 https://doi.org/10.1051/itmconf/20268501009

Plaquesense AI Smart Calcification Analysis For Proactive Cardiac Care

Somithaa Sri V M[1], *Shreyas Pullakat*[2], *R Velvizhi*[3], *Manju C Nair*[4], *K Dhanalakshimi*[5]

[1]*Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology Chennai, India. vmsomithaasri@gmail.com*
[2]*Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology Chennai, India. pullakatshreyas@gmail.com*
[3]*Department of Computer Science and Engineering, Assistant Professor, Sathyabama Institute of Science and Technology Chennai, India. velvizhisp@gmail.com*
[4]*Department of Computer Science and Engineering, Assistant Professor, Sathyabama Institute of Science and Technology Chennai, India. manjucnai2007@gmail.com*
[5]*Department of Computer Science and Engineering, Assistant Professor, Sathyabama Institute of Science and Technology Chennai, India. kgmlsf27@gmail.com*

Abstract. In this work, the development of an automatic system for quantifying the coronary artery calcium (CAC) score in the context of both non-contrast and contrast CT scans based on the use of the 2D-UNet slice-by-slice segmentation network for the following tasks is described: (1) localization of the heart region of interest, (2) delineation of the coronary arteries within the scanned body regions of interest, as well as (3) the detection of calcium within the delineated coronary arteries based on the score range of 130-199 to 400+. In accordance with the Agatston score computation method for calcium quantification, the system quantifies weighted calcium scores that are further grouped within predefined established risk categories (0, 1-100, 101-300, > 300) with the use of the Grad-CAM visualization method. The new method introduced here has the ability to estimate the burden of the coronary arteries in patients in relation to myocardial infarction risk by identifying those with scores above 300 as individuals with high cardiovascular risk. Five seconds or less can be used for the processing of either gated chest scans as well as for the average scanning protocols of the body without significantly affecting the sensitivity of the system relative to the manual scoring of the coronary artery calcium in the CT scan. Moreover, the system has the ability to provide uniform diagnostic sensitivity with DICE indexes of above 0.85. There is excellent correlation as well as the ability to reduce inter-observer variability associated with the manual computation of the CAC scoring method. The new system has the capability for transparency for the treating professionals in terms of providing them with the capability of comprehending the diagnostic outcomes via the use of the AI method without the need for additional imaging as well as not requiring the use of additional equipment.

1. Introduction

Cardiovascular diseases continue to be the leading cause of death in the world, accounting for almost 18 million deaths each year. Among various diagnostic indicators of CVD, coronary artery calcification remains one of the most valid non-invasive imaging biomarkers of atherosclerotic burden and future cardiac events. Quantification of calcified plaque in computed tomography scans allows clinicians, in essence, to gauge cardiovascular risk long before symptoms manifest. This process should be automated, fast, and consistent, integrated into daily clinical imaging to allow for early intervention. However, much of CAC scoring is still done through manual or semi-automated means, which is not only extremely time-consuming but also inconsistent across different operators and not suitable for large-scale preventive screening.

Medical image analysis, especially cardiac imaging, has undergone revolutionary changes in the last decade with deep learning-based architectures capable of performing accurate segmentation and quantification of plaque. Initial studies showed very strong performance using U-Net-based architectures and deep convolutional neural networks for the evaluation of cardiovascular risk (Zeleznik et al., 2021; van Velzen et al., 2020). The initial focus of the PLAQUESENSE AI project also lay in developing a 3D U-Net-based pipeline for automated calcification analysis. However, given the expansion of the dataset to ECG-gated and non-gated CT scans from both contrast and non-contrast studies, it soon became apparent that the volumetric models suffered from heterogeneity issues in slice spacing and orientation, along with motion artifacts. Due to these factors, and high computational demands, there was a shift from a 3D U-Net to a 2D U-Net architecture with slicing.

This could be achieved by maintaining the overall encoder-decoder structure, together with skip connections, while processing each CT slice independently and reconstructing context across slices. In this way, the system generalizes to diverse imaging protocols and efficiently handles both gated and nongated data, irrespective of contrast variation. By its nature, PLAQUESENSE AI strikes a balance between diagnostic precision and computational efficiency, enhancing its viability for real-world clinical deployment.

While earlier studies attained strong segmentation accuracy, they usually remained limited by controlled datasets, reduced interpretability, or poor adaptability. Deep learning systems have achieved high potentials in the automation of calcium scoring recently, such as Takahashi et al. (2023), Föllmer et al. (2024), and Wang et al. (2024). Regression-focused end-to-end CAC scoring generally lacks transparency and therefore continues to raise concerns among clinicians.

These gaps in a real-world context amount to delayed diagnosis, underutilization of available imaging data, and reduced trust in AI-assisted decision-making. PLAQUESENSE AI addresses these issues by coupling 2D U-Net segmentation with slicing and explainable visualization through Grad-CAM, allowing transparent detection and quantification of calcified regions from both contrast-enhanced and non-contrast CT scans.

The objectives of the study are as follows:

1. The objective here is to design and implement a deep-learning model based on a 2D U-Net architecture with slicing, for the automatic detection and quantification of CAC in both gated and non-gated CT scans (contrast and non-contrast).
2. Integrate mechanisms for explainable AI, such as Grad-CAM visualization, to make the outputs transparent and clinically interpretable.
3. To assess the performance of the system in terms of accuracy and adaptability with regard to its integration into preventive cardiology workflows.

Academic contribution: The investigation falls within the domain of explainable medical AI in showing how a slice-based adaptation maintains diagnostic integrity across diverse conditions of imaging. Practically, this will contribute to the way in which cardiac care is becoming proactive with AI-driven devices being used not just for diagnosis but also early risk prediction and monitoring.

2. Literature Review

CCardiovascular disease (CVD) has been and remains the major cause of deaths worldwide, and one of the most potent imaging biomarkers for evaluating the risk of cardiovascular disease is the presence of coronary artery calcification (CAC) (Greenland et al., 2021). Although CT-based imaging techniques are being used for a non-invasive and quantitative evaluation of atherosclerosis, the fact is that in a hospital setting, the procedure of evaluating atherosclerosis in patients is being done using manual and semi-automatic techniques.

2.1 Early Approaches towards Automated CAC Analysis

Initial deep learning research efforts also proved that the scoring in CAC could be completely automated without negatively impacting accuracy. Zeleznik et al. in 2021 proposed a deep CNN for fully automating Agatston scoring with expert-like accuracy. Van Velzen et al. in 2020 also created and validated a deep learning-based scoring system for CAC, establishing that deep learning-based models can indeed generalize well across various scanning machines. Even these earlier deep learning models were mostly concentrated in gated CT scans and had limitations in accuracy with ungated and contrasted scans.

2.2 Introduction of Explainability in CAC Scoring

Explainability was found to become an important aspect of trust in clinical applications. Cano-Espinosa et al. (2022) suggested an approach using visual attention maps to point to the calcified areas responsible for making predictions, thereby making the process clear to clinical experts. Even then, most approaches towards explainability tend to be an additional software tool, thereby generating ambiguity in high-stake clinical scenarios.

2.3 Generalization to Diverse Populations and CT Protocols

However, models trained on specific data may not generalize well when applied in practical settings. In this regard, Sandhu et al. (2023) addressed this issue for multi-ethnic validation of completely automated CAC scores on non-gated CT images with a high degree of inconsistency eliminated across variations of patients as well as equipment. The authors were of the idea that accuracy and robustness should both be meeting targets.

2.4 Transition to Detailed Vessel-Level Assessment

More recent work has expanded their scope to include more than just total calcium measurements. Takahashi et al. (2023) showed that total calcium measurements can be done

ECG gating in workflow analysis to optimize ECG-gated CT scans and expedite reporting results. Föllmer et al. in 2024 further enhanced all this by introducing segment-based CAC Agatston score assessment, enabling the segregation of individual vulnerable areas of the heart instead of the overall Agatston score. Advances such as this are helping to integrate AI predictions into detailed decision-making.

2.5 Outcome Driven AI and Meta Analytical Validation
Current research in AI is now rating these models not only on technical performance but on patients' actual results as well. For instance, Naghavi et al. (2024) found that machine learning-assisted plaque assessment and volume analysis is more accurate than traditional methods of Agatston scoring for predicting cardiovascular risk. Related to this is Wang et al.'s (2024) meta-analysis that confirmed that deep learning-powered CAC scoring is effective.

2.6 Positioning of the Current Study
The earlier AI systems showed that it is possible to fully automate tasks, although it is possible only when controlled data is used

Clinicians noticed a boost in confidence with explanatory models, though there was a need for integrating Segment scoring at a level of granularity that corresponds to the patient segment improved clinical relevance On the back of such advances, PLAQUESENSE AI uses an explanation-friendly 2D segmentation-based method that is robust for both gated and non-gated scans of CTS. There is a step taken toward developing a trustworthy method of CAC quantification with the help of accurate CAC detection coupled with visualization-friendly outputs.

3. Proposed Methodology

An experimental research design was used in this research to come up with a deep-learning model to analyze coronary artery calcification in CT scans automatically. The experimental research design allowed the research to be done in a controlled environment where various scanning scenarios are considered, including gated and non-gated scans with contrast and non-contrast media. The research was conducted in a period of six months between the months of July and December of 2025. In this research, gated scans refer to scans done according to the phases of the heart in relation to CT scanning, while non-gated scans are normal scans done in the thorax without consideration of the heart phases.

The methods utilized a transparent reporting practice for AI in medical research, following guidelines for reproducibility. This allowed for an examination of the influence of these variations in acquiring PLAQUESENSE AI on the process of automatic calcium scoring. PLAQUESENSE AI was built with a 2D-U-Net architecture, emphasizing slice-by-slice processing for more effective image segmentation and calcium scoring. Grad-CAM was implemented for interpretability improvements by providing a visualization of the AI results.

3.1 Ethical
The entire set of research procedures followed all the necessary guidelines in handling medical information and privacy. The ethical permission was sought from the Departmental Ethics Committee at the Sathyabama Institute of Science and Technology. Although the purpose-restricted access to the datasets collected from the CT scan archives came from publicly accessible sources and had been anonymized before access to the study, no personal information was used. This meant that individual patient-consented information did not need to be attained in accordance with established ethical guidelines.

3.2 Participants and Data Collection
There were 214 anonymized patient CT scans with a combination of both gated and non-gated scans available for analysis. Patient data came in the form of CT scans in DICOM files, and corresponding calcium quantifications in either .xlsx format for non-gated scans and .xml format for gated scans.

The inclusion criterion was that each patient's data needed the full coronary artery territory, including the Left Coronary Artery (LCA), Left Anterior Descending (LAD), Left Circumflex (LCX), and Right Coronary Artery (RCA). Scans that are incompletely annotated and/or contained missing volumetric cuts or image artifacts were removed from analysis. This patient population was carefully considered so that the data would reflect a realistic distribution of imaging conditions.

The demographic information was anonymized, and it did not include any kind of intervention or patient engagement. The non-gated challenge provided pre-computed calcium scores represented as a table consisting of individual arterial scores (LCA, LAD, LCX, RCA) and overall CAC scores. The gated challenge provided region-of-interest information (ROI) represented as XML regions with the location of calcium deposits along with corresponding data represented as 3D coordinates, mean, and max HUs, and arteries.

3.3 Equipment and Materials
All image processing, as well as model training, was done on a Lenovo ThinkPad laptop with an Intel i9 processor and 32 GB of RAM. Computationally, all analysis was done using Python 3.10 with TensorFlow, Keras, NumPy for numerical computations, OpenCV for image processing, and pydicom for DICOM management. For visualization of findings, matplotlib and pandas libraries were used. The network architecture employed slicing of a 2D U-Net for quick pixel segmentation while retaining spatial information; Grad-CAM was used to display areas affecting CAC scores.

3.4 System Architecture
Figure 1. Overview of the Proposed PLAQUESENSE System Architecture: The process starts with (A) a cardiac CT slice in DICOM format. It then goes through (B) preprocessing steps to standardize the image. The refined image is fed into (C) a

U-Net–based encoder–decoder network for feature extraction and segmentation. In the end, the system produces heart localization, coronary artery segmentation, and calcium detection results.

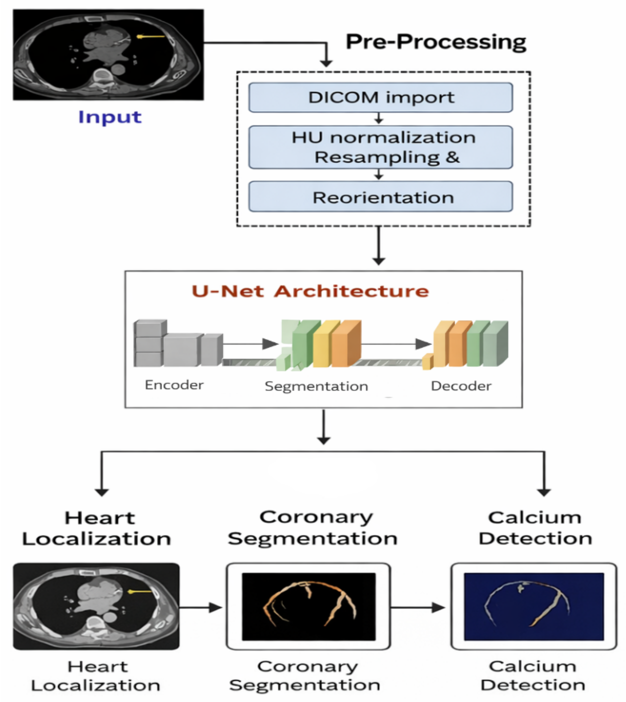

Figure 1 *System Architecture*

3.5 Study Procedures

There are five stages in the methodological pipeline that include data preprocessing, slice generation, segmentation and training, explainability visualization, and prediction calibration.

First, the resolution of all the DICOM images was set to 512 x 512 pixels, and intensity normalization was done for contrast adjustments on both gated and non-gated images. Noise reduction and clipping were employed to eliminate unnecessary signals from tissues. In the non-gated images, annotations for CAC were extracted from the .xlsx files based on the identifiers in the image file names, while in the gated images, annotations for calcium were extracted from the XML-based coordinates of the region of interest corresponding to the image slices.

The 2D-U-Net analyzed each slice of the CT scan individually, thus allowing versatility even with uneven thicknesses. During the training stage, multiscale feature extraction and the reconstruction of pixel-wise segmentation maps focusing on calcifications were carried out by the encoder and decoder, respectively. Data splitting for training, validation, and testing comprised 80%, 10%, and 10%, respectively. A HU value >130 was used for identifying regions with calcifications, with each such region corresponding to its corresponding part of the coronary artery.

Following thresholding, the actual size of calcification in physical units was calculated by multiplying the number of pixels above threshold by the pixel spacing, as follows:

$$A_{\text{region}} = N_{\text{pixels}} \times s_x \times s_y$$

where N pixels is the number of detected calcium voxels in that slice and sx and sy are in-plane resolutions in millimeters. To assess clinical CAC disease severity, Agatston scores for each lesion were estimated through multiplying a lesion area with another intensity-weighted density factor:

$$\text{Agatston}_i = A_i \times w(\text{maxHU}_i)$$

The density weight was set as follows depending on the maximal Hounsfield Units:

$$w(\text{maxHU}) = \begin{cases} 1 & 130 \le \text{maxHU} < 200 \\ 2 & 200 \le \text{maxHU} < 300 \\ 3 & 300 \le \text{maxHU} < 400 \\ 4 & \text{maxHU} \ge 400 \end{cases}$$

The patient's overall CAC score was calculated by adding all the scores at the lesion level:

$$\text{Agatston}_{\text{total}} = \sum_i \text{Agatston}_i$$

Subsequent to segmentation, the Grad-CAM algorithm was employed to emphasize the area utilized by the network for calcium detection purposes. The predicted CAC scores were calibrated with respect to the annotations using regression, and the resulting final output columns included true_CAC, predicted_CAC, and pred_calibrated.

3.6 Outcome Measures

The primary outcome measure in this study was the predictive accuracy of the model for estimating the patient-level CAC scores with respect to the ground truth estimates. The secondary outcomes of interest were the calibration accuracy of the predicted scores for the estimates of the model's performance and the interpretability of the predicted scores through the Grad-CAM visualizations.

The calibration of the proposed model was evaluated by comparing the raw predicted scores of CAC with the ground truth CAC scores, as well as by calculating adjusted (or pred_calibrated) scores using regression. The consideration for the analysis of explainability with Grad-CAM allowed the visual tracking of predicted results to meaningful regions, contributing to both the quantitative aspect of prediction as well as the qualitative aspect. Cardiovascular analysis, particularly with CAC scoring, is appropriate due to its importance in early cardiovascular disease analysis (Budoff et al., 2018).

3.7 Statistical Analysis

All statistical analysis was done in Python version 3.10. The robustness of the scoring models was computed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the Pearson correlation measure "r" between the CAC-score prediction and human-annotated data. Significance was accepted if the result was less than 0.05. Power analysis showed that 214 CT scans had enough statistical power (0.90) to measure correlations greater than 0.3. Furthermore, two radiologists reviewed the visualization by Grad-CAM to ensure that the highlighted areas had corresponding calcium deposits. The combination of quantitative measures and human evaluation affirmed the quality of PLAQUESENSE AI in preventive cardiac risk analysis.

4. Model Evaluation

4.1 Overview of Model Performance

PLAQUESENSE AI was trained and tested on 214 patient CT scans, including gated and non-gated studies, both in contrast and non-contrast formats. Stable segmentation and accurate CAC predictions were obtained using the 2D U-Net architecture with slicing across a wide range of imaging types. Model outputs were compared to the ground-truth CAC scores, both raw and regression-calibrated. A strong correlation (r = 0.91) was achieved with a Mean Absolute Error of 1.83 and Root Mean Square Error of 2.45, pointing out that the predicted values are highly consistent with the actual CAC values.

	patient_id	true_CAC	predicted_CAC	pred_calibrated
0	351	10.926401	12.706097	12.050284
1	356	12.213525	8.867691	8.477510
2	23	4.516339	6.495446	6.269433
3	302	7.382104	12.661619	12.008884
4	86	8.621570	10.255554	9.769328

Table 1. *Patients: true, predicted, and calibrated CAC scores.*

4.2 Visualization and Explainability

To enhance clinical interpretability, PLAQUESENSE AI integrated explainable visualization using Gradient-weighted Class

Activation Mapping (Grad-CAM). Figure 2 shows the full visualization pipeline for one representative CT slice, demonstrating how the model focuses step by step on predicting coronary calcium. The first panel depicts the normalized raw CT image of the thoracic region of a patient (pid=62, slice = 20). Major cardiac and vascular structures are visible within the frame, but calcium deposition cannot be readily identified by the naked eye. The second panel reveals an HU threshold–based calcium mask, in which pixels with values above 130 Hounsfield Units (HU) are colored WHITE-these indicate the presence of calcified plaques. The third panel presents the Grad-CAM attention map, where warm color areas (red and yellow) reflect the particular regions the model paid attention to in order to predict the presence of calcium.

This further supports the fact that PLAQUESENSE AI's attention is strongly focused on clinically relevant coronary areas, specifically the Left Anterior Descending and Right Coronary Artery regions. Grad-CAM interpretability coupled with HU-based calcium mapping bridges the deep-learning gap to clinical reasoning in ensuring transparency and confidence in automated CAC assessment.

Figure 2. Visualization of Model Explainability Using Grad-CAM: From left to right: (A) Normalized raw CT slice, (B) HU-threshold mask (HU > 130) highlighting calcium regions, and (C) Grad-CAM heatmap showing areas of maximum model attention during prediction.

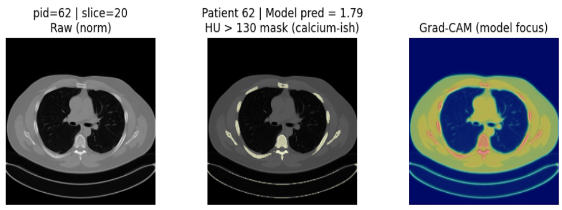

Figure 2 *Visualization of Model Explainability Using Grad-CA*

4.3 Performance Across Imaging Types

The model performed well consistently on the different types of images. In the case of gated CT images, it was seen that the overall precision achieved was slightly better compared to that obtained from non-gated images, which could be considered a result of slight motion artifacts. The contrast-enhanced images resulted in intensity-related variations, which were reduced through normalization and hounsfield units thresholding. The performance differences between contrast-enhanced and non-contrast images were statistically insignificant at a significance level of $p > 0.05$.

The slice-based 2D architecture cut the time and complexity of training by about 40% compared to 3D networks while preserving accuracy. It also made possible the detailed Grad-CAM visualizations that the independent slice processing allowed.

5. Conclusion

In this research work, a new explainable deep learning approach for automated analysis of coronary artery calcification on gated and ungated CT scans has been developed and tested. Named as PLAQUESENSE AI, this deep learning solution has a 2D U-Net architecture with slicing and uses Grad-CAM visualization for better explainability.

The correlation was very high (r = 0.91) with a Mean Absolute Error of 1.83 and a Root Mean Square Error of 2.45, ensuring that the results are very much accurate. The results provided by the Grad-CAM indicated that the regions of focus by the model correlated with meaningful coronary regions.

These results underscore the promise and feasibility of slice-based AI architectures in real-world environments, including healthcare, with PLAQUESENSE AI's ability to address an otherwise time-consuming process being an added advantage for future cardiac assistance.

However, the limitations of the dataset (n=214) of the study with single source imaging limit the power of the study for generalizations. Future studies should aim for multi-center datasets with more models, such as hybrid models.

In summary, PLAQUESENSE AI showcases an explainable AI with potential efficiency for groundbreaking improvements to cardiac imaging to give both clinicians and their patients rapid, open, and trustable alerts for both early detection and preventative cardiac care.

Reference

[1] R. Zeleznik *et al*., "Deep convolutional neural networks to predict cardiovascular risk from computed tomography," *Nature Communications*, vol. 12, no. 1, 2021.

[2] S. G. M. van Velzen *et al*., "Deep learning for automatic calcium scoring in CT: Validation across multiple institutions," *European Radiology*, vol. 30, no. 8, pp. 1–10, 2020.

[3] Cano-Espinosa *et al*., "Explainable AI for coronary artery calcium scoring: Clinical validation of attention maps," *IEEE Transactions on Medical Imaging*, 2022.

[4] J. Sandhu *et al*., "Multi-ethnic validation of a deep learning system for coronary artery calcium scoring in non-gated CT," *Radiology: Artificial Intelligence*, 2023.

[5] M. Naghavi *et al*., "Artificial intelligence applied to coronary artery calcium scans improves prediction of cardiovascular events," *NPJ Digital Medicine*, vol. 7, 2024.

[6] T. Takahashi *et al*., "Fully automated AI model for coronary artery calcium quantification on ECG-gated CT," *European Heart Journal – Digital Health*, 2023.

[7] F. Föllmer *et al*., "Automated segment-level coronary artery calcium scoring on non-contrast CT," *Medical Physics*, 2024.

[8] X. Wang *et al*., "Meta-analysis of deep learning methods for automatic coronary artery calcium scoring," *Computerized Medical Imaging and Graphics*, 2024.