Opportunistic Screening for Thoracic Aortic Calcification on Non-Dedicated CT via Deep Learning

Catharine De Vita Graves[1], Antonildes N. Assunção-Jr[1], Gabriela Liberato de Sousa[1], Roberto N. Dantas-Jr[1], Carla F. G. Da Silva[1], Cesar H. Nomura[1], Marco A. Gutierrez[1]
[1] Instituto do Coração HCFMUSP, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brazil.

Abstract

Thoracic Aortic Calcification (TAC) is a significant predictor of cardiovascular events, but it's often overlooked in routine CT scans. This study explores the use of these non-dedicated scans for opportunistic TAC detection and quantification.

Our model was trained and validated on a dataset comprising 661 chest CT exams retrospectively collected from patients undergoing imaging for clinical indications. TAC regions were annotated by experts to serve as ground truth.

We developed a deep learning model based on an enhanced 3D U-Net with residual blocks to automatically segment TAC and predict the Agatston calcium score. Using expert annotations as a reference, our model achieved strong performance, with a Pearson's correlation coefficient (ρ) of 0.87 ± 0.06 and a coefficient of determination (R^2) of 0.76 ± 0.11.

These results demonstrate that deep learning can reliably assess TAC from standard CT scans, turning incidental findings into actionable clinical insights for proactive cardiovascular risk assessment.

1. Introduction

Thoracic Aortic Calcification (TAC) is a marker of atherosclerosis, characterized by the deposit of calcium in the thoracic aorta wall [1]. It serves as a predictor of adverse cardiovascular outcomes, including stroke [1], [2]. TAC can be detected incidentally in computed tomography (CT) scans performed for other clinical purposes, such as pulmonary evaluation or coronary assessment, rather than being the primary reason for imaging. its presence and extent are often underreported during routine radiological assessments of chest or abdominal computed tomography (CT) scans performed for non-cardiac reasons [2].

An automated tool capable of quantifying TAC offers an opportunity for early cardiovascular risk assessment and intervention. Accurate and automated detection of TAC can turn incidental findings into actionable insights, allowing clinicians to identify high-risk individuals without needing additional dedicated cardiac imaging.

Deep learning is an artificial intelligence method that learns from data and can be used to solve tasks automatically. Many studies have been published using deep learning to detect TAC in CT scans [3]–[7].

Research in this area explores various approaches; some concentrate solely on quantifying aortic calcium as a cardiovascular risk indicator [3], [4], while others include multiple cardiac structures such as coronary arteries and heart valves [5]–[7].

In this study, we examine the potential of using routinely acquired, non-dedicated CT scans for opportunistic detection and quantification of TAC. We developed a deep learning model based on the 3D U-Net architecture enhanced with residual blocks to automatically segment TAC and predict the Agatston calcium score. Our approach enables reliable quantification of calcification from standard CT scans, demonstrating the feasibility of integrating deep learning-based TAC evaluation into routine imaging workflows.

2. Methodology

This study used CT scans routinely collected at the Heart Institute (INCOR), located in São Paulo, Brazil, as part of a clinical research protocol approved by the Institutional Review Board (IRB) under #5970/24/181. INCOR is a specialized institute focuses on cardiac and pulmonary diseases. We collected 661 chest CT exams with arterial calcification acquired clinically from 2013 to 2023 for multiple purposes. Each exam was annotated by specialists, indicating the calcium segmentation and the Agatston score. The dataset includes images from patients aged 19 to 96, with a slight female majority (55%).

Each exam was divided into windows measuring 256x256x32 pixels. We segmented the mediastinum using the methodology described in [8]. We selected at least seven windows around the heart, using the mediastinum center as a reference. Then we selected enough windows to cover the entire mediastinum. Only windows with aorta calcifications were used during the training phase; in the test phase, all windows were used to perform the inference, and the test performance was evaluated using the entire exam.

We used five-fold cross-validation for model evaluation. The dataset was split into five equal parts

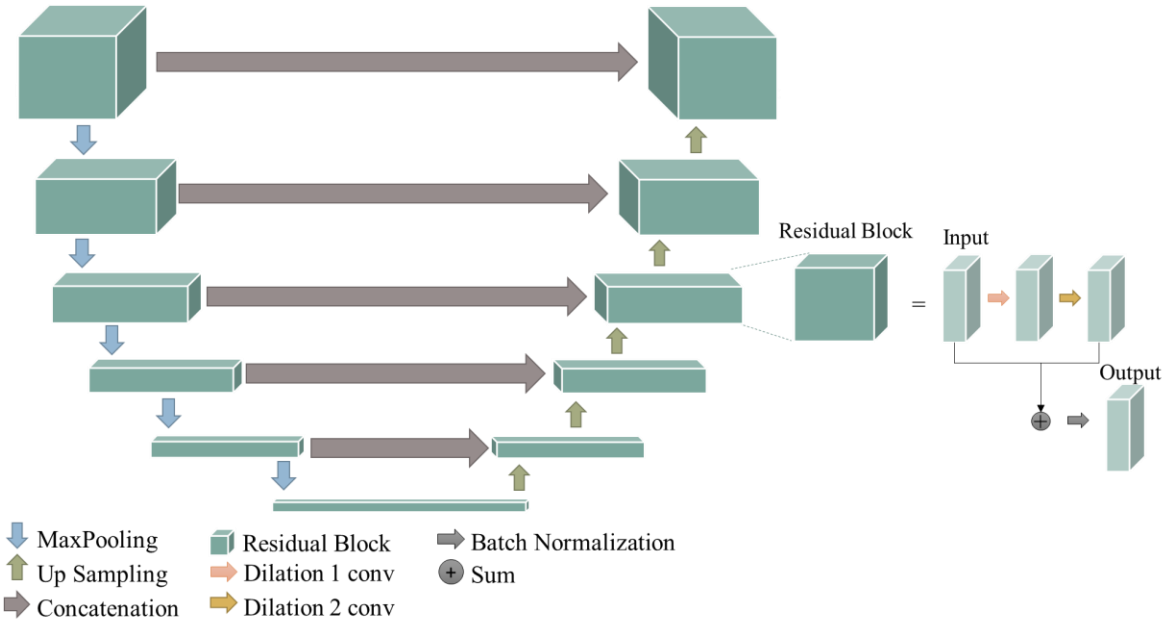

Figure 1: Proposed ResUnet model architecture.

(folds). In each of the five iterations, one fold was designated as the test set, while a different fold was used for validation during training. The remaining three folds were used to train the model.

We developed a model based on the U shape of the Unet [9] with residual blocks [10], schematized in Figure 1. The proposed ResUnet architecture consists of five downsampling levels and five upsampling levels. The encoder and decoder are connected at each level to reduce the semantic gap and enhance learning convergence. Each level features a residual block with dilated convolutions to expand the receptive field. The number of filters in each convolutional layer varies depending on the level, starting at 16 in the first encoder level and doubling at each subsequent level. The final layer uses a softmax activation function to produce the segmentation output. We compared the proposed model (ResUnet) with the U-Net 3D as a baseline, adapted from [9] to perform 3D process.

It is important to notice that the proposed model is composed of fewer parameters compared to the Unet3D, as can be seen in Table 1.

Table 1: Number of parameters in each model.

Model	Parameters
ResUnet	18,289,396
Unet3D	90,311,300

The models were designed to segment calcifications within the aorta and subsequently calculate the Agatston calcium score using equation 1.

$$\boldsymbol{Agatston\ score} = \sum_{i}^{N} \boldsymbol{A_i \alpha_i} \qquad (1)$$

Where A is the lesion area, α is a weight factor that depends on the lesion's maximum Hounsfield Unit (HU), and N is the number of lesions in the exam [11].

Performance was evaluated using Pearson's correlation coefficient (ρ) and the coefficient of determination (R^2), comparing the calculated Agatston scores obtained from the models' outputs with those of the expert radiologist annotations as groud-truth labels.

3. Results

In Table 2, the comparison of R^2 and ρ between the calcium score annotated by the specialist and those obtained using the analysed models, ResUnet and Unet3D, are presented. The metrics are shown with their averages and standard deviations across the test folders.

Table 2: Comparison of ResUnet and Unet3D performance, using R^2 and Pearson's correlation coefficient (ρ).

Model	R^2	ρ
ResUnet	**0.76 ± 0.11**	**0.87 ± 0.06**
Unet3D	0.71 ± 0.13	0.84 ± 0.08

Figures 2 and 3 present scatter plots comparing the true calcium scores (annotated by a specialist) with the scores predicted by the ResUnet and Unet3D models, respectively. The plots display results from the first fold only. To improve the visualization of small calcium score values, the figures are shown on a logarithmic scale. Zero

scores were adjusted to 0.1 to allow proper representation on the logarithmic scale

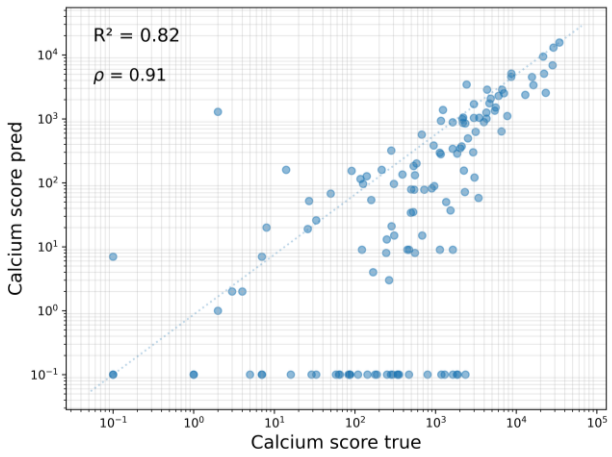

Figure 2: Log scale scatter plot between the true calcium score and the predicted by the ResUnet model.

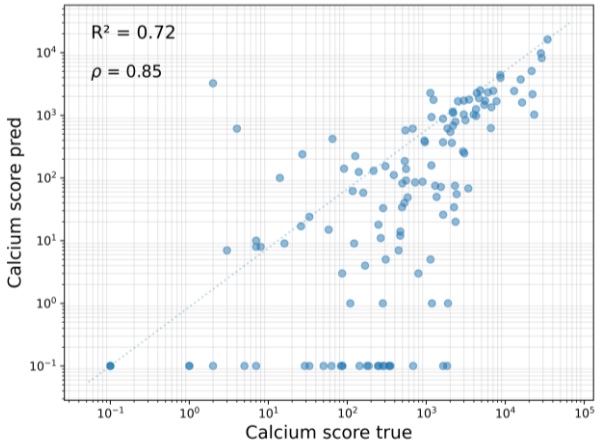

Figure 3: Log scale scatter plot between the true calcium score and the predicted by the Unet3D model.

3. Discussion

The proposed model demonstrates superior performance compared to the Unet3D architecture, as evidenced by the average among the folders presented on Table 2, $R^2 = 0.76 \pm 0.11$ vs 0.71 ± 0.13 and $\rho=0.87 \pm 0.06$ vs 0.84 ± 0.08.

The scatter plot for the first fold obtained with the ResUnet (Figure 2), compared with the one obtained with the Unet3D (Figure 3), shows less data dispersion and higher coefficients of determination (R^2=0.82 vs. 0.72) and correlation (ρ=0.91 vs. 0.85). However, both models occasionally produce false negatives, reporting zero calcium where it exists. A potential source of this error could be the method used to select the windows around the mediastinum. Since the aorta can occupy a large space around the mediastinum, this approach may exclude significant areas of the examination, which could affect the model's accuracy.

Our proposed model is significantly more efficient than the Unet3D, with nearly five times fewer parameters, as shown in Table 1. This reduction is crucial for two reasons: it lowers the computational resources required for training and inference in routine clinical settings, and it enhances the model's ability to generalize to new image distributions that may differ slightly from the training data. This improved generalization capability reduces the risk of overfitting.

4. Conclusion

This work introduces a compact and efficient deep learning model for the automated segmentation and quantification of the thoracic aortic calcium score in routinely acquired images.

These findings support the feasibility of applying deep learning for reliable, automated assessment of TAC in routine CT exams, without requiring dedicated cardiac imaging. By transforming incidental data into clinically relevant information, this approach enables proactive cardiovascular risk assessment, with the potential to improve patient outcomes and enhance healthcare efficiency.

Our study highlights the transformative role of artificial intelligence in extracting unexploited clinical value from existing medical imaging, supporting a shift toward more preventive, scalable, and data-driven cardiovascular care.

Acknowledgments

This work was supported by Canon Medical System do Brasil Ltda. and Zerbini Foundation as part of the research project “Automatic Quantification of Prognostic Biomarkers in Cardiovascular Disease Using Deep Learning on Chest Computed Tomography Images”.

References

[1] D. M. Hermann *et al.*, “Thoracic aortic calcification is associated with incident stroke in the general population in addition to established risk factors,” *Eur. Hear. J. - Cardiovasc. Imaging*, vol. 16, no. 6, pp. 684–690, Jun. 2015, doi: 10.1093/ehjci/jeu293.

[2] Y. Brodov *et al.*, “Calcium score of the entire thoracic aorta is an independent predictor of all-

cause mortality in patients referred to chest computed tomography," *Int. J. Cardiovasc. Imaging*, vol. 40, no. 1, pp. 177–183, Oct. 2023, doi: 10.1007/s10554-023-02971-2.

[3] P. M. Graffy, J. Liu, S. O'Connor, R. M. Summers, and P. J. Pickhardt, "Automated segmentation and quantification of aortic calcification at abdominal CT: application of a deep learning-based algorithm to a longitudinal screening cohort," *Abdom. Radiol.*, vol. 44, no. 8, pp. 2921–2928, Aug. 2019, doi: 10.1007/s00261-019-02014-2.

[4] F. N. Guilenea, M. E. Casciaro, A. F. Pascaner, G. Soulat, E. Mousseaux, and D. Craiem, "Thoracic aorta calcium detection and quantification using convolutional neural networks in a large cohort of intermediate-risk patients," *Tomography*, vol. 7, no. 4, pp. 636–649, Dec. 2021, doi: 10.3390/tomography7040054.

[5] N. Lessmann *et al.*, "Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks with Dilated Convolutions," *IEEE Trans. Med. Imaging*, vol. 37, no. 2, pp. 615–625, Feb. 2018, doi: 10.1109/TMI.2017.2769839.

[6] G. Singh *et al.*, "End-to-end, pixel-wise vessel-specific coronary and aortic calcium detection and scoring using deep learning," *Diagnostics*, vol. 11, no. 2, Feb. 2021, doi: 10.3390/diagnostics11020215.

[7] S. G. M. M. van Velzen *et al.*, "Deep Learning for Automatic Calcium Scoring in CT: Validation Using Multiple Cardiac CT and Chest CT Protocols," *Radiology*, vol. 295, no. 1, pp. 66–79, Apr. 2020, doi: 10.1148/radiol.2020191621.

[8] C. V. Graves *et al.*, "Enhancing calcium score quantification in cardiac images through robust mediastinum segmentation," in *Medical Imaging 2024: Computer-Aided Diagnosis*, Apr. 2024, p. 22, doi: 10.1117/12.3006910.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *MICCAI2015*, 2015, pp. 234–241.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[11] A. S. Agatston, W. R. Janowitz, F. J. Hildner, N. R. Zusmer, M. Viamonte, and R. Detrano, "Quantification of coronary artery calcium using ultrafast computed tomography.," *J. Am. Coll. Cardiol.*, vol. 15, no. 4, pp. 827–832, 1990, doi: 10.1016/0735-1097(90)90282-T.