Multimodal Deep Learning Framework for Automatic Atherosclerosis Risk Classification, Assessment, and Monitoring

Catania Cecere Eleonora
Politecnico di Torino
Turin, Italy
327305

Cevoli Elisa
Politecnico di Torino
Turin, Italy
330775

Cordoba Acosta Tatiana Camila
Politecnico di Torino
Turin, Italy
320757

Ottomano Claudia
Politecnico di Torino
Turin, Italy
331014

***Abstract*— This study presents a multimodal framework for assessing and monitoring atherosclerosis by integrating deep learning and statistical models with imaging, physiological signals, and traditional risk factors. The proposed approach enhances early detection, risk classification, and personalized treatment through AI-driven decision fusion and a biocybernetics feedback loop. Explainability and uncertainty quantification are incorporated to improve model transparency, reliability, and clinical trust in AI-driven prediction.**

Keywords—Atherosclerosis, CAC, IMT, c-PWV, Deep Learning, Multimodal Analysis, Risk Assesment, Explainability, Uncertainty Quantification.

I. INTRODUCTION

Atherosclerosis (AS) is a chronic, systemic, and immune-inflammatory disease that affects medium and large-caliber arteries. This condition leads to thickening of the arterial wall due to the accumulation of lipids, inflammatory cells, and fibrous tissue, causing a progressive reduction of the vascular lumen and impairment of blood flow. AS initially develops asymptomatically, then progresses to manifest with symptoms and clinical complications as the lesions advance. Atherogenesis begins with endothelial dysfunction, which increases the permeability of the arterial wall [1] and reduces nitric oxide production, compromising vasodilation. This promotes the accumulation of oxidized LDL, triggering inflammation and activation of monocytes and macrophages. Initially, fatty streaks form, followed by the proliferation of smooth muscle cells that generate a stabilizing fibrous cap. However, cellular apoptosis and the action of metalloproteinases can weaken the plaque, making it unstable and prone to rupture, leading to thrombus formation and potential vascular occlusion. The complications of AS vary in severity: initially, narrowing of the arterial lumen may cause chronic ischemia in affected regions or angina pectoris. If the plaque ruptures and forms an occlusive thrombus, more severe acute events such as myocardial infarction, ischemic stroke, or critical limb ischemia may occur, with the risk of tissue necrosis and amputation. Additionally, arterial stiffness and loss of elasticity can contribute to aneurysm formation, increasing the risk of vascular rupture. [2]

The main risk factors for the development of AS include advanced age, hypertension, hyperlipidemia, diabetes mellitus, obesity, smoking, and physical inactivity [3]. The World Health Organization reports that atherosclerosis is the leading cause of death worldwide, affecting both developed and developing countries. Cardiovascular diseases account for 32% of all deaths, with an estimated 17.9 million people dying each year.

Studies have shown that the risk of cardiovascular complications is associated with parameters such as intima-media thickness (IMT) [4], increased pulse wave velocity (PWV) [5], and the presence of arterial calcifications [6] — indicators of vascular stiffness and disease progression. Early diagnosis and continuous monitoring of AS are crucial for preventing severe complications. Common diagnostic tools include ultrasound (US) for IMT assessment, carotid-femoral PWV (cfPWV) measurement to assess the propagation time of the pulse wave between the carotid and femoral arteries, and computed tomography (CT) for coronary artery calcium (CAC) evaluation [2].

Emerging and promising techniques include the use of artificial intelligence to automatically extract diagnostic parameters from acquired images and signals. This approach helps support less experienced operators, improving diagnostic reliability and reducing operator variability, thereby contributing to a more standardized and precise assessment of disease progression.

The aim of this study is to develop an innovative framework for the assessment and monitoring of AS, integrating multimodal and multiscale approaches to improve early lesion detection, assess complication risk, and optimize therapeutic strategies. The final framework integrates the model into a biocybernetics loop to provide continuous feedback. Using uncertainty quantification (UQ) techniques and AI explainability (XAI), will help healthcare professionals tailor treatment strategies based on the principles of personalized and precise medicine.

II. DATA ACQUISITION AND RAPRESENTATION

The framework should include patients with significant risk factors such as age, family history, hypertension, dyslipidemia (elevated LDL, low HDL), diabetes, obesity, and smoking. Patients with prior cardiovascular events, like heart attacks or strokes, should also be included due to their higher risk of progression.

The cohort should consist of adults aged 40-85 years, with an emphasis on heterogeneity. The 40-75 age group is crucial for early detection of subclinical AS, while the 75-85 age group is important for monitoring the progression of existing cardiovascular issues. Monitoring should be longitudinal, including periodic exams with traditional risk factors, TC images, US images and PPG signals. The acquired data must be properly preprocessed and labelled by experts. This

approach enables early identification of at-risk patients and facilitates timely interventions to prevent complications, improving long-term prognosis with personalized treatment strategies.

A. *Traditional risk factors*

1) *MESA*

Cardiovascular risk factors are traditionally divided into modifiable and non-modifiable factors. Non-modifiable factors are those that cannot be directly influenced but can only be mitigated through preventive measures. These include advanced age, male sex for young individuals, and family history.

On the other hand, modifiable risk factors are those influenced by lifestyle and include hypertension, which can damage arteries; high LDL and low HDL levels, particularly in combination with hypertension; diabetes, as it causes cardiovascular complications; high body mass index and waist circumference; and smoking habits. [7].

The Multi-Ethnic Study of Atherosclerosis (MESA) is a prospective study aimed at better understanding subclinical cardiovascular diseases and their progression in a multi-ethnic population. It has led to the development of a risk score that considers the multi-ethnic nature of the studied population for coronary heart disease (CHD). This score integrates traditional risk factors with the CAC score.

The MESA Risk Score estimates the 10-year risk of coronary events, such as myocardial infarction, cardiac arrest, confirmed angina requiring revascularization, and death from CHD. The MESA Risk Score calculator is available online and can be used to assist clinicians in communicating risk to patients and determining risk-based therapeutic strategies [8]

B. *Imaging*

1) *Computed Tomography (CT)*

The CAC score is a quantitative measure of calcifications present in the coronary arteries. These calcifications are indicators of subclinical AS, meaning the presence of atherosclerotic plaques that have not yet become symptomatic. Coronary calcium quantification is performed through a non-contrast CT scan. This technique has the advantage of providing good spatial resolution and fast execution, but it requires exposure to ionizing radiation, is sensitive to motion artifacts, and has a high cost. [9]

The most used method for calculating the score is the Agatston method, which considers the extent and density of the calcifications, identifying coronary calcifications as voxels with a Hounsfield unit (HU) value greater than or equal to 130 [10].

2) *Ultrasound (US)*

An increase in IMT is closely linked to AS, making it a valuable indicator of atherosclerotic risk [4]. US is the primary method for detecting carotid stenosis and assessing lesion severity due to its availability, non-invasiveness, cost-effectiveness, and lack of ionizing radiation [11]. Images are acquired using a transducer emitting high-frequency sound waves, which reflect off body structures and are detected by the transducer. This provides real-time imaging of the common carotid artery (CCA) transverse sections for stenosis analysis. For accurate measurements, the ultrasound beam must be perpendicular to the adventitia. While US is rapid and non-invasive, it requires skilled operators to ensure high-quality results.

C. *Signals*

1) *Photoplethysmography (PPG)*

Arterial stiffness is a key risk factor for CVD [12], and cf-PWV is a key parameter used to assess it [5]. To non-invasively estimate cf-PWV, we used PPG signals [13], a cost-effective technique that detects blood volume changes through light, making it particularly valuable for studying arterial stiffness.

PPG signals were collected from digital, radial, and brachial arteries and transformed into spectrograms using the Short-Time Fourier Transform (STFT). This method divides temporal signals into overlapping segments, analyzing their frequency content over time and producing detailed time-frequency representations. To ensure consistency and compatibility with the model proposed in the task, PPG signals were processed using 76 samples per segment, a Hamming window, and zero padding. The resulting spectrograms were resized to a standard resolution of 224×224 pixels, where darker regions indicate higher frequency power and lighter regions indicate lower frequency power. This method is particularly valuable for studying arterial stiffness; however, it has some limitations: motion artifacts, fluctuating sensor contact, and optical disturbances.

III. MULTIMODALITY AND MULTISCALE INTEGRATION

Each of the parameters identified in the previous section has limitations in predicting cardiovascular risk. A more accurate risk estimation can be achieved by integrating all parameters together.

A. *MESA Score*

For the calculation of the MESA score, we have decided to rely on the official MESA calculator, a statistical model based on multivariate linear regression (MLR). The calculator provides a continuous percentage output representing the 10-year risk of developing CHD. The risk classification is discretized following Ridker et al.'s recommendations [14]: low (<5%), intermediate (6-20%), and high (>20%) risk.

MESA calculator requires traditional risk factors and the calcium score in Agatston units as inputs. To automate the calculation of the Agatston score and save time, a deep learning (DL) model has been identified that can compute this score directly from CT images.

B. *TC Model*

As proposed by Winkel DJ et al. [15] (Fig. 1), the study aims to develop a DL classification method for computing the Agatston score from CT scans by identifying and classifying CAC based on their location in the coronary arteries. It is structured into two main stages: (i) identification of voxels with HU $\geq$ 130 corresponding to CAC and (ii) assignment of branch labels (LM, LAD, CX, RCA) to the identified voxels, where each label corresponds to a specific coronary artery region. A U-Net model is employed to segment and extract the heart region from CT images (slice thickness $\leq$ 0.3 mm). The CT volume is cropped to the heart region, and a coronary territory map is aligned with the heart anatomy to assign prior probabilities for voxel classification. Voxels with

intensity ≥ 130 HU in coronary artery regions are selected as candidates. For each candidate voxel, a 32x32 image patch centered on the voxel is extracted, along with its spatial coordinates (x, y, z). These inputs, along with the local territory map, are processed by a DL model consisting of two components: i) CNN based on a ResNet architecture, which processes the 32x32 image patch and the coronary territory map using 3x3 convolutions, batch normalization, and ReLU activation; ii) Fully Connected Neural Network, which processes the voxel's spatial coordinates through four layers with batch normalization and ReLU activation. The model is trained using Stochastic Gradient Descent as the optimizer, with an initial learning rate of 10^3, dynamically adjusted based on validation loss. The outputs of the two networks are concatenated and passed through a sigmoid-activated neuron to compute the probability of the voxel being CAC. After classifying the voxels, the extent of calcifications and the Agatston score are determined by calculating the area of the segmented voxels. This approach effectively integrates spatial and anatomical features to improve the accuracy of CAC identification and classification.

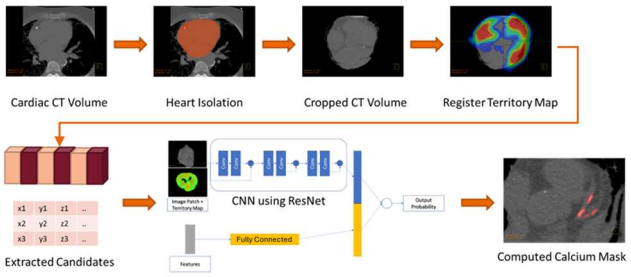

Fig. 1. Winkel DJ et al. DL pipeline for CAC identification, classification, and Agatston score computation from CT scans. It includes heart segmentation (U-Net), voxel extraction (HU intensity), and classification (ResNet CNN).

C. *US Model*

Following the approach by Liu M. et al. [16] (Fig. 2), US images were used as input for the CANet classification model. The model is based on a modified U-Net architecture, incorporating an EfficientNet encoder pretrained on the ImageNet dataset to enhance the extraction of high-dimensional features. The encoder is integrated with the decoder using skip connection, facilitating the integration of high-level semantic features with low-level features, thus improving segmentation accuracy. The Adam optimizer with an adaptive learning rate is used: the learning rate is reduced by a factor of 10 if the Dice coefficient on the validation set doesn't improve for 10 consecutive epochs. Input images are resized to 896×672 pixels and normalized using min-max scaling to maintain consistency in data processing. Furthermore, data augmentation methods are applied during training to mitigate overfitting and enhance the model's capacity to generalize new data. Upon completing the segmentation task, which identifies the area of the lumen-intima ($Area_{LI}$) and the area of the media-adventitia ($Area_{MA}$), following Mengmeng L. et al, the CCA stenosis severity rate is calculated (1). $Area^*_{MA}$ considers a 0.5 mm reduction on the outer edge of $Area_{MA}$, as recommended by the guidelines of the American Society of Echocardiography [17].

$$\text{Stenosis severety rate (\%)} = \left(1 - \frac{Area_{LI}}{Area^*_{MA}}\right) \times 100 \quad (1)$$

The stenosis can be classified as mild (0-49%), moderate (50-69%), or severe (70% or greater) based on the stenosis severity rate (%) [18].

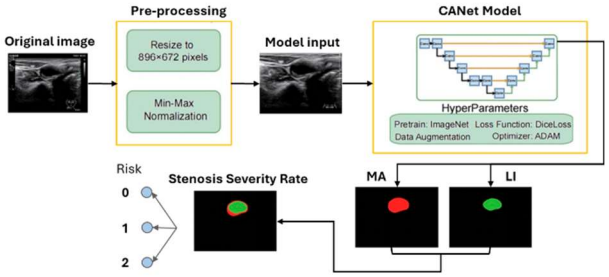

Fig. 2. CANet classification model: U-Net with an EfficientNet encoder.

D. *PPG Model*

Following the approach proposed by Abrisham et al. [13] (Fig. 3), PPG signals transformed into spectrograms were used as input for a DL classification model based on a modified pre-trained ResNet-18 architecture. This neural network, composed of 18 convolutional layers organized into five blocks, was adapted for regression tasks by replacing the softmax layer with a single neuron. To optimize model performance, specific hyperparameters were selected: a learning rate of 0.001, the Adam optimizer, 200 epochs, a patience value of 20, and a batch size of 64. Once cf-PWV was estimated, the following risk classification was applied: Low risk for cf-PWV < 10.7 m/s, indicating lower arterial stiffness and a relatively low risk of cardiovascular and all-cause mortality; Moderate risk for cf-PWV between 10.7 and 11.5 m/s, where the risk is elevated but not critical, and further clinical parameters may be considered for a more precise evaluation; and High risk for cf-PWV > 11.5 m/s, indicating significant arterial stiffness and a high risk of cardiovascular and all-cause mortality [19].

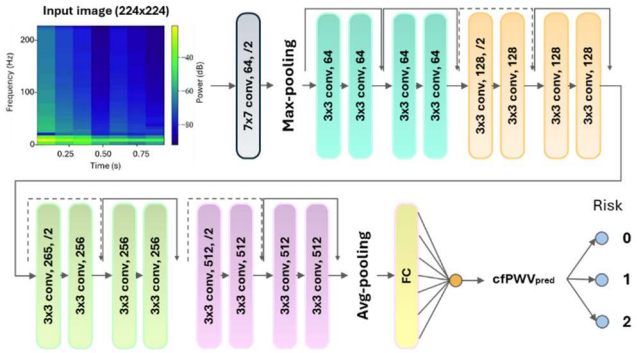

Fig. 3. DL classification model based on a modified pre-trained ResNet-18 architecture by K. P. Abrisham et al.

IV. Task Selection

A. *Decision Fusion*

The proposed framework aims to assess the risk of severe complications associated with AS using an advanced multimodal risk classification system, providing a valuable tool for preventive healthcare and personalized treatment planning.

The dataset, which includes traditional risk factors, medical images, and physiological signals, is split into training, validation, and test sets using stratified sampling with an 8:1:1 ratio. The training set is used to build the three proposed

DL models, while the validation set is used for hyperparameter tuning. The test set is used to evaluate the models' ability to generalize to unseen data and to quantify uncertainty.

For each model the per-class recall is computed to evaluate the accuracy of the predicted classes. The risk outputs from all models are combined using a decision fusion method to generate a single final risk classification.

To assess uncertainty in the DL models, we employ Monte Carlo (MC) Dropout, reactivating dropout layers during inference to perform multiple predictions. The prediction variance reflects the model's uncertainty and is combined with each model's accuracy to determine a weight for the final integration. The proposed model pipeline is reported in Fig. 4.

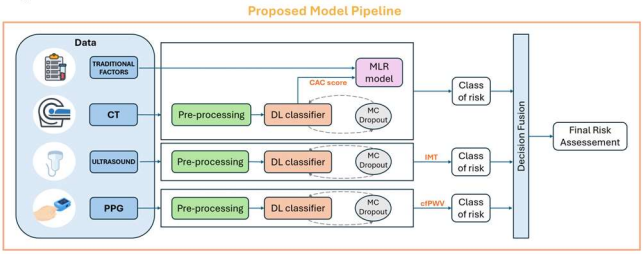

Fig. 4. Proposed model pipeline integrating automatic risk classification from medical images, physiological signals, and traditional risk factors. Each modality undergoes pre-processing before being analyzed by a DL classifier, enhanced with MC Dropout for uncertainty estimation. The outputs are then combined through a decision fusion mechanism to improve overall classification accuracy. Traditional risk factors, along with CAC data, are integrated via a multiple logistic regression (MLR) model to enhance predictive performance.

The overall framework integrates the predictive risk model within a biocybernetics loop, allowing continuous feedback to the medical team. This integration ensures that the model dynamically adapts to changes in the monitored subject's condition, enabling prompt adjustments to personalized therapies based on the model's evolving predictions.

To enhance transparency and support clinical decision-making, explainable AI is integrated into the biocybernetics loop by generating saliency maps each time the model updates its risk prediction. These maps highlight key areas in input images, allowing clinicians to track where the model's focus is directed. By displaying these maps, the framework provides clear justifications for therapy adjustments, fostering trust in the model and enabling data-driven decisions for personalized treatment (Fig. 5).

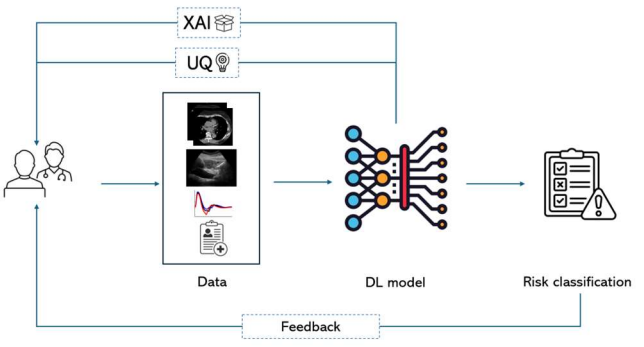

Fig. 5. Biocybernetics loop integrating a DL risk prediction model, XAI, and UQ to enhance transparency and reliability in a continuous feedback system for personalized treatment adaptation.

REFERENCES

[1] A. M. Markin, I. A. Sobenin, A. V. Grechko, D. Zhang, and A. N. Orekhov, 'Cellular Mechanisms of Human Atherogenesis: Focus on Chronification of Inflammation and Mitochondrial Mutations', *Front. Pharmacol.*, vol. 11, p. 642, May 2020, doi: 10.3389/fphar.2020.00642.

[2] P. Libby *et al.*, 'Atherosclerosis', *Nat. Rev. Dis. Primer*, vol. 5, no. 1, p. 56, Aug. 2019, doi: 10.1038/s41572-019-0106-z.

[3] K. Lechner *et al.*, 'Lifestyle factors and high-risk atherosclerosis: Pathways and mechanisms beyond traditional risk factors', *Eur. J. Prev. Cardiol.*, vol. 27, no. 4, pp. 394–406, Mar. 2020, doi: 10.1177/2047487319869400.

[4] P.-J. Touboul, 'Clinical impact of intima media measurement', *Eur. J. Ultrasound*, vol. 16, no. 1–2, pp. 105–113, Nov. 2002, doi: 10.1016/S0929-8266(02)00050-2.

[5] Y. Ben-Shlomo *et al.*, 'Aortic Pulse Wave Velocity Improves Cardiovascular Event Prediction', *J. Am. Coll. Cardiol.*, vol. 63, no. 7, pp. 636–646, Feb. 2014, doi: 10.1016/j.jacc.2013.09.063.

[6] R. Nicoll and M. Y. Henein, 'The predictive value of arterial and valvular calcification for mortality and cardiovascular events', *IJC Heart Vessels*, vol. 3, pp. 1–5, Jun. 2014, doi: 10.1016/j.ijchv.2014.02.001.

[7] The Global Cardiovascular Risk Consortium, 'Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality', *N. Engl. J. Med.*, vol. 389, no. 14, pp. 1273–1285, Oct. 2023, doi: 10.1056/NEJMoa2206916.

[8] R. L. McClelland *et al.*, '10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors', *J. Am. Coll. Cardiol.*, vol. 66, no. 15, pp. 1643–1653, Oct. 2015, doi: 10.1016/j.jacc.2015.08.035.

[9] M. Mahesh, 'The Essential Physics of Medical Imaging, Third Edition.', *Med. Phys.*, vol. 40, no. 7, p. 077301, Jul. 2013, doi: 10.1118/1.4811156.

[10] H. S. Hecht, 'Coronary Artery Calcium Scanning', *JACC Cardiovasc. Imaging*, vol. 8, no. 5, pp. 579–596, May 2015, doi: 10.1016/j.jcmg.2015.02.006.

[11] L. Saba *et al.*, 'Ultrasound-based carotid stenosis measurement and risk stratification in diabetic cohort: a deep learning paradigm', *Cardiovasc. Diagn. Ther.*, vol. 9, no. 5, pp. 439–461, Oct. 2019, doi: 10.21037/cdt.2019.09.01.

[12] G. F. Mitchell *et al.*, 'Arterial Stiffness and Cardiovascular Events: The Framingham Heart Study', *Circulation*, vol. 121, no. 4, pp. 505–511, Feb. 2010, doi: 10.1161/CIRCULATIONAHA.109.886655.

[13] K. P. Abrisham, K. Alipour, B. Tarvirdizadeh, and M. Ghamari, 'Deep Learning-Based Estimation of Arterial Stiffness from PPG Spectrograms: A Novel Approach for Non-Invasive Cardiovascular Diagnostics', in *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA: IEEE, Jul. 2024, pp. 1–7. doi: 10.1109/EMBC53108.2024.10782553.

[14] P. M. Ridker, P. W. F. Wilson, and S. M. Grundy, 'Should C-Reactive Protein Be Added to Metabolic Syndrome and to Assessment of Global Cardiovascular Risk?', *Circulation*, vol. 109, no. 23, pp. 2818–2825, Jun. 2004, doi: 10.1161/01.CIR.0000132467.45278.59.

[15] D. J. Winkel *et al.*, 'Deep learning for vessel-specific coronary artery calcium scoring: validation on a multi-centre dataset', *Eur. Heart J. - Cardiovasc. Imaging*, vol. 23, no. 6, pp. 846–854, Jun. 2022, doi: 10.1093/ehjci/jeab119.

[16] M. Liu *et al.*, 'A deep learning-based calculation system for plaque stenosis severity on common carotid artery of ultrasound images', *Vascular*, p. 17085381241246312, Apr. 2024, doi: 10.1177/17085381241246312.

[17] M.-P. Dubuisson and A. K. Jain, 'A modified Hausdorff distance for object matching', in *Proceedings of 12th International Conference on Pattern Recognition*, Jerusalem, Israel: IEEE Comput. Soc. Press, 1994, pp. 566–568. doi: 10.1109/ICPR.1994.576361.

[18] W. N. Kernan *et al.*, 'Guidelines for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association', *Stroke*, vol. 45, no. 7, pp. 2160–2236, Jul. 2014, doi: 10.1161/STR.0000000000000024.

[19] I. Sequí-Domínguez, I. Cavero-Redondo, C. Álvarez-Bueno, D. P. Pozuelo-Carrascosa, S. Nuñez De Arenas-Arroyo, and V. Martínez-Vizcaíno, 'Accuracy of Pulse Wave Velocity Predicting Cardiovascular and All-Cause Mortality. A Systematic Review and Meta-Analysis', *J. Clin. Med.*, vol. 9, no. 7, p. 2080, Jul. 2020, doi: 10.3390/jcm9072080.