Coronary Plaque Localization and Characterization using Self-Attention Tree-Graph embeddings

Mario Viti, Hugues Talbot, Nicolas Gogin

Coronary Plaque Localization and Characterization using Self-Attention Tree-Graph embeddings

Mario Viti[*1,2] MARIO.VITI@CENTRALESUPELEC.FR
[1] *GE Healthcare*
[2] *CentraleSupélec, Université Paris-Saclay, Inria*
Hugues Talbot[†2] HUGUES.TALBOT@CENTRALESUPELEC.FR
[2] *CentraleSupélec, Université Paris-Saclay, Inria*
Nicolas Gogin[‡1] NICOLAS.GOGIN@GE.COM
[1] *GE Healthcare*

Editors: Under Review for MIDL 2022

Abstract

Cardiac computed tomography angiography (CCTA) provides a non invasive imaging solution that reliably depicts the anatomical features of coronary artery diseases (CAD). Despite many successes with deep learning applications to medical imaging, recently developed methods for the automated detection of coronary plaque only suggests its feasibility as results are not yet mature enough for application deployment. The task is mainly to localize plaques and to characterize their composition depending on their appearance. Plaque composition can be divided into three classes: calcified, mixed and non-calcified. This study proposes a novel architecture for plaque detection and characterization. Its performances are tested against 2 state-of-the-art methods for coronary analysis. All trained models are evaluated and tested in the same setting using a proprietary dataset of 205 manually annotated CCTA cases. Our proposed method addresses two issues: unlike related methods it seamlessly models bifurcations and secondly it does not rely on multi-planar reformatted (MPR) visualization techniques which are by construction sensitive to centerline detection.

Keywords: Cardiac CT angiography, Major adverse cardiac event, prediction, deep learning.

1. Introduction

Coronary artery plaques present features of varying nature with corresponding functional values. In recent studies (Williams et al., 2019) non-calcified plaques were identified as a coronary obstruction independent factor associated with major adverse coronary events (MACE). Non-calcified plaques appear in patients with no flow-limiting diseases and are visually appreciable in CCTA by clinicians. The gold-standard reference method for non invasive functional analysis in CCTA is the Fractional Flow Reserve estimated via CT (Lee et al., 2016)(Pontone et al., 2016) (FFR_{CT}). Recent studies(Stuijfzand et al., 2020) shows that multivariate regression model based on calcium score(Agatston et al., 1990), stenosis degree and visually assessed anatomical plaque composition is as effective at predicting

[*] Phd Student
[†] Phd Director
[‡] Phd Advisor

MACE as the FFR_{CT}, while both have superior predictive power compared to the sole stenosis degree. Advantageously, anatomical features are reliably interpretable by clinicians, therefore plaque localization and characterization is necessary in order to achieve automatic clinical diagnosis.

1.1. Related methods

RCNN (Zreik et al., 2019) method processes the MPR coronary volumes (Kanitsar et al., 2006) as a sequence of 3D cubes interleaved by a fixed step for coronary plaque characterization and stenosis detection. Each cube is individually processed by a 3D-CNN for local feature extaction to be aggregated into a sequence and processed together by a GRU sequence classifier. In (Denzinger et al., 2019) the authors propose to replace the local feature extractor by *Radiomics* features (Lambin et al., 2012) computed from the vessel wall segmentation, these features are then processed together by the same GRU sequence classifier. Subsequently in (Denzinger et al., 2020) the authors proposed to use as input two perpendicular views of the MPR volume to a 2D CNN classifier based on VGG-16 features extractor (Cimpoi et al., 2015). The method achieved similar results to the 3D RCNN in predicting revascularization and obstruction for a given input lesion, but with a cost-effective pipeline. The TR (Ma et al., 2021) method has been recently proposed as the state of the art for obstructive stenosis ($> 50\%$ occlusion) detection against both (Lambin et al., 2012) and (Zreik et al., 2019). All presented method makes use of the extracted centerline and the MPR reconstruction as a preprocessing step.

2. Method

In our proposed CNN+ATT method, 3D patches $28 \times 28 \times 28$ with isotropic resolution of $0.35\,mm$ are sampled from the CCTA volume along a previously extracted centerline without MPR straightned reconstruction. Each cube is individually processed by a 3D-CNN for local feature extaction. The 3D-CNN is composed of 4 convolutional blocks interleaved with max pooling to achieve spatial reduction: $28^3 \mapsto 14^3 \mapsto 7^3 \mapsto 3^3$. Each convolutional block has 2 convolutional layers with residual connections and non-linear activation and batch normalization. The last convolutional layers is pooled by a global average pooling layer to obtain a fixed size feature vector (256 channels). The Self-Attention aggregation block has 2 multi-head self attention blocks with residual connections in sequence followed by a fully connected layer for classification. The number of layers is tuned experimentally on the evaluation set to be as small as possible without loss of performance.

Positional context information is input to the model by computing positional embeddings from the tree-graph structure of the centerline (Fig. 1). Given the graph $\mathcal{G} = (V, E), V = v_1, \ldots, v_n, E \in V \times V$ a k-order neighborhood of a node v_i is $\mathcal{N}_i = v_j \in V | d(v_i, v_j) \leq k$. Using the self-attention technique from (Vaswani et al., 2017), the layer operating on such neighbourhood is defined as follows:

$$v_i^0 = v_i + pos_i$$
$$Q_i^l = \sigma(W_Q v_i^{l-1})$$
$$K_i^l = \sigma(W_K \mathcal{N}_i^{l-1})$$

$$V_i^l = \sigma(W_V \mathcal{N}_i^{l-1})$$
$$v_i^{l+1} = \text{Softmax}(Q_i^l {K_i^l}^\top / \sqrt{d_k}) V_i^l \quad (1)$$

Where W_K, W_Q, W_V are trainable parameters and σ is a non linear activation function, where we overload the row vector multiplication to operate on sets as the result of $v^{l\geq 1}$ is the same independently of the ordering of $\mathcal{N}_i$. Positions pos_i are computed locally by exploiting the directed tree structure using a simple signed hop count from the center node (+1 if distal,-1 if proximal), this scalar function is then mapped to vectors of dimension d_k by using sin functions at different frequencies (Vaswani et al., 2017) to obtain positional embeddings. In order to avoid information for the center node getting diluted in the aggregation step only the center node value vector V_i is used compute the output while the whole neighbourhood query Q and key K vector are used to compute the self-attention matrix. There are many options for positional embeddings in graphs: a simple hop count suffices for this application as it holds the desirable property of a constant representation for all sorts of configurations unlike eigen-vector decomposition (Dwivedi and Bresson, 2020) and it is order invariant unlike learnable embeddings.

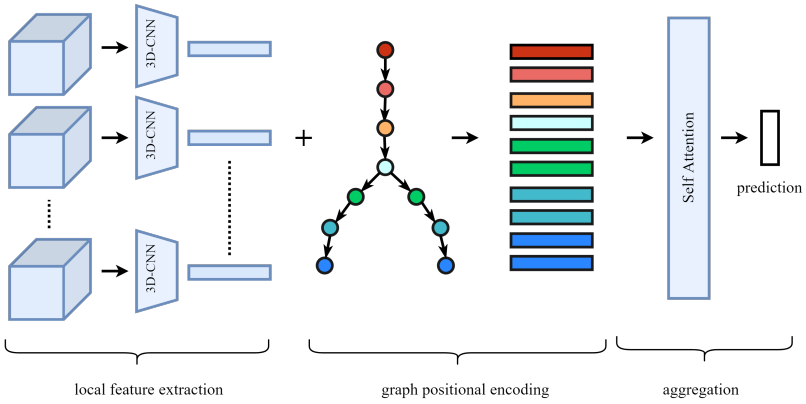

Figure 1: 3D patches are extracted from the CCTA volume by sampling points from the centerline, for each patch a 3D-CNN computes a local features vector, the graph distance is used. The centerline is extracted as a preprocessing step.

3. Experiments

3.1. Data

The dataset used in this work includes 205 coronary CT angiography (CCTA) scans, collected from clinical sites in France and Italy. Images were acquired with tube voltage ranging from 100 kVp to 120 kVp, a current from 600 mA to 1000 mA, a pixel spacing from $0.35mm$ to $0.48mm$ and $0.625mm$ slice thickness. Each patient underwent both CCTA and anatomical tests (Agatston et al., 1990) and was assigned a CAD-RADS (Hecht et al., 2017) score by trained radiologists. Each scan is paired with annotated coronary centerlines: for each lesion a starting and ending point is marked so that each point of the centerline is

associated to a label: 0 for healthy, 1 for calcified, 2 for mixed, and 3 for non-calcified. All annotations are based on clinical reports issued during medical review. Manual annotations were carried out to match the clinical report and the final annotations were submitted to a trained cardiologist for review. A subset of 69 patients using CAD-RADS as stratification criteria to match the global dataset plaque distribution (Fig. 2). The collection was carried out to obtain a CAD-RADS uniform distribution: 32, 50, 46, 41 and 36 patients from 0 to 4. The datasets contains a higher-than-normal concentration of plaques as CAD-RADS distribution usually follows an exponentially-decaying function (Fig. 2). For example in (Zreik et al., 2019) 28 non-calcified annotated plaques where found in while for this study 104 non-calcified plaques where annotated.

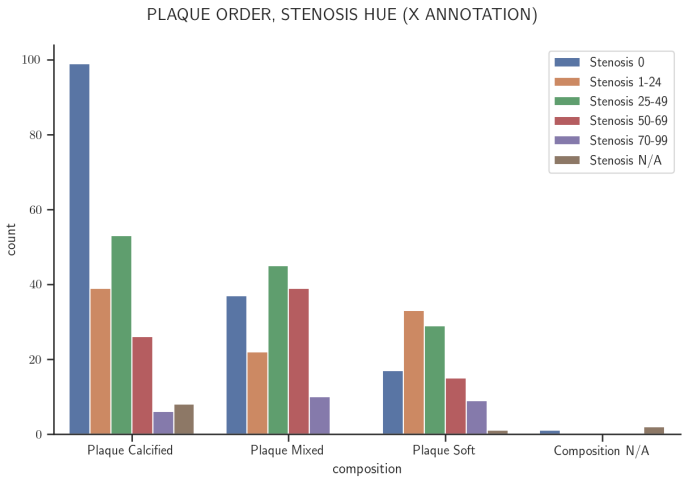

Figure 2: The 69 patients test-set annotation distribution is obtained by CAD-RADS as stratification criteria.

3.2. Training

Because of the high class imbalance, each data point is weighted so that at each iteration t batches are uniformly distributed among healthy and (any) plaque and at $t+1$ among and healthy and non-calcifed plaque. The loss is a combination of binary cross-entropy for the healthy against (any) plaque class and multiclass cross-entropy for healthy, calcified, mixed and non-calcified. Each term is weighted by a 0.5 scalar factor. For TR and RCNN, patches are arranged in a sequence while for our CNN+ATT the patches are arranged into a neighbourhoods of order k. To allow all networks the same batching, sequence size is kept fixed while for our CNN+ATT the neighbours number is clipped to 2k.

3.3. Testing

90 patients are used for training, 46 patients for hyper-parameter tuning and 69 for testing. The proposed method is tested against 2 models RCNN (Zreik et al., 2019) and TR (Ma et al., 2021). RCNN is currently the state of the art for detection and characterization of

coronary plaques while TR has been tested against RCNN for occlusive stenosis detection (>50% stenosis). All models take as input 3d patches of size 28x28x28 with a isotropic voxel size of 0.35 x 0.35 x 0.35 mm^3. For TR and RCNN patches where sampled from the MPR volume while for our CNN+ATT patches where sampled from the original CCTA volume. Each model is evaluated on a test set comprising segments of fixed length 50.4 mm. Each segment is labeled with the highest-interest occurring label in the following order: 0 for healthy, 1 for calcified, 2 for mixed, 3 for non-calcified. For prediction, the most confident inferred class overlapping the portion of highest-interest occurring label is assigned to the segment. The test set has a non uniform distribution of healthy, calcified, mixed and non-calcified plaques (Fig. 2). To capture the performance with respect to multiple classes, sensitivity and specificity are evaluated in different label combinations: calcified against healthy, mixed against healthy, non-calcified against healthy, all against healthy where all means the union of calcified, mixed and non-calcified (Fig. 3, Tab. 3.3).

For the RCNN the overlapping has not been used as the GRU unit last output is used for the prediction of each segment already. The length of the segment is not arbitrary, it was chosen to accommodate the TR architecture that uses sequences of 30 cubic patches of side 28 voxels interleaved by a 4 voxel step $((30-1)*4+28)\text{voxel}*0.35\frac{mm}{\text{voxel}} = 144\text{voxel}*0.35\frac{mm}{\text{voxel}} = 50.4mm$. CNN+ATT uses a smaller neighbourhood thus the inference is repeated for each location of the segment with the same interleaving step.

By applying the same highest-interest label overlapping strategy as assignment criteria the evaluation is also carried out at the coronary level. As most coronaries have plaques only correct label assigned is considered and performance measured by C-Index or concordance score and confusion matrix (Tab. 3.3).

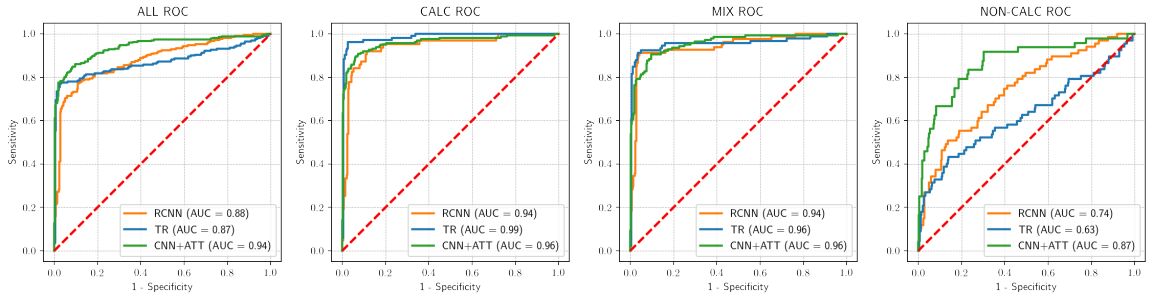

Figure 3: Segment level ROC curves

3.4. Centerline robustness

One of the major concerns about this method is the robustness to the centerline extraction quality. In order to test how results are affected by centerline quality, our method underwent evaluation using increasingly noisy centerline coordinates, to ensure that the content perceived by the model remains unchanged under random shifts the consistency of intermediate representation local feature extractoed by the 3D-CNN is measured using cosine alignment metrics. Each coordinate point p around the centerline is perturbed by a uniform noise $u \sim U(-s, s)^3$. Embeddings e_p consistency is measured using the cosine similarity as noise increases $cossim(e_p, e_{p+u})$(Fig.5). The results suggest that the method is consistent

Segment level	All				Calc			
	sens.	spec.	prec.	f1	sens.	spec.	prec.	f1
RCNN(Zreik et al., 2019)	0.83	0.72	0.69	0.75	0.82	0.94	0.91	0.86
TR(Ma et al., 2021)	0.86	0.52	0.59	0.70	0.93	**0.97**	**0.96**	**0.94**
CNN+ATT	**0.88**	**0.83**	**0.79**	**0.83**	**0.94**	0.82	0.79	0.86
	Mix				No-Calc			
	sens.	spec.	prec.	f1	sens.	spec.	prec.	f1
RCNN(Zreik et al., 2019)	0.88	**0.96**	0.91	0.86	0.67	0.65	0.57	0.61
TR(Ma et al., 2021)	0.91	0.94	**0.92**	**0.91**	0.56	0.58	0.48	0.52
CNN+ATT	**0.94**	0.83	0.79	0.85	**0.87**	**0.70**	**0.66**	**0.75**

Table 1: Segment level plaque localization.

Coronary Level	PREDICTED			
RCNN(Zreik et al., 2019)	Healthy	Calc	Mix	No-Calc
Healthy	**59**	5	3	6
Calc	8	4	**12**	10
Mix	20	6	**29**	12
No-Calc	15	3	11	**14**
C-Index	0.700			
TR(Ma et al., 2021)	Healthy	Calc	Mix	No-Calc
Healthy	**60**	2	1	10
Calc	3	10	**14**	7
Mix	21	13	**23**	10
No-Calc	20	10	**13**	10
C-Index	0.697			
CNN+ATT	Healthy	Calc	Mix	No-Calc
Healthy	**52**	2	9	10
Calc	3	**11**	10	10
Mix	6	5	**50**	6
No-Calc	12	3	10	**18**
C-Index	**0.720**			

Table 2: Coronary level plaque characterization.

up to a threshold of 2.25 mm shift from the original ideal centerline, this result is reasonable considering that the average coronary radius is about 2.5 mm.

4. Discussion

We presented an evaluation of a novel architecture that exploits the tree-graph structure of inputs data. The proposed method outmatches the state of the art for characterization of coronary plaques and non-calcified plaque detection. Non-calcified plaque remains the hardest to detect as its attenuation can be indistinguishable from its surrounding tissue, for example myocardial lipid tissue or in distal low-contrasted coronary portions (Fig. 4).

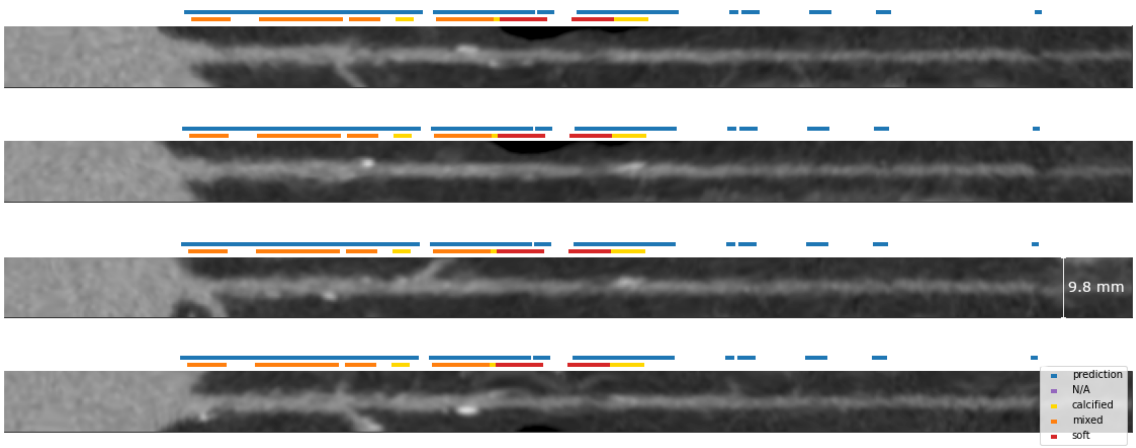

Figure 4: A descriptive example show strengths and weaknesses of the approach, in the proximal portion of the coronary plaque are correctly localized while distal low-contrasted coronary portion presents sparse false detections. Distal coronary portions with luminal diameter of $\leq 1.5mm$ may be considered non-clinical and are often excluded from automatic analysis (Zreik et al., 2019)

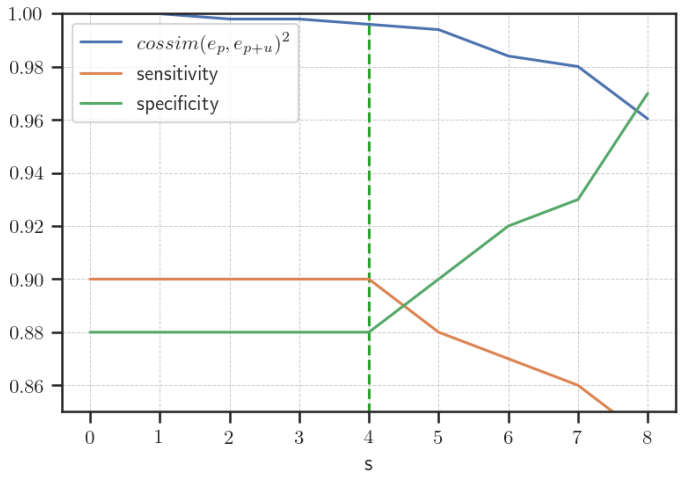

Figure 5: On the x-axis the centerline perturbation term s in voxels. The y-axis shared by sensitivity, specificity and a local features e alignement metric, which together represent the consistency of intermediate and output representation of the inputs as the centerline is distrubed around its original location.

The golden standard for medical application is visual inspection carried out by trained cardiologists, however as more data becomes available automatic coronary plaque detection will eventually become more mature. Consensus among experts in the medical community remains an issue as a high inter observer variance affects the quality of data annotations. The uncertainty is correlated with the spatial resolution of CT scanners. High resolution scans may be used to mitigate the issue.

5. Compliance with Ethical Standard

5.1. Human rights

The authors declare that the work described has been carried out in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans.

5.2. Informed consent and patient details

The authors declare that neither this report nor the collected data contain any personal information that could lead to the identification of the patient, data has been collected upon agreement with the clinical sites and patients.

Acknowledgments

Centre Cardiologique du Nord, Hôpitaux Universitaires Henri Mondor, Azienda Ospedaliero Universitaria Torino.

References

Arthur S. Agatston, Warren R. Janowitz, and Hildner. Quantification of coronary artery calcium using ultrafast computed tomography. *Journal of the American College of Cardiology*, 15(4):827–832, March 1990. ISSN 07351097. doi: 10.1016/0735-1097(90)90282-T. URL https://linkinghub.elsevier.com/retrieve/pii/073510979090282T.

M. Cimpoi, Subhransu Maji, and A. Vedaldi. Deep filter banks for texture recognition and segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3828–3836, 2015.

Felix Denzinger, Michael Wels, Nishant Ravikumar, et al. Coronary Artery Plaque Characterization from CCTA Scans Using Deep Learning and Radiomics. In Dinggang Shen, Tianming Liu, and Terry M. Peters, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, volume 11767, pages 593–601. Springer International Publishing, Cham, 2019. ISBN 9783030322502 9783030322519. doi: 10.1007/978-3-030-32251-9_65. URL http://link.springer.com/10.1007/978-3-030-32251-9_65.

Felix Denzinger, Michael Wels, and Katharina Breininger. Deep Learning Algorithms for Coronary Artery Plaque Characterisation from CCTA Scans. In Thomas

Tolxdorff, Thomas M. Deserno, Heinz Handels, Andreas Maier, Klaus H. Maier-Hein, and Christoph Palm, editors, *Bildverarbeitung für die Medizin 2020*, pages 193–198. Springer Fachmedien Wiesbaden, Wiesbaden, 2020. ISBN 9783658292669 9783658292676. doi: 10.1007/978-3-658-29267-6_42. URL http://link.springer.com/10.1007/978-3-658-29267-6_42.

Vijay Prakash Dwivedi and Xavier Bresson. A generalization of transformer networks to graphs. *CoRR*, abs/2012.09699, 2020. URL https://arxiv.org/abs/2012.09699.

Harvey S. Hecht, Paul Cronin, and Michael J. Blaha. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. *Journal of Cardiovascular Computed Tomography*, 11(1):74–84, January 2017. ISSN 19345925. doi: 10.1016/j.jcct.2016.11.003. URL https://linkinghub.elsevier.com/retrieve/pii/S1934592516302854.

Armin Kanitsar, Dominik Fleischmann, and Rainer Wegenkittl. Diagnostic Relevant Visualization of Vascular Structures. In Georges-Pierre Bonneau, Thomas Ertl, and Gregory M. Nielson, editors, *Scientific Visualization: The Visual Extraction of Knowledge from Data*, pages 207–228. Springer-Verlag, Berlin/Heidelberg, 2006. ISBN 9783540260660. doi: 10.1007/3-540-30790-7_13. URL http://link.springer.com/10.1007/3-540-30790-7_13.

Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, and Sara Carvalho. Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4):441–446, March 2012. ISSN 09598049. doi: 10.1016/j.ejca.2011.11.036. URL https://linkinghub.elsevier.com/retrieve/pii/S0959804911009993.

Ji Hyun Lee, Dalio Institute of Cardiovascular Imaging, New York-Presbyterian Hospital, New York, NY, USA, and Bríain ó Hartaigh. Fractional Flow Reserve Measurement by Computed Tomography: An Alternative to the Stress Test. *Interventional Cardiology Review*, 11(2):105, 2016. ISSN 1756-1477. doi: 10.15420/icr.2016:1:2. URL https://www.icrjournal.com/articleindex/icr.2016:1:2.

Xinghua Ma, Gongning Luo, Wei Wang, and Kuanquan Wang. Transformer Network for Significant Stenosis Detection in CCTA of Coronary Arteries. In Marleen de Bruijne, Philippe C. Cattin, Stéphane Cotin, Nicolas Padoy, Stefanie Speidel, Yefeng Zheng, and Caroline Essert, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, volume 12906, pages 516–525. Springer International Publishing, Cham, 2021. ISBN 9783030872304 9783030872311. doi: 10.1007/978-3-030-87231-1_50. URL https://link.springer.com/10.1007/978-3-030-87231-1_50.

Gianluca Pontone, Daniele Andreini, and Andrea I. Guaricci. Rationale and design of the PERFECTION (comparison between stress cardiac computed tomography PERfusion versus Fractional flow rEserve measured by Computed Tomography angiography In the evaluation of suspected cOroNary artery disease) prospective study. *Journal of Cardiovascular Computed Tomography*, 10(4):330–334, July 2016. ISSN 19345925. doi:

10.1016/j.jcct.2016.03.004. URL https://linkinghub.elsevier.com/retrieve/pii/S1934592516300405.

Wijnand J. Stuijfzand, Alexander R. van Rosendael, and Fay Y. Lin. Stress Myocardial Perfusion Imaging vs Coronary Computed Tomographic Angiography for Diagnosis of Invasive Vessel-Specific Coronary Physiology: Predictive Modeling Results From the Computed Tomographic Evaluation of Atherosclerotic Determinants of Myocardial Ischemia (CREDENCE) Trial. *JAMA Cardiology*, 5(12):1338, December 2020. ISSN 2380-6583. doi: 10.1001/jamacardio.2020.3409. URL https://jamanetwork.com/journals/jamacardiology/fullarticle/2769554.

Ashish Vaswani, Noam M. Shazeer, and Niki Parmar. Attention is all you need. *ArXiv*, abs/1706.03762, 2017.

Michelle C. Williams, Alastair J. Moss, and Marc Dweck. Coronary Artery Plaque Characteristics Associated With Adverse Outcomes in the SCOT-HEART Study. *Journal of the American College of Cardiology*, 73(3):291–301, January 2019. ISSN 07351097. doi: 10.1016/j.jacc.2018.10.066. URL https://linkinghub.elsevier.com/retrieve/pii/S0735109718392106.

Majd Zreik, Robbert W. van Hamersvelt, and Jelmer M. Wolterink. A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography. *IEEE Transactions on Medical Imaging*, 38(7):1588–1598, July 2019. ISSN 0278-0062, 1558-254X. doi: 10.1109/TMI.2018.2883807. URL https://ieeexplore.ieee.org/document/8550784/.