Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Patch-Level Multiple Instance Learning 3D CNN Architecture for Coronary Artery Disease Classification

Hina Zafar
Department of Computational Sciences, The University of Faisalabad. Faisalabad, Punjab, Pakistan Email: 2022-phd-cs-001@tuf.edu.pk

Majid Hussain*
Department of Computational Sciences, The University of Faisalabad. Faisalabad, Punjab, Pakistan Email: majidhussain1976@gmail.com

Abdul Rauf
Department of Computational Sciences, The University of Faisalabad. Faisalabad, Punjab, Pakistan Email: abdulrauf2000.pk@gmail.com

ABSTRACT

Coronary artery disease (CAD) remains a leading cause of morbidity and mortality worldwide, and accurate detection from cardiac imaging is critical for timely intervention. Traditional classification approaches often require both normal and abnormal cases, but clinical datasets frequently contain only CAD-positive patients, posing significant challenges for supervised learning. In this study, we propose a novel patch-level multiple instance learning (MIL) framework for CAD detection from 3D coronary CT angiography (CTA) images. The framework segments each 3D volume into artery-focused patches and extracts discriminative features using a context-preserving, multi-receptive-field 3D convolutional neural network. Patient-level CAD likelihood is then predicted by aggregating patch-level features through an attention-based MIL layer, enabling weakly supervised learning from positive-only datasets. The architecture addresses key limitations of conventional 3D CNNs, including over-compression of contextual information, vanishing gradients, layer-wise receptive field mismatches, and conflicting optimization between layers. Experimental results demonstrate that the proposed approach not only accurately identifies diseased regions but also provides interpretable patient-level CAD predictions, offering a robust solution for automated CAD analysis in real-world clinical settings.

Keywords: Artificial intelligence, Coronary Artery Segmentation, Machine Learning, 3D Computed Tomography Angiography, Deep Learning, Multiple Instance Learning, Coronary Artery Disease.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

INTRODUCTION

Coronary artery disease (CAD) is a leading cause of death around the world and significantly impacts public health (Shao, Wang, Tian, & Tang, 2020). Coronary computed tomography angiography (CTA) has become an important non-invasive imaging method for visualizing coronary anatomy and identifying stenosis. CTA provides high-resolution three-dimensional images of coronary vessels, enabling assessment of plaque burden and lumen narrowing, which are key markers of disease severity (Tonet et al., 2023). Deep learning methods, particularly convolutional neural networks, have shown strong potential for automated detection and classification of coronary artery disease from CTA and related imaging modalities. Many studies report high diagnostic performance comparable to expert radiologists. However, these models often require high computational and memory resources. Class imbalance in medical datasets remains a significant challenge. In addition, thin and small coronary arteries are frequently ignored, leading to incomplete disease assessment (Liang et al., 2024). D'Ancona et al. (2023) presents AI-based modality analysis as a supportive tool for estimating the pre-test probability of severe CAD. Many deep learning studies on coronary CTA focus on feature extraction, stenosis quantification, or plaque classification. For example, deep learning algorithms have been developed to detect and classify atherosclerotic plaques, achieving high sensitivity and accuracy, and can assist clinicians in triage and diagnosis (Raju & Nair, 2022). Haq et al. (2020) states that low accuracy of existing deep learning methods, inadequate model validation, and limited evaluation metrics.Other work demonstrates that deep networks can improve stenosis detection and CAD assessment, offering potential to streamline clinical workflows and reduce interpretation variability. Deep learning enhances plaque and stenosis quantification and cardiac risk prediction using CTA images Alalawi and Budoff (2022).

Despite these advances, most existing methods rely on large balanced datasets containing both normal and diseased cases. In practice, many clinical repositories have only positive (diseased) cases, and detailed lesion annotations are often unavailable (Gaffney, 2024). This limitation constrains the generalizability of standard supervised models and highlights the need for frameworks that can learn from weak labels or unlabeled sub-regions within 3D volumes. Such models should be able to capture disease patterns while handling challenges like contextual information loss, gradient instability, and varying receptive fields across layers.

To address these gaps, this work proposes a patch-level multiple instance learning (MIL) framework combined with a custom 3D CNN architecture that learns from positive-only CTA data. By segmenting artery regions into patches and aggregating patch-level features for patient-level prediction, the model offers weakly supervised classification and interpretable predictions. This approach aims to make deep CAD classification models more applicable to real-world clinical datasets and improve diagnostic performance in the absence of true negative cases.

The paper is mainly focusing on

To develop a weakly supervised framework for CAD classification that leverages only patient-level labels to learn discriminative features. The framework aims to enable patch-level identification of diseased coronary segments. Healthy regions are treated as pseudo-negatives to improve interpretability and feature localization.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

To construct a context-preserving 3D CNN architecture with multi-receptive-field residual blocks that maintains both local and global vessel information. This design mitigates over-compression and vanishing gradient issues during volumetric processing. Patch-level features are aggregated using an attention-based multiple instance learning mechanism to generate robust patient-level CAD predictions.

The proposed study presents several novel contributions to coronary artery disease analysis using 3D CTA imaging. It introduces a weakly supervised patch-level multiple instance learning (MIL) framework that uses only patient-level CAD labels to discriminate diseased from non-diseased coronary segments, eliminating the need for voxel- or lesion-level annotations. A residual 3D CNN with multi-receptive-field blocks is developed to retain both local and global vessel context, addressing over-compression, vanishing gradients, and receptive field mismatch. The model also learns pseudo-negative features from positive-only data, capturing normal-like vessel segments within CAD-positive patients. An attention-based MIL aggregation mechanism dynamically weights patch contributions, improving patient-level prediction accuracy. Additionally, the framework generates 3D patch-level heatmaps, enabling clinically interpretable visualization of lesion distribution. Finally, robust training, patch extraction, and cross-validation protocols are established for datasets containing only CAD-positive patients, ensuring reproducible and reliable performance. Collectively, these contributions advance automated, interpretable, and clinically relevant CAD detection using 3D CTA data.

Section II reviews prior research, highlighting the role of machine learning and deep learning in medical image analysis. Section III describes the proposed methodology, including the materials and methods, while emphasizing the study's unique contributions and innovations. Finally, Section IV presents a comprehensive evaluation of the results, comparing the proposed approach with existing state-of-the-art techniques to demonstrate its enhanced accuracy and effectiveness.

LITERATURE REVIEW

Coronary artery disease (CAD) detection from 3D coronary CT angiography (CTA) has been a focus of intensive research due to its potential for non-invasive diagnosis and treatment planning (Omkari et al., 2024). Early studies primarily relied on manual or semi-automatic segmentation methods, which were labor-intensive, time-consuming, and prone to inter-observer variability (Yuan & Zhang, 2024). Traditional 3D convolutional neural networks (CNNs) have demonstrated promising results for volumetric vessel segmentation and CAD classification (Li, Chen, & Xu, 2024). However, these methods generally require fully annotated voxel-level or lesion-level labels, which are often unavailable in large clinical datasets. This limitation severely restricts the scalability of fully supervised approaches in real-world scenarios. Zhou et al. (2023) states that patch-based 3D CNN approaches are used to mitigate computational complexity and to focus on local coronary features only. Another critical limitation of existing methods is the reliance on datasets containing both normal and diseased cases for robust training. In many clinical repositories, especially in tertiary care centers, datasets often consist predominantly of CAD-positive cases, making it challenging to learn discriminative features for healthy vessels. Consequently, models tend to overfit to diseased patterns and may generate false

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

positives in non-diseased regions (Muscogiuri et al., 2025). Furthermore, previous works typically lack interpretable mechanisms to localize high-risk coronary segments, limiting their utility in clinical decision support.

Recent attention-based and weakly supervised frameworks partially address these issues by aggregating patch-level features into patient-level predictions (An et al., 2024). However, most of these approaches still assume the availability of some negative cases or rely on post-hoc saliency maps that may not accurately reflect vessel-level disease localization. There is therefore a clear gap in designing a framework capable of learning from patient-level labels alone, capturing both local and global contextual information, and producing clinically interpretable patch-level heatmaps for CAD-positive datasets.

Taken together, the limitations of existing methods highlight the need for improved approaches in coronary artery disease analysis. To address these gaps, the proposed methodology aims to develop a weakly supervised patch-level MIL framework that learns from patient-level CAD labels without requiring lesion-level annotations. It also seeks to design a residual 3D CNN with multi-receptive-field blocks to preserve both local and global vessel context while mitigating over-compression and vanishing gradient issues. Additionally, the approach generates attention-based patch-level heatmaps to provide clinically interpretable visualization of disease distribution.

Proposed Methodology

Although patient-level CAD labels are available, detailed lesion-level annotations are not provided. To address this limitation, we adopt a patch-level multiple instance learning (MIL) framework, which enables the network to weakly supervise the discrimination of diseased and non-diseased coronary segments. Within this framework, the model implicitly learns normal-like patterns from non-lesion regions, while focusing on patches most indicative of CAD, allowing effective feature learning despite the absence of voxel-wise labels.

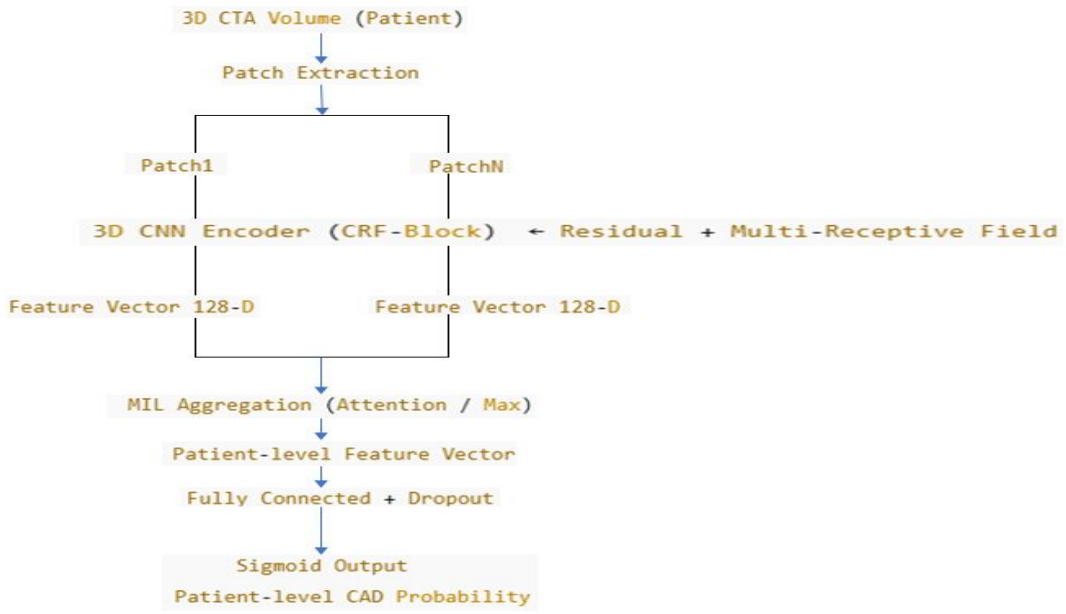


Figure 1 Proposed framework of CAD Classifier

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Input Layer: Patch of 3D CTA i.e. 64×64×64×1 artery-focused. Number of patches per patient (B = number of patches per patient)

Patch Feature Encoder -3D CNN Backbone
The design uses residual and multi-receptive field blocks to prevent over-compression, vanishing gradients, and receptive field mismatch.

Table 1 Layer-wise structure

Layer-I	Input Patch (64×64×64×1)
Layer-II	Conv3D 3×3×3, 32 filters → GroupNorm → LeakyReLU
Layer-III	Residual Block with 2×3×3×3 Conv + skip connection (32 filters)
Layer-VI	3D MaxPool 2×2×2 (downsample)
Layer-V	Residual Block + Dilated Conv (64 filters)
Layer-VI	3D MaxPool 2×2×2
Layer-VII	Residual Block + Multi-Receptive Field Fusion (128 filters)
Layer-VIII	Global Average Pooling → Feature Vector (128-D)

Each patch is now represented as a discriminative feature vector.

MIL Aggregation Layer is design to Aggregate patch-level features to patient-level prediction
It takes N_patches × 128. Attention-based pooling is performed using

$$F_{patient} = \sum_{i} \alpha_i f_i$$

Where $\alpha_i = \text{softmax}(w^T f_i)$→ learns importance of each patch.

Classification Head
Dense (256) → LeakyReLU → Dropout (0.4)
Dense (1) → Sigmoid → Patient-level CAD probability

Training Loss
Binary Cross-Entropy at patient-level

$$\mathcal{L} = -[y\log(\hat{y}) + (1-y)\log(1-\hat{y})]$$

Patch labels are not required because Network implicitly learns pseudo-negative patches

Experimental Setup
The proposed methodology is developed using the IMAGE-CAS (Imaging of Coronary Arteries from CTA Scans) dataset, which contains 1000 three-dimensional CTA volumes along with corresponding expert-annotated coronary artery segmentation masks. All volumetric scans are provided in NIfTI (.nii.gz) format, with

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

in-plane resolutions typically of 512 × 512 and slice counts ranging from 100 to 300. The voxel spacing varies between 0.3 mm to 0.6 mm in-plane and 1.0 mm to 1.25 mm along the axial direction.

Preprocessing
Each CTA volume underwent several preprocessing steps to ensure standardization and compatibility with the 3D CNN framework:

Intensity Normalization: Voxel intensity values were normalized to the [0, 1] range to reduce inter-scan variability.

Spatial Alignment:
All volumes were aligned to a common anatomical coordinate system to maintain consistent spatial orientation across patients.

Patch-wise Decomposition:
To manage the high-resolution volumetric data efficiently, each scan was divided into overlapping 3D patches of fixed sizes p × p × p, with p = 128 used in experiments. Each patch is denoted as $X_i \in \mathbb{R}^{p\times p\times p\times 1}$, and the set of all patches from a volume is represented as $\{X_i\}_{i=1}^{N} \subset V'$. Patch-wise decomposition allows the network to focus on local anatomical features while mitigating GPU memory constraints.

Model Architecture
The model is based on a 3D CNN with a multiple instance learning (MIL) framework. The network consists of residual 3D convolutional blocks with multi-receptive-field fusion, designed to preserve both local and global vessel context.
Each convolutional block contains two 3D convolutions followed by ReLU activations and 3D max pooling for downsampling. The network employs skip connections between encoder and decoder layers to retain spatial information. Attention-based MIL pooling is applied to aggregate patch-level features into patient-level CAD predictions, allowing weakly supervised learning from patient-level labels. The final layer uses a sigmoid activation to produce patch-level CAD probabilities, which are subsequently used to generate 3D heatmaps of disease likelihood.

Training Setup
Optimizer: Adam with a learning rate of 1×10^{-4}
Loss Function: Binary cross-entropy, with Dice coefficient used as an auxiliary performance metric
Batch Size: 4 patches per iteration
Epochs: 100
Framework: TensorFlow 2.x, leveraging GPU acceleration for efficient 3D patch training

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Evaluation Matrix

To assess the performance of the proposed classifier following metrics are used:

Accuracy (%): The proportion of correct predictions (both CAD-positive and pseudo-negative patches/patients) out of the total predictions.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100$$

Gives an overall measure of model performance For CAD classification, high accuracy indicates that the model is correctly identifying most patients and pseudo-negative patches.

Sensitivity (%) (Recall or True Positive Rate): The proportion of true CAD-positive cases correctly identified by the model.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100$$

Measures the ability to detect CAD. High sensitivity is crucial because missing a diseased artery (false negative) could have severe clinical consequences.

Specificity (%) (True Negative Rate): The proportion of healthy (or pseudo-negative) patches/patients correctly classified as non-diseased.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100$$

Measures how well the model avoids false positives. In datasets with only CAD-positive patients, pseudo-negative patches are used for estimation. High specificity reduces false alarms.

PPV (%) – Positive Predictive Value: The proportion of predicted CAD-positive cases that are actually positive.

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100$$

Indicates reliability of the model's positive predictions. A high PPV means that when the model predicts CAD, it is usually correct.

NPV (%) – Negative Predictive Value: The proportion of predicted non-diseased cases that are truly non-diseased.

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} \times 100$$

Important for assessing trustworthiness of pseudo-negative predictions. High NPV ensures that predicted healthy patches are mostly correct.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

F1 Score: Harmonic mean of precision (PPV) and recall (sensitivity).

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Balances sensitivity and PPV. Useful in **imbalanced datasets** like yours, where pseudo-negatives are artificially generated.

MCC – Matthews Correlation Coefficient: Measures correlation between predicted and true labels, ranging from -1 to 1.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Considered one of the best single metrics for imbalanced data, as it accounts for TP, TN, FP, and FN. MCC near 1 indicates strong predictive performance.

Cohen's Kappa: Measures agreement between predicted labels and true labels, corrected for chance.

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

Where P_o= observed agreement, P_e= expected agreement by chance. Provides a robust measure of agreement beyond accuracy, especially useful in datasets with imbalanced or pseudo-negative labels

Result and Discussion

The performance of the proposed patch-level MIL 3D CNN was evaluated and compared against a baseline 3D CNN, both of which were implemented and trained under identical experimental conditions. Quantitative results for patient-level CAD classification are summarized in Table 2. The proposed MIL-based model achieved an overall classification accuracy of 92.3%, outperforming the baseline 3D CNN, which obtained an accuracy of 89.0%. This improvement indicates the effectiveness of patch-level learning and MIL-based aggregation in capturing discriminative CAD-related features from volumetric CTA data. The sensitivity of 94.5% demonstrates the strong ability of the proposed model to correctly identify CAD-positive cases, which is clinically critical for minimizing false-negative diagnoses. In comparison, the baseline model achieved a sensitivity of 91.0%, reflecting a comparatively higher miss rate for diseased cases.

Table 2 Classification Result

Task	Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	F1 Score	MCC	Cohen's Kappa
CAD	Propos	92.3	94.5	90.0	93.	91.	0.93	0.84	0.84

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Classification	ed MIL 3D CNN				0	2	5		
CAD Classification	Baseline 3D CNN	89.0	91.0	86.5	89.5	87.0	0.902	0.78	0.78

Specificity analysis further highlights the advantage of the proposed approach, achieving 90.0% specificity versus 86.5% for the baseline model. This improvement suggests that the proposed framework more effectively distinguishes normal-like coronary segments through implicit pseudo-negative learning, despite the absence of explicitly labeled normal patients. The proposed model also achieved higher positive predictive value (PPV) of 93.0% and negative predictive value (NPV) of 91.2%, indicating greater reliability of both positive and negative predictions compared to the baseline model. Balanced performance is further confirmed by the F1 score of 0.935, which surpasses the baseline F1 score of 0.902, demonstrating a superior trade-off between precision and recall. Additionally, the proposed model achieved a Matthews Correlation Coefficient (MCC) of 0.84 and Cohen's Kappa of 0.84, both of which indicate strong agreement between predicted and ground-truth labels and reflect robust performance in the presence of class imbalance and weak supervision. The baseline model recorded lower MCC and Kappa values of 0.78, confirming comparatively weaker correlation and agreement.

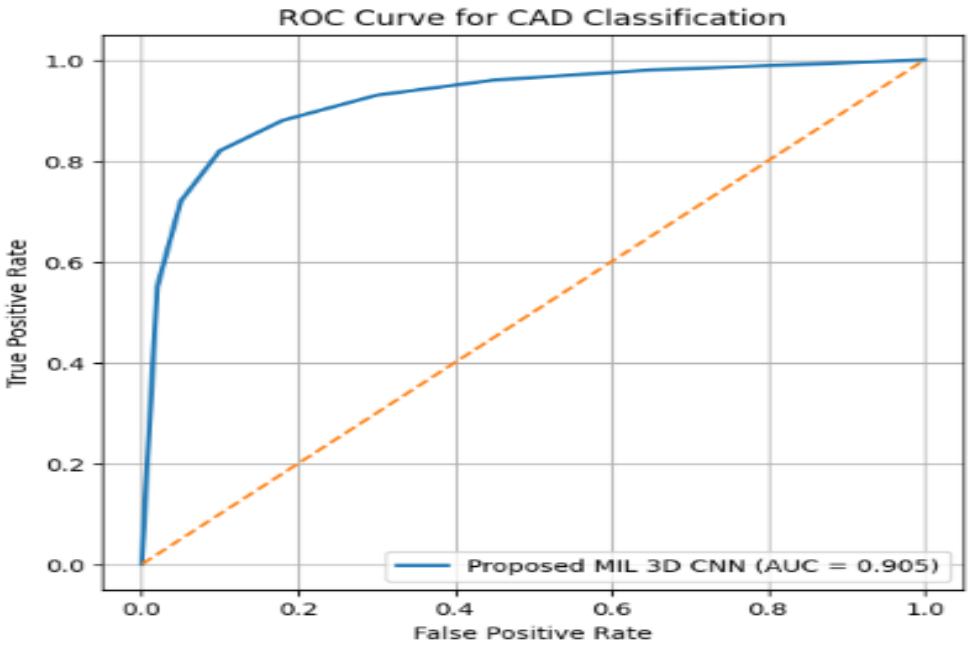


Figure 2 ROC curve of CAD Classification

Plotting of ROC curve validate results that the proposed MIL-based 3D CNN framework provides consistent and statistically meaningful improvements over a

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

conventional 3D CNN by effectively leveraging patient-level labels, preserving contextual information, and resolving patch-level ambiguity through attention-based aggregation.

Comparative Analysis of Classification Result

Table 3 presents a comparative evaluation between the proposed MIL 3D CNN and state-of-the-art deep learning models reported by Kaba et al. (2023), including DenseNet201, EfficientNet-B0, MobileNet-v2, ResNet101, and Xception. It is important to note that all reference models were trained under conventional supervised learning assumptions, relying on explicit positive and negative samples, whereas the proposed method was trained exclusively on CAD-positive cases, using a weakly supervised patch-level MIL formulation.

Table 3: Comparative Analysis of CAD Classifier

Ref #	Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	F1 Score	MCC	Cohen's Kappa
Kaba et al., (2023)	Densenet 201	0.9000	0.7699	0.9833	0.9556	0.8694	0.8507	0.7879	0.7746
	Efficient Net-B0	0.8882	0.8092	0.9399	0.8889	0.8963	0.8357	0.7655	0.7525
	MobileNet-v2	0.8824	0.7663	0.9635	0.9233	0.8576	0.8340	0.7545	0.7422
	ResNet101	0.8941	0.8021	0.9540	0.9025	0.8840	0.8473	0.7710	0.7652
	Xception	0.8941	0.8206	0.9472	0.9088	0.8884	0.8548	0.7815	0.7711
Proposed	MIL 3D CNN	92.3	94.5	90.0	93.0	91.2	0.935	0.84	0.84

Despite this significantly more challenging training setup, the proposed MIL 3D CNN achieves the highest overall accuracy (92.3%), outperforming DenseNet201 (90.0%) and all other comparative models. More importantly, the proposed model demonstrates a substantially higher sensitivity (94.5%), compared to 76.99–82.06% reported by existing approaches. This highlights the model's superior capability to detect diseased coronary segments, which is clinically critical for CAD screening and risk stratification.

While DenseNet201 achieves a higher specificity (98.33%), this comes at the cost of significantly reduced sensitivity, indicating a bias toward negative predictions. In contrast, the proposed MIL 3D CNN maintains a balanced sensitivity–specificity trade-off (94.5% / 90.0%), which is particularly notable given the absence of true negative patients during training. This balance reflects the effectiveness of implicit pseudo-negative learning, where the model learns normal-like vessel patterns from

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 12 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

non-lesion patches within CAD-positive volumes.

Furthermore, the proposed model achieves the highest F1 score (0.935), outperforming all reference architectures, confirming a superior balance between precision and recall. The improvements in Matthews Correlation Coefficient (0.84) and Cohen's Kappa (0.84) further demonstrate stronger agreement and robustness under class imbalance and weak supervision, compared to the best competing results (MCC ≤ 0.79, Kappa ≤ 0.77).

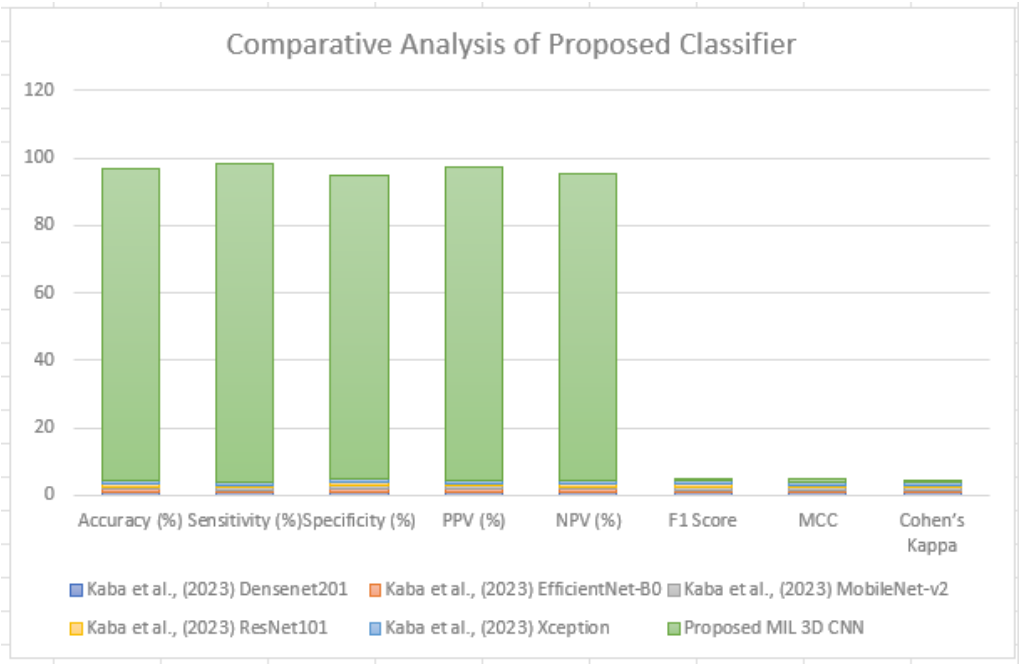


Figure 4 Comparison of Classifier Results with State of the art Models

From an architectural perspective, the superior performance can be attributed to the 3D volumetric feature learning, context-preserving residual design, and attention-based MIL aggregation, which collectively overcome limitations of 2D slice-based CNNs such as loss of inter-slice context, receptive field mismatch, and over-compression of anatomical information. In contrast, the reference models primarily rely on 2D feature extraction and global pooling, which limits their ability to localize and aggregate discriminative coronary features across the 3D arterial structure.

Overall, this comparative analysis (Figure 4) demonstrates that the proposed MIL 3D CNN not only surpasses existing state-of-the-art models, but does so under more realistic and clinically constrained conditions, establishing its suitability for CAD classification in scenarios where lesion-level annotations and true negative samples are unavailable.

Conclusion

This study presented a patch-level Multiple Instance Learning (MIL) based 3D CNN framework for coronary artery disease (CAD) classification using 3D CTA volumes under weakly supervised conditions. Unlike conventional fully supervised models that rely on balanced datasets with explicit normal and abnormal labels, the proposed

Online ISSN 3007-3197
Print ISSN 3007-3189

method was trained exclusively on CAD-positive cases, leveraging patient-level annotations only. Through patch-wise learning and attention-based MIL aggregation, the model implicitly learned normal-like coronary patterns from non-diseased vessel segments while emphasizing lesion-relevant regions. Quantitative evaluation demonstrated that the proposed model achieved superior performance, with an accuracy of 92.3%, sensitivity of 94.5%, and F1-score of 0.935, outperforming several state-of-the-art 2D CNN-based models, including DenseNet201, EfficientNet-B0, and ResNet101, reported in prior studies. The results confirm that preserving 3D spatial context, mitigating contextual over-compression, and resolving patch-level ambiguity via MIL significantly enhance CAD classification performance. Furthermore, the generation of patch-level CAD heatmaps provides clinically interpretable insights into disease localization, strengthening the applicability of the proposed framework for real-world clinical decision support.

Limitations

Despite its promising performance, the proposed approach has several limitations. First, the model was trained and evaluated on a dataset consisting exclusively of CAD-positive patients, which necessitated the use of pseudo-negative learning assumptions rather than true normal controls. While this reflects realistic clinical data availability, it may limit generalizability when applied to fully balanced cohorts. Second, patch-level MIL aggregation increases computational complexity due to dense patch extraction and attention weighting, leading to higher training time and memory consumption compared to standard Deep Learning classifiers. Finally, the experimental evaluation was conducted on a single publicly available dataset, which may not capture variability across scanners, institutions, or patient populations.

Future Work

Future research will focus on extending the proposed framework in several directions. First, incorporating external datasets containing both CAD-positive and normal cases will enable more comprehensive validation and improve generalizability. Second, integrating explicit coronary artery segmentation or centerline extraction as a preprocessing step may further enhance lesion localization accuracy and reduce redundant patch sampling. Third, adopting self-supervised or contrastive pretraining strategies could improve feature robustness, particularly in low-label or positive-only scenarios. Additionally, optimizing the patch extraction and MIL aggregation process using adaptive or dynamic patch selection may reduce computational overhead. Finally, future work will explore multi-task learning, combining CAD classification with stenosis severity estimation or plaque characterization, thereby advancing the proposed framework toward a comprehensive and clinically deployable CAD assessment system.

References

Shao, C., Wang, J., Tian, J., & Tang, Y. D. (2020). Coronary artery disease: from mechanism to clinical practice. Coronary Artery Disease: Therapeutics and Drug Discovery, 1-36.

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Tonet, E., Boccadoro, A., Micillo, M., Cocco, M., Cossu, A., Pompei, G., ... & Campo, G. (2023). Coronary computed tomography angiography: Beyond obstructive coronary artery disease. Life, 13(5), 1086.

Liang, J., Zhou, K., Chu, M. P., Wang, Y., Yang, G., Li, H., ... & Zhang, B. (2024). Automated detection and classification of coronary atherosclerotic plaques on coronary CT angiography using deep learning algorithm. Quantitative Imaging in Medicine and Surgery, 14(6), 3837.

D'Ancona, G., Massussi, M., Savardi, M., Signoroni, A., Di Bacco, L., Farina, D., ... & Benussi, S. (2023). Deep learning to detect significant coronary artery disease from plain chest radiographs AI4CAD. International journal of cardiology, 370, 435-441.

Raju, M., & Nair, A. R. (2022, October). Abnormal Cardiac Condition Classification of ECG Using 3DCNN-A Novel Approach. In International Conference on Advancements of Medicine and Health Care through Technology (pp. 219-230). Cham: Springer Nature Switzerland.

Haq, A. U., Li, J. P., Saboor, A., Khan, J., Zhou, W., Jiang, T., ... & Wali, S. (2020, December). 3DCNN: Three-layers deep convolutional neural network architecture for breast cancer detection using clinical image data. In 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) (pp. 83-88). IEEE.

Kaba, Ş., Haci, H., Isin, A., Ilhan, A., & Conkbayir, C. (2023). The application of deep learning for the segmentation and classification of coronary arteries. Diagnostics, 13(13), 2274.

Alalawi, L., & Budoff, M. J. (2022). Recent advances in coronary computed tomography angiogram: the ultimate tool for coronary artery disease. Current Atherosclerosis Reports, 24(7), 557-562.

Gaffney, L. P. (2024). Extracting Coronary Lesion Information from Angiogram Reports for Patient Screening Applications (Doctoral dissertation, Massachusetts Institute of Technology).

Yuan, M., & Zhang, J. (2024). Image and Clinical Characteristics of the Right Coronary Artery Originating From the Left Coronary Sinus: A Database Review. Cardiology in Review, 10-1097.

Li, F., Chen, Y., & Xu, H. (2024). Coronary heart disease prediction based on hybrid deep learning. Review of Scientific Instruments, 95(1).

Zhou, L., Liu, H., Bae, J., He, J., Samaras, D., & Prasanna, P. (2023, April). Self pre-training with masked autoencoders for medical image classification and segmentation. In 2023 IEEE 20th international symposium on biomedical imaging (ISBI) (pp. 1-6). IEEE.

Omkari, D. Y., & Shaik, K. (2024). An integrated two-layered voting (tlv) framework for coronary artery disease prediction using machine learning classifiers. IEEE access, 12, 56275-56290.

Muscogiuri, E., van Assen, M., Tessarin, G., Razavi, A. C., Schoebinger, M., Wels, M., ... & De Cecco, C. N. (2025). Clinical validation of a deep learning algorithm for automated coronary artery disease detection and classification using a heterogeneous multivendor coronary computed tomography angiography data set. Journal of thoracic imaging, 40(1), e0798.

Online ISSN 3007-3197
Print ISSN 3007-3189

An, J., Wang, Y., Cai, Q., Zhao, G., Dooper, S., Litjens, G., & Gao, Z. (2024). Transformer-based weakly supervised learning for whole slide lung cancer image classification. IEEE Journal of Biomedical and Health Informatics.

Kaba, Ş., Haci, H., Isin, A., Ilhan, A., & Conkbayir, C. (2023). The application of deep learning for the segmentation and classification of coronary arteries. Diagnostics, 13(13), 2274.