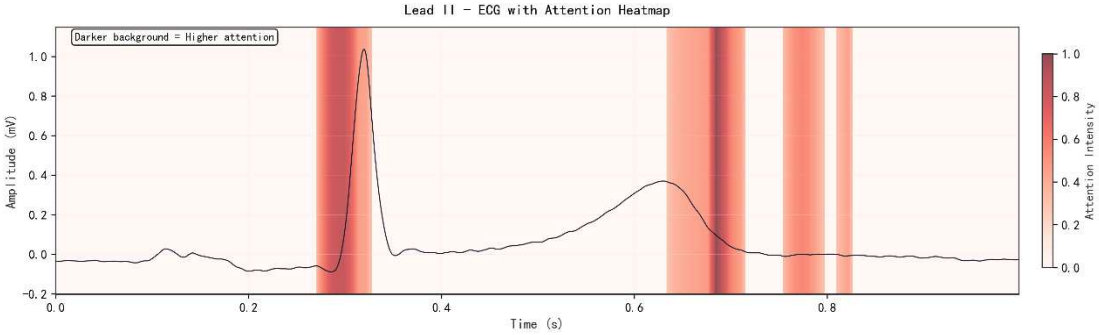


Fig. 6 Grad-CAM visualization of an ECG

Discussion

This study evaluated whether resting ECG waveforms could complement structured CCTA report features when sensitivity is prioritised for severe RCA stenosis. In internal patient-grouped cross-validation, multimodal fusion increased sensitivity from 0.391 to 0.849 relative to the structured-report model. Specificity decreased from 0.837 to 0.690. The central result is therefore a sensitivity-specificity trade-off, not a uniform improvement across all diagnostic objectives.

The intended future role is a secondary screening or triage aid within a clinically referred population. Every patient in this cohort had ECG, CCTA, and CAG, which created a selected, relatively high-risk sample. The results cannot be extrapolated to asymptomatic screening, primary

care, or patients who never undergo invasive angiography. The structured-report comparator also used four extracted stenosis values and should not be equated with full image-based CCTA interpretation.

The fusion architecture offers a plausible way to combine coarse anatomical context with electrical information. FiLM allowed CCTA report features to condition the ECG representation instead of entering only at the classifier. However, the incremental performance does not prove that the ECG branch captured RCA-specific ischaemia. Correlated clinical information, optimisation choices, and referral patterns remain alternative explanations.

The CCTA-positive subgroup analysis was ancillary to the sensitivity-oriented primary result. Reclassification of 38 false-positive patients suggests that fusion may also improve classification for some overestimated CCTA findings. The analysis did not fully quantify lost true positives within the same decision pathway. It also lacked a prespecified threshold, decision-curve analysis, and prospective outcomes. It therefore cannot support claims of fewer invasive angiograms or improved net benefit.

High sensitivity does not make the model a rule-out test. A mean sensitivity of 0.849 indicates that some angiography-defined severe cases remained undetected at the reported operating point. In addition, the 0.5 patient-level threshold was not locked in an independent cohort to meet a prespecified sensitivity target. Future studies should select the operating threshold according to an explicit clinical cost function, then evaluate calibration, false-negative consequences, and workflow effects in untouched data.

Several limitations constrain interpretation. The retrospective single-centre design, selected CAG-referred cohort, and absence of independent external validation limit generalisability. Labels came from clinical CAG reports rather than blinded core-laboratory or quantitative angiographic review. Structured CCTA inputs omitted original images and plaque features. Multiple ECGs per patient may have produced unequal training contributions despite patient-grouped folds. Model selection and evaluation occurred within the same cross-validation framework, and no nested test cohort was available. Fold means lacked patient-level confidence intervals and formal paired

comparisons. Calibration curves, decision-curve analysis, prespecified screening thresholds, and prospective clinical-utility data were also unavailable. Finally, the study addressed only the RCA.

Conclusions

We developed a multimodal deep-learning model that combined resting ECGs with structured CCTA report features to identify CAG-report-defined RCA diameter stenosis of at least 70%.

Internal patient-grouped cross-validation showed a sensitivity of 0.849 and an AUC of 0.831. Relative to the structured-report model, the sensitivity gain was accompanied by lower specificity. These results support further study of ECG-CCTA report fusion as a sensitivity-oriented secondary screening aid in selected referral populations.

The present study does not establish safe rule-out, fewer missed diagnoses, reduced invasive testing, or improved outcomes. Independent multicentre validation should lock the threshold in advance and evaluate calibration, false-negative consequences, and clinical net benefit before deployment.

Methods

Dataset construction and preprocessing

Study design, data source, and cohort selection

This single-centre retrospective study used the clinical database of the cardiovascular centre at The First Affiliated Hospital with Nanjing Medical University. The study included patients treated between October 2019 and March 2025. Adults were eligible if they had a resting 12-lead ECG, CCTA, and a formal CAG report from the same episode of care. Patients were excluded if any examination was unavailable. For repeated examinations, the most recent CCTA and CAG were selected. This selected referral cohort does not represent population-based screening.

Only ECGs acquired within seven days before or after CAG were included to improve temporal alignment between the electrical signals and reference anatomy. This window was intended to reduce mismatch caused by disease progression or treatment interventions[21].

The final cohort comprised 1,904 eligible ECG records from 760 patients. Each record was matched to structured CCTA report features and a CAG label. Some patients contributed multiple ECGs. Training and patient-level aggregation therefore incorporated the interval between each ECG and CAG.

Reference standard and target outcome

RCA stenosis severity was obtained from formal CAG reports issued by the cardiovascular centre. With reference to CAD-RADS 2.0 stenosis ranges[22], we extracted the maximum RCA luminal diameter stenosis. When a report described multiple segments, the most severely stenosed segment determined the label.

The task was defined as patient-level binary classification. For patient j, RCA diameter stenosis of at least 70% in the CAG report was labelled positive for severe stenosis, $y_{RCA}^{(j)} = 1$. Stenosis below 70% was labelled negative, $y_{RCA}^{(j)} = 0$. This anatomical threshold supported model development and does not define revascularisation eligibility or a clinical rule-out decision.

Structured CCTA report features

Formal CCTA reports provided the maximum reported stenosis of the left main coronary artery (LM), left anterior descending artery (LAD), left circumflex artery (LCX), and RCA. These four report-derived values formed the anatomical input to the structured-report and multimodal models.

CCTA reports expressed stenosis as an exact percentage, an interval such as 50% – 70%, or a qualitative category such as mild, moderate, or severe. The preprocessing workflow mapped these descriptions to standardised percentage values using CAD-RADS 2.0 stenosis ranges and scaled them to [0,1] by min – max normalisation.

Each patient's structured CCTA information was therefore represented as a four-dimensional vector:

$$C^{(j)} = \left[c_{LM}^{(j)}, c_{LAD}^{(j)}, c_{LCX}^{(j)}, c_{RCA}^{(j)}\right]$$

This four-dimensional representation retained only a coarse summary from clinical reports. It

did not include original CCTA images, calcium burden, plaque morphology, remodelling, or other image-level features.

ECG preprocessing

All ECGs were standard resting 12-lead recordings (I, II, III, aVR, aVL, aVF, and V1 – V6). Before model input, the raw signals underwent resampling, filtering, beat selection, and amplitude standardisation to reduce acquisition-system differences and noise.

ECG signals generated by different acquisition systems were resampled to 1,000 Hz to standardise temporal resolution.

A 0.5 – 100-Hz Butterworth band-pass filter was then applied to remove low-frequency baseline drift and high-frequency noise.

R peaks were detected using the Pan – Tompkins algorithm[23], and cardiac cycles containing the P – QRS – T complex were extracted around each R peak. Beats from ECGs with regular rhythm and acceptable signal quality were averaged to generate a representative beat; for records with arrhythmia or unstable signals, a fixed-length raw-signal segment of 1,000 samples was extracted.

The final ECG input tensor had dimensions of 12 leads × 1,000 samples. The amplitude of each lead was z-score standardised to improve comparability across patient records.

Temporal-decay weighting

Some ECGs and CAG examinations were not obtained on the same day. Temporal-decay weighting was therefore introduced during training to represent the interval between each record and the reference examination.

Let Δt denote the absolute interval in days between ECG acquisition and CAG. The weight for each ECG was defined as w_i:

$$w_i = e^{-\frac{\Delta t}{\tau}}$$

The time constant τ was set to three days. ECGs closer to CAG consequently contributed more to the loss, while all included records remained within the seven-day window.

Temporal-decay weighting represents the modelling assumption that more recent examinations may better reflect status at the reference time point[21]. Its contribution was evaluated through ablation, but a performance difference cannot by itself prove that earlier ECGs necessarily contain label noise.

Model architecture: ECG backbone conditioned by structured CCTA reports

The multimodal framework used ECG as its principal input and structured CCTA report features as conditioning information. The design retained temporal waveform information while allowing coarse anatomical findings to modulate the ECG representation. Figure 7 shows the overall architecture.

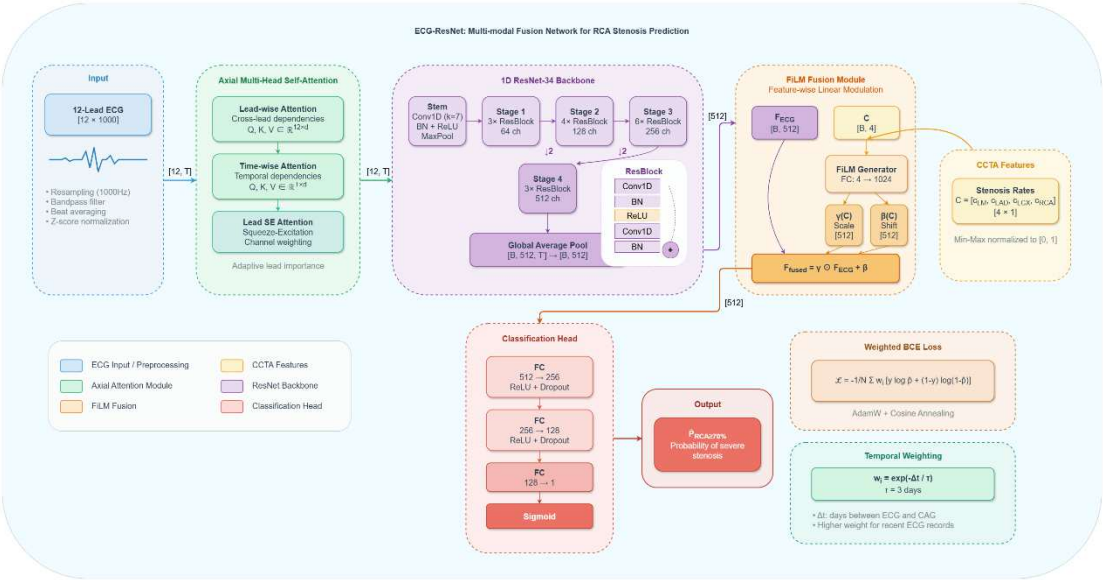


Fig. 7 Architecture for ECG and structured CCTA report fusion

ECG feature-extraction backbone

A one-dimensional residual network (1D ResNet-34) was used as the ECG feature-extraction backbone. One-dimensional convolutions are suited to direct modelling of ECG time series[24].

The network comprised an initial convolutional layer and four residual stages, each containing multiple residual blocks, with the number of channels increasing progressively from 64 to 512. This hierarchy was used to extract temporal representations over progressively longer timescales from local waveform morphology.

Residual connections supported gradient propagation during deep-network training.

Axial attention and lead modelling

Because convolutional layers have limited capacity to model long-range dependencies and cross-lead relationships, an axial multi-head self-attention (axial MHSA) module was introduced before the ResNet backbone[25]. The module received a 12 × 1,000 tensor and comprised four stacked layers. Each layer contained separate lead-axis and time-axis self-attention branches, each with four attention heads and a feature dimension, d_model, of 128.

Attention was computed separately along the lead and temporal axes to model inter-lead relationships and dependencies between time points within each lead.

This factorisation reduced the computational cost of global two-dimensional attention while preserving the lead-by-time structure of the ECG.

The model further used a squeeze-and-excitation (SE)-style lead-attention module[26] to adaptively weight representations from the 12 leads. The mean and standard deviation of each lead were calculated along the temporal dimension and concatenated into a 24-dimensional descriptor. This descriptor was mapped to 12 weights by fully connected layers (24→16→12, with ReLU activation in the hidden layer). After Softmax normalisation, the weights were multiplied element-wise by the time series from each lead, and the weighted tensor was passed to the ResNet backbone. These weights reflect only the model's internal allocation of attention and should not be interpreted as the clinical or causal importance of individual leads.

Multimodal fusion using FiLM

Structured CCTA report features were fused through conditional modulation rather than concatenation only at the classifier input.

Feature-wise linear modulation (FiLM)[27] used the CCTA vector to generate channel-wise scaling and offset parameters for the ECG representation.

Given an ECG representation F_{ecg} and CCTA vector C, the FiLM generator used C to produce scaling $\gamma(C)$ and offset $\beta(C)$, yielding F_{fused}:

$$F_{fused} = \gamma(C) \odot F_{ecg} + \beta(C)$$

where $\odot$ denotes element-wise multiplication.

The CCTA report features therefore conditioned the ECG feature distribution through channel-wise scaling and shifting. They were not appended only at the classifier input.

Ablation analysis assessed whether this conditional fusion contributed to internal performance.

Training, internal validation, and performance measures

Training strategy

The model was implemented in PyTorch. Binary cross-entropy (BCE) was used as the base loss for the binary classification task, and temporal-decay weights adjusted the contribution of each ECG record to the training loss. The loss was defined as:

$$L = -\frac{1}{N}\sum_{i=1}^{N} w_i \left[y_{RCA}^{(j)} \log \hat{p}_i + (1 - y_{RCA}^{(j)}) \log(1 - \hat{p}_i) \right]$$

Here, N is the number of records in a batch. $\hat{p}_i$ is the predicted probability of severe RCA stenosis for record i, and w_i is its temporal weight.

To address class imbalance, we compared class-weighted BCE and variants such as focal loss[28]. Based on training stability and validation performance, the final model used weighted random sampling in combination with standard BCE without further increasing loss-function complexity.

The AdamW optimiser[29] was used with an initial learning rate of 3×10^{-5} and a weight decay of 1×10^{-3}. The first three epochs used warm-up, followed by cosine annealing[30] to reduce the learning rate to 1×10^{-6}. The batch size was 8, the random seed was 42, automatic mixed precision was used, and training continued for 40 epochs. The best epoch for each fold was selected according to patient-level F1 score for RCA stenosis of at least 70% in that fold's internal validation set; data from the other folds were not used to select the epoch or hyperparameters.

Internal validation and performance measures

We used patient-grouped stratified fivefold cross-validation for internal evaluation. Stratification used the highest stenosis grade across the four coronary arteries. All ECGs from one patient remained in the same fold to reduce patient-level leakage. In each rotation, approximately 80% of patients were used for training and 20% for validation and model selection. The study did not include a separate external or nested test cohort. We therefore report the fold-wise validation mean and standard deviation as internal performance estimates.

Each fold was evaluated using patient-level sensitivity, specificity, accuracy, F1 score, AUC, and Brier score. For a potential secondary-screening interpretation, sensitivity and specificity were considered jointly to characterise the case-finding trade-off. Record-level probabilities were aggregated before calculating the primary patient-level metrics.

For patients with multiple ECG predictions, temporal-decay weights w_i were used to calculate the weighted patient-level probability $\hat{p}_{patient}$. The primary patient-level analysis used a threshold of 0.5.

This aggregation gave greater weight to ECGs obtained closer to CAG. Because the threshold was not independently locked for a target sensitivity, the reported operating point should be interpreted as an internal estimate rather than a deployable screening threshold.

List of abbreviations

RCA: right coronary artery; CCTA: coronary computed tomography angiography; ECG: electrocardiogram; CAG: invasive coronary angiography; AUC: area under the receiver operating characteristic curve; F1: F1 score; AI: artificial intelligence; ROC: receiver operating characteristic; PR: precision–recall; AP: average precision; PPV: positive predictive value; NPV: negative predictive value; Grad-CAM: gradient-weighted class activation mapping; QRS: QRS complex; ST–T: ST segment and T wave; FiLM: feature-wise linear modulation; CAD-RADS: coronary artery disease reporting and data system; LM: left main coronary artery; LAD: left anterior descending artery; LCX: left circumflex artery; 1D: one-dimensional; ResNet: residual network; MHSA: multi-head self-attention; SE: squeeze-and-excitation; ReLU: rectified linear unit; BCE:

binary cross-entropy; AdamW: Adam optimizer with decoupled weight decay.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the First Affiliated Hospital of Nanjing Medical University (approval no. 2024-SR-882.A1) and was conducted in accordance with the Declaration of Helsinki. Owing to the retrospective nature of the study and the use of anonymised clinical data, the requirement for informed consent was waived.

Consent for publication

The manuscript contains no identifiable personal information, images, or videos; therefore, this section is not applicable.

Availability of data and materials

The patient-level ECG, CCTA report, and CAG report data used in this study are not publicly available because of patient privacy, the scope of informed consent, and institutional data-governance requirements. The data are available from the corresponding author on reasonable request. Analysis code is available from the GitHub project page: https://github.com/Hyyygit/Hyyygit-AI-ECG-CCTA-Fusion-Classification; the public repository contains no patient data.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Key Research and Development Program of China through the project “Regional Comprehensive Application Demonstration of Digital and Intelligent Technologies for Proactive Health Services” (Grant No. 2023YFC3605800).

Authors' contributions

Yue He collected and processed the data, developed and trained the models, performed the analyses, and drafted the manuscript. Qiuyang Yang contributed to study design, data interpretation, and manuscript revision. Yong Li contributed to clinical data interpretation and manuscript revision. Shenqi Jing provided methodological guidance and assisted with model evaluation. Yun Liu conceived and supervised the study, provided overall guidance on methodology and clinical interpretation, coordinated the research process, and critically revised the manuscript. Tao Shan contributed to study design, clinical collaboration, and manuscript revision. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the clinicians and technical staff of the cardiovascular centre at The First Affiliated Hospital with Nanjing Medical University for their assistance with clinical data retrieval and data management. The authors also acknowledge the faculty members and clinical experts at Nanjing Medical University for their guidance on the medical background, clinical interpretation, and clinical relevance of this study.

References

1. Nedkoff L, Briffa T, Zemedikun D, Herrington S, Wright FL. Global Trends in Atherosclerotic Cardiovascular Disease. *Clinical Therapeutics* 2023;45:1087-1091
2. Shao C, Wang J, Tian J, Tang Y-d. *Coronary Artery Disease: From Mechanism to Clinical Practice*. In: Wang M, ed. *Coronary Artery Disease: Therapeutics and Drug Discovery*. Singapore: Springer Singapore, 2020:1-36
3. Trimarchi G, Pizzino F, Paradossi U, Gueli IA, Palazzini M, Gentile P, et al. Charting the Unseen: How Non-Invasive Imaging Could Redefine Cardiovascular Prevention. *Journal of Cardiovascular Development and Disease* 2024;11:245
4. Saxton A, Chaudhry R, Manna B. *Anatomy, thorax, heart right coronary arteries*. In. *StatPearls [Internet]*: StatPearls Publishing, 2025
5. Ghobrial M, Haley HA, Gosling R, Rammohan V, Lawford PV, Hose DR, et al. The new role of diagnostic angiography in coronary physiological assessment. *Heart* 2021;107:783-789

6. Marano R, Rovere G, Savino G, Flammia FC, Carafa MRP, Steri L, et al. CCTA in the diagnosis of coronary artery disease. *La radiologia medica* 2020;125:1102-1113
7. Zuza I, Nadarevic T, Jakljevic T, Bartolovic N, Kovacic S. The Effect of Severe Coronary Calcification on Diagnostic Performance of Computed Tomography-Derived Fractional Flow Reserve Analyses in People with Coronary Artery Disease. *Diagnostics (Basel)* 2024;14
8. Birnbaum Y, Wilson JM, Fiol M, de Luna AB, Eskola M, Nikus K. ECG Diagnosis and Classification of Acute Coronary Syndromes. *Annals of Noninvasive Electrocardiology* 2014;19:4-14
9. Wu Z, Guo C. Deep learning and electrocardiography: systematic review of current techniques in cardiovascular disease diagnosis and management. *BioMedical Engineering OnLine* 2025;24:23
10. Khalifa A, Abdulaziz M, Parvez SS, Salem A, Panhwer HS, Saleem U, et al. The Digital Revolution in Cardiac Ischemia: Artificial Intelligence (AI)-Enhanced Detection, Diagnosis, and Risk Stratification. *Cureus* 2025;17:e95059
11. Elias P, Poterucha TJ, Rajaram V, Moller LM, Rodriguez V, Bhave S, et al. Deep Learning Electrocardiographic Analysis for Detection of Left-Sided Valvular Heart Disease. *Journal of the American College of Cardiology* 2022;80:613-626
12. Park J, Kim J, Kang S-H, Lee J, Hong Y, Chang H-J, et al. Artificial intelligence-enhanced electrocardiography analysis as a promising tool for predicting obstructive coronary artery disease in patients with stable angina. *European Heart Journal - Digital Health* 2024;5:444-453
13. Ding C, Yao T, Wu C, Ni J. Advances in deep learning for personalized ECG diagnostics: A systematic review addressing inter-patient variability and generalization constraints. *Biosensors and Bioelectronics* 2025;271:117073
14. Xu X, Li J, Zhu Z, Zhao L, Wang H, Song C, et al. A Comprehensive Review on Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis. *Bioengineering* 2024;11:219
15. Yang XY, Li YM, Wang JY, Jia YH, Yi Z, Chen M. Utilizing multimodal artificial intelligence to advance cardiovascular diseases. *Precis Clin Med* 2025;8:pbaf016
16. Okabe S, Nakamura S, Hashimoto N, Onoda N, Murata T, Yamaoka Y, et al. Associating electrocardiographic abnormalities with coronary artery disease: insights into microvascular dysfunction from dynamic CT perfusion. *European Radiology* 2025;35:7096-7108
17. Suchá D D. The plaque factor: Can coronary CT angiography plaque analysis link stenosis and ischemia? *European Radiology* 2025;35:6268-6270
18. Heusch G. Myocardial ischemia/reperfusion: Translational pathophysiology of ischemic heart disease. *Med* 2024;5:10-31
19. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization.

In:*Proceedings of the IEEE international conference on computer vision*, 2017; 618-626
20. Alencar JNd, Matos VFdA, Scheffer MK, Felicioni SP, Marchi MFND, Martínez-Sellés M. ST segment and T wave abnormalities: A narrative review. *Journal of Electrocardiology* 2024;85:7-15
21. Rossi A. *Electrocardiogram in Ischemic Heart Disease*. In. *Ischemic Heart Disease: From Diagnosis to Treatment*: Springer, 2023:171-189
22. Cury RC, Leipsic J, Abbara S, Achenbach S, Berman D, Bittencourt M, 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). *JACC Cardiovasc Imaging* 2022;15:1974-2001
23. Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering* 1985;BME-32:230-236
24. Khan F, Yu X, Yuan Z, Rehman Au. ECG classification using 1-D convolutional deep residual neural network. *PLOS ONE* 2023;18:e0284791
25. Ho J, Kalchbrenner N, Weissenborn D, Salimans T. Axial Attention in Multidimensional Transformers. In, 2019
26. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In:*Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018; 7132-7141
27. Perez E, Strub F, De Vries H, Dumoulin V, Courville A. Film: Visual reasoning with a general conditioning layer. In:*Proceedings of the AAAI conference on artificial intelligence*, 2018
28. Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In:*Proceedings of the IEEE international conference on computer vision*, 2017; 2980-2988
29. Loshchilov I, Hutter F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101* 2017
30. Loshchilov I, Hutter F. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983* 2016