



# Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis

Robbert W. van Hamersvelt<sup>1</sup> · Majd Zreik<sup>2</sup> · Michiel Voskuil<sup>3</sup> · Max A. Viergever<sup>2</sup> · Ivana Išgum<sup>2</sup> · Tim Leiner<sup>1</sup>

Received: 29 May 2018 / Revised: 23 August 2018 / Accepted: 2 October 2018 / Published online: 12 November 2018  
© The Author(s) 2018

## Abstract

**Objectives** To evaluate the added value of deep learning (DL) analysis of the left ventricular myocardium (LVM) in resting coronary CT angiography (CCTA) over determination of coronary degree of stenosis (DS), for identification of patients with functionally significant coronary artery stenosis.

**Methods** Patients who underwent CCTA prior to an invasive fractional flow reserve (FFR) measurement were retrospectively selected. Highest DS from CCTA was used to classify patients as having non-significant ( $\leq 24\%$  DS), intermediate (25–69% DS), or significant stenosis ( $\geq 70\%$  DS). Patients with intermediate stenosis were referred for fully automatic DL analysis of the LVM. The DL algorithm characterized the LVM, and likely encoded information regarding shape, texture, contrast enhancement, and more. Based on these encodings, features were extracted and patients classified as having a non-significant or significant stenosis. Diagnostic performance of the combined method was evaluated and compared to DS evaluation only. Functionally significant stenosis was defined as  $\text{FFR} \leq 0.8$  or presence of angiographic high-grade stenosis ( $\geq 90\%$  DS).

**Results** The final study population consisted of 126 patients (77% male,  $59 \pm 9$  years). Eighty-one patients (64%) had a functionally significant stenosis. The proposed method resulted in improved discrimination ( $\text{AUC} = 0.76$ ) compared to classification based on DS only ( $\text{AUC} = 0.68$ ). Sensitivity and specificity were 92.6% and 31.1% for DS only ( $\geq 50\%$  indicating functionally significant stenosis), and 84.6% and 48.4% for the proposed method.

**Conclusion** The combination of DS with DL analysis of the LVM in intermediate-degree coronary stenosis may result in improved diagnostic performance for identification of patients with functionally significant coronary artery stenosis.

## Key Points

- Assessment of degree of coronary stenosis on CCTA has consistently high sensitivity and negative predictive value, but has limited specificity for identifying the functional significance of a stenosis.
- Deep learning algorithms are able to learn complex patterns and relationships directly from the images without prior specification of which image features represent presence of disease, and thereby may be more sensitive to subtle changes in the LVM caused by functionally significant stenosis.

---

Robbert W. van Hamersvelt and Majd Zreik contributed equally to this work.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00330-018-5822-3>) contains supplementary material, which is available to authorized users.

✉ Robbert W. van Hamersvelt  
R.W.vanHamersvelt-3@umcutrecht.nl

<sup>2</sup> Image Sciences Institute, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

<sup>3</sup> Department of Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

<sup>1</sup> Department of Radiology, University Medical Center Utrecht, Utrecht University, P.O. Box 85500, 3508 GA Utrecht, the Netherlands

- *Addition of deep learning analysis of the left ventricular myocardium to the evaluation of degree of coronary artery stenosis improves diagnostic performance and increases specificity of resting CCTA. This could potentially decrease the number of patients undergoing invasive coronary angiography.*

**Keywords** Artificial intelligence · Myocardial ischemia · Coronary artery disease · Computed tomography angiography

### Abbreviations

AUC	Area under the receiver operating characteristic curve
CABG	Coronary artery bypass grafting
CAD	Coronary artery disease
CCTA	Coronary computed tomography angiography
DL	Deep learning
DS	Degree of stenosis
FFR	Fractional flow reserve
ICA	Invasive coronary angiography
LVM	Left ventricular myocardium
PCI	Percutaneous coronary intervention

### Introduction

Assessment of degree of stenosis (DS) in coronary arteries using coronary computed tomography angiography (CCTA) is an accepted diagnostic tool for the detection and exclusion of coronary artery disease (CAD), with consistently high sensitivity and negative predictive value [1–3]. However, it has limited specificity in indicating the functional significance of a stenosis [2, 3]. Invasive fractional flow reserve (FFR) is currently the reference standard to indicate functional significance of a coronary stenosis and to guide treatment [4]. However, due to its invasive nature and high cost, adoption of invasive FFR in clinical practice is limited, and the search for a non-invasive method that would determine functional significance of a stenosis continues.

To address the limited specificity of CCTA, new techniques have been developed to obtain information about the functional significance of a stenosis in a non-invasive way. FFR derived from CT (FFR<sub>CT</sub>) is an emerging method which has shown promising results [5–7]. By simulating flow and pressure through the coronary arteries, a virtual FFR value is obtained. More recently, promising results have been obtained with analysis of myocardial perfusion from resting CT [8–16]. Even though it is well known that perfusion defects are more pronounced under conditions of hyperemia [17, 18], prior studies have shown the feasibility and accuracy of identification of patients with a functionally significant coronary artery stenosis with resting CCTA only [8–13]. With these approaches, functional information is obtained without the need for an additional stress perfusion acquisition, thereby saving radiation and contrast medium dose, lowering risk, and reducing examination duration and cost. In recent studies, approaches exploiting

machine learning have been proposed in which the left ventricular myocardium (LVM) in resting CCTA is analyzed and used to classify patients with regard to the presence of functionally significant coronary artery stenosis [8, 9, 13].

In classical machine learning, discriminant features describing the LVM, such as hypo-attenuation and changes in myocardial wall thickness, are manually designed by an expert. Subsequently, these features are used in an algorithm that is built to classify patients according to presence of functionally significant coronary artery stenosis [8, 9]. In contrast to the approaches using expert engineered image features, we recently proposed a deep learning (DL) algorithm, whereby the LVM features that discriminate patients with and without functionally significant coronary artery stenosis are independently learned by the algorithm directly from the image [13]. The current study expands on our previous work [13] by applying a combined method of visual stenosis grading on CCTA and only applying the DL-based analysis to the intermediate-degree stenosis. In contrast to classical machine learning-based approaches, the DL algorithm is able to independently learn generic and complex LVM patterns, and could potentially be more sensitive to changes in the LVM caused by functionally significant stenosis [19, 20].

Thus, the aim of the current study was to evaluate the added value of resting CCTA LVM deep learning analysis over coronary DS evaluation only, for identification of patients with functionally significant coronary artery stenosis.

### Materials and methods

#### Study design and population

This was a single center, retrospective, observational study. Between 2012 and 2016, 136 consecutive patients who underwent a CCTA within 1 year prior to an invasive FFR measurement were selected. Patients were excluded ( $n = 10$ ) when automatic LVM segmentation failed or when a significant part was not or incorrectly segmented, following grading defined by Abadi et al [21]. This was the only exclusion criteria used to indirectly exclude scans with bad image quality, suboptimal contrast, motion artifact, or any other artifact. No further in- or exclusion criteria were applied. This study was approved by our local institutional review board at the University Medical Center Utrecht in the Netherlands (protocol number 15/608). The need for informed consent was

waived by the institutional review board, since the study was performed on already available imaging data.

### CCTA image acquisition

All patients underwent CCTA on a 256-slice CT system with a collimation of  $128 \times 0.625$  mm (Brilliance iCT, Philips Healthcare). In accordance with the Society of Cardiovascular Computed Tomography guidelines [22], beta-blockers and/or nitroglycerine was administered to target a heart rate of 60 beats per minute (bpm). Patients were imaged using either an ECG-triggered step-and-shoot protocol ( $\leq 60$  bpm) or a retrospectively ECG-gated spiral protocol (heart rate  $> 60$  bpm). An intravenous contrast agent (70–80 mL) (iopromide, Ultravist 300, Bayer Healthcare), followed by a mix (60–67 mL) of 30% contrast agent and 70% saline, followed by saline (30–40 mL) was administered using a flow rate of 6 ( $< 80$  kg) or 6.7 mL/s ( $\geq 80$  kg). Bolus tracking with 7-s delay was used for acquisition (threshold of 100 HU for step-and-shoot or 200 HU for spiral). For step-and-shoot, images were acquired at 78% of the R-R interval with a rotation time of 0.27 s. A weight-dependent tube voltage and corresponding current of 80, 100, or 120 kVp and 195, 210, and 210 mAs were used in patient weighing  $< 50$  kg, 50–80 kg, and  $> 80$  kg, respectively. For spiral, a heart rate-dependent pitch (0.16–0.18) and rotation time (0.27–0.33) were used. Weight-dependent tube voltage and corresponding tube current were 100, 120, or 120 kVp and 600, 600, and 700 mAs for patients weighing  $< 65$  kg, 65–80 kg, and  $> 80$  kg, respectively. Images acquired at 75% of the R-R interval were selected for further analysis. For both protocols, axial images were reconstructed with a slice thickness of 0.9 mm at 0.45-mm increment using iterative reconstruction (iDose<sup>4</sup>, Philips Healthcare). Maximum DS for each coronary artery was visually scored in consensus by two observers and categorized in 0%, 1–24%, 25–49%, 50–69%, and  $\geq 70\%$  or non-diagnostic, in accordance with dedicated guidelines [23, 24].

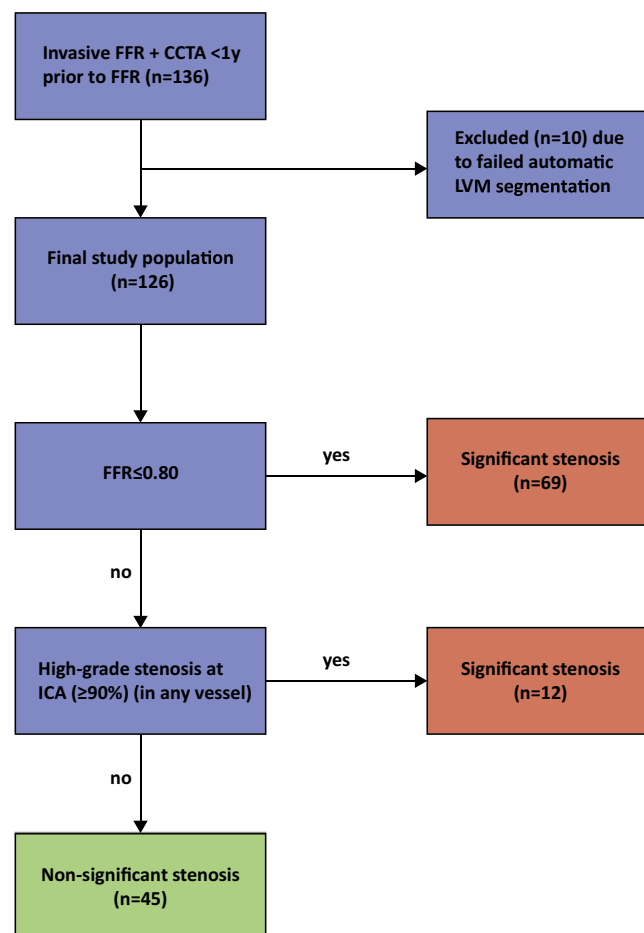
### Invasive coronary angiography and FFR

All patients underwent invasive coronary angiography (ICA) and FFR according to standard clinical guidelines. FFR measurements were performed in the coronary arteries when deemed clinically indicated. Measurements were performed under adenosine-induced hyperemia (140  $\mu\text{g}/\text{kg}$  as continuous intravenous infusion), and FFR was measured as distal as possible in the target vessel using a coronary pressure wire (Pressure<sup>TM</sup> Certus<sup>TM</sup> Wire, St. Jude Medical). The FFR value was automatically calculated by dividing the pressure measured distal to the stenosis by the pressure measured at the level of the guiding catheter in mmHg. In case FFR was  $\leq 0.80$ , the stenosis was considered significant. In the vessels with high-grade coronary stenosis at ICA ( $\geq 90\%$  DS),

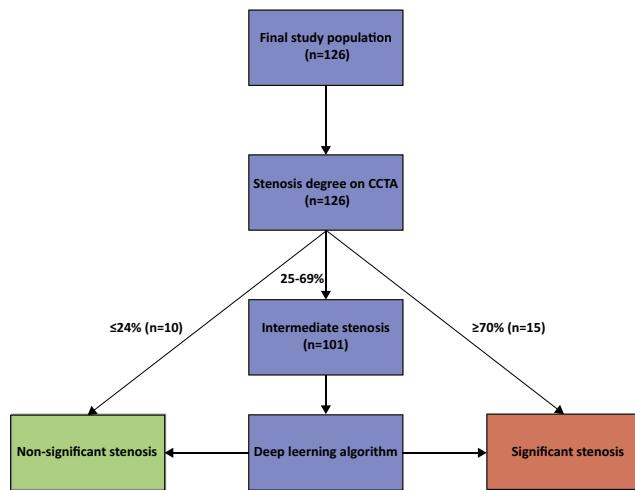
invasive FFR was not performed, but stenoses were deemed significant based on DS alone [25]. Therefore, when evaluating ischemia on a patient basis, either a significant FFR ( $\leq 0.8$ ) or high-grade stenosis at ICA ( $\geq 90\%$  DS) was used to indicate the presence of functionally significant stenoses (Fig. 1).

### CCTA and deep learning-based image analysis

To evaluate the diagnostic performance of adding LVM DL analysis to evaluation of DS, patients were subdivided into three different groups based on the highest degree of coronary stenosis at CCTA (group 1, DS  $\leq 24\%$ ; group 2, DS 25–69%, and group 3,  $\geq 70\%$  DS). Patients with  $\leq 24\%$  DS at CCTA ( $n = 10$ ) were considered to have a non-significant stenosis, and patients with  $\geq 70\%$  DS at CCTA ( $n = 15$ ) were considered to have a significant stenosis. Patients with intermediate stenosis at CCTA (25–69% DS,  $n = 101$ ) were referred for fully automatic DL analysis of the LVM to indicate whether a functionally significant stenosis was present (Fig. 2). With the DL method, resting CCTA images were automatically



**Fig. 1** Flowchart indicating ischemia on a patient basis. CCTA, coronary computed tomography angiography; FFR, fractional flow reserve; ICA, invasive coronary angiography; LVM, left ventricular myocardium; n, number of patients



**Fig. 2** Flowchart of analysis on a patient basis. CCTA, coronary computed tomography angiography; n, number of patients

analyzed using the method as described by Zreik et al [13]. The prior technical study [13] focused on development of the DL algorithm, while assessing the effect of the different building blocks. The current clinical study expands and validates the previous study in two ways. First, by employing a clinically more relevant reference standard to indicate functional significant stenosis on a patient level: either a significant FFR ( $\leq 0.8$ ) or high-grade stenosis at ICA ( $\geq 90\%$  DS). Second, a combined method of visual stenosis grading on CCTA and only applying the DL-based analysis to the intermediate-degree stenosis is applied. In this way, the DL method could be trained and evaluated on this subset of most challenging patient category of intermediate stenosis only. A full description of the DL method can be found elsewhere [13]; a graphical summary is presented in Fig. 3, and a more detailed description about feature extraction in Supplement Fig. 1. First, LVM was automatically segmented on all CT slices using a multiscale convolutional neural network (Fig. 4). Subsequently, the LVM was characterized (encoded) on all CT slices by the algorithm in an unsupervised manner using a convolutional auto-encoder. Using these encodings, which likely contain information regarding shape, texture, contrast enhancement, and more, features were extracted to represent the whole LVM as a volume. Based on these features, patients were classified according to the presence of functionally significant coronary stenosis using a support vector machine classifier. For the current study, the DL method was trained and validated on the subset of patients with intermediate stenosis ( $n = 101$ ) using an invasively measured FFR  $\leq 0.80$  or angiographic high-grade stenosis ( $\geq 90\%$  DS) as an indicator for presence of functionally significant stenosis. All patients with intermediate stenosis ( $n = 101$ ) were classified in 10-fold stratified cross-validation experiments. For each 10-fold cross-validation, the dataset was divided into ten stratified subsets whereby one subset was used as validation set and

the other nine subsets were used for training. This process was repeated ten times, whereby each subset was used once as validation set. To evaluate robustness of the method, the 10-fold cross-validation was repeated 50 times, with randomization of the data after each repetition. To allow for evaluation of the combined method (DS and LVM DL on intermediate stenosis) on the entire set of patients, patients found to be non-significant ( $\leq 24\%$  DS) and significant ( $\geq 70\%$  DS) based on DS on CCTA were assigned probability values of 0 s and 1 s, respectively. For patients with intermediate stenosis, the output probability of the LVM DL model in each cross-validation experiment was employed.

## Statistical analysis

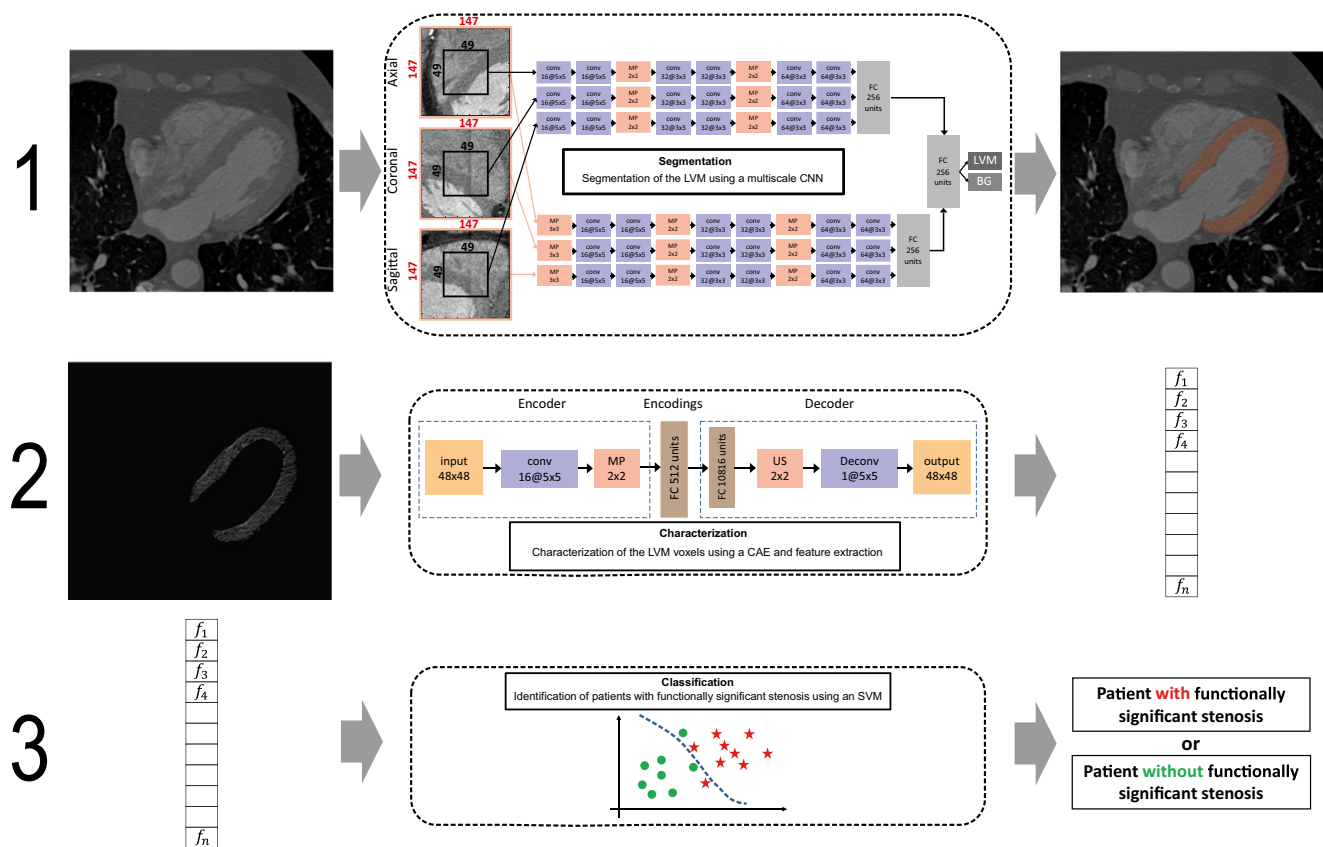
Diagnostic performance of the proposed method (DS evaluation combined with DL analysis of LVM on patients with intermediate stenosis at CCTA) for predicting the presence of functionally significant stenosis on a patient basis was evaluated and compared to the diagnostic performance of DS evaluation alone. Diagnostic performance was evaluated using accuracy, sensitivity, specificity, negative predictive value, positive predictive value, and area under the receiver operating characteristic curve (AUC). For the evaluation of the DS combined with DL analysis, diagnostic performance was evaluated on all 50 cross-validation experiments, and the average  $\pm$  standard deviation (SD) was used to represent diagnostic performance. For DS only, diagnostic performance is presented with 95% confidence interval (CI). Because prior events and interventions could influence texture changes, an additional sensitivity analysis was performed after exclusion of patients with prior events and interventions (prior myocardial infarction (MI), prior PCI, and/or prior CABG,  $n = 23$ ). The Shapiro-Wilk test was used to evaluate normality of the data. Continuous values are listed as mean with SD and categorical values as percentages, unless stated otherwise. A  $p$  value  $< 0.05$  was used to indicate statistical significance. IBM SPSS version 21.0 (IBM corp.) and MedCalc Statistical Software version 17.7.2 (MedCalc Software BVBA) were used for statistical analyses.

## Results

### Patient characteristics

In total, 136 patients were eligible, of which 10 were excluded due to incorrect or failed automatic LVM segmentation. The final study population consisted of 126 patients (77% male, mean age  $59 \pm 9$  years). Baseline characteristics are listed in Table 1.





**Fig. 3** Graphical summary of the utilized deep learning method. The described DL method includes three stages. First (1), LVM is automatically segmented using a multiscale convolutional neural network (CNN). The multiscale CNN includes two identical streams; each analyzes a set of triplanar input patches taken at a single scale (black and red squares) from the axial, coronal, and sagittal image slices with the target voxel in their centers. Each set is processed with a combination of convolutional (conv), max-pooling (MP), and fully connected (FC) layers. In the end, single voxels are classified as myocardium or background. This was performed on all CT slices to segment the whole myocardium. All LVM voxels are analyzed in the next stage. Second (2), LVM was characterized (encoded) on all CT slices by the algorithm in an unsupervised manner using a convolutional auto-encoder (CAE). The

CAE contains two parts, an encoder and a decoder. The encoder compresses the data to a lower dimensional representation by convolutional and max-pooling layers. The decoder expands the compressed form to reconstruct the input data by deconvolutional (deconv) and upsampling (US) layers. To represent the entire LVM, statistics over encodings of all LVM voxels (on all CT slices) are used as features. Third (3), based on the extracted features, patients are classified with a support vector machine (SVM) to those with or without functionally significant coronary artery stenosis. BG, background; CAE, convolutional auto-encoder; CCTA, coronary computed tomography angiography; CNN, convolutional neural network; conv, convolutional; deconv, deconvolutional; FC, fully connected; LVM, left ventricular myocardium; MP, max-pooling; SVM, support vector machine; US, upsampling

### Invasive coronary angiography and FFR

The median number of days (IQR) between CCTA and invasive FFR was 33 (40). Eighty-one patients (81/126, 64%) had a functionally significant stenosis, of which 69 patients (69/81, 85%) had an FFR  $\leq 0.80$ . In the remaining twelve patients (12/81, 15%), lowest FFR value was  $> 0.80$  and a high-grade stenosis on ICA ( $\geq 90\%$  DS) was present in another coronary artery (Fig. 1). The distribution of severity of disease on a patient-by-patient basis is depicted in Fig. 5.

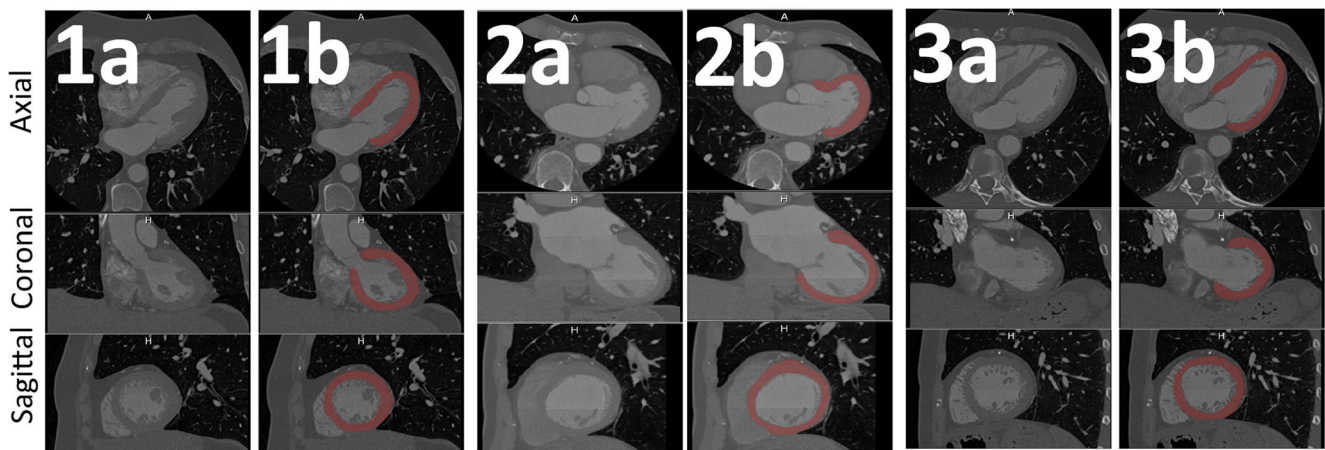
### Degree of stenosis on CCTA

Maximum DS on a patient basis was 0%, 1–24%, 25–49%, 50–69%, and  $\geq 70\%$  in 2, 8, 10, 91, and 15 patients, respectively. In

20 patients, at least one segment was scored non-diagnostic, of which 18 patients had a significant stenosis ( $\geq 50\%$  DS) in another segment. In the remaining two patients, maximum DS in another segment was 25–49%, and PCI was present.

### Diagnostic performance

The diagnostic performance of analysis by evaluation of DS alone to indicate functionally significant stenosis was moderate with an AUC of 0.68 (95%CI 0.59–0.76) (Fig. 6). When applying a threshold of  $\geq 50\%$  DS to indicate functional significance of a stenosis, sensitivity was 92.6% (95%CI 84.6–97.2%) and specificity 31.1% (95%CI 18.2–46.6%). Changing the threshold to lower or higher DS resulted in commensurate changes in sensitivity or specificity (Table 2). In ten patients (10/126, 8%),



**Fig. 4** Example of fully automatic left ventricular myocardium segmentation. Coronary CTA images of three randomly selected patients (1, 2, and 3). For each patient, the left column (a) depicts a conventional image, and the right column (b), the fully automatic LVM

segmentation overlay. LVM was automatically segmented on all CT slices using a multiscale convolutional neural network. For each patient, an example of one axial slice (top), one coronal slice (middle), and one sagittal slice (bottom) is shown. LVM, left ventricular myocardium

the maximum DS at CCTA was  $\leq 24\%$  (Fig. 2). In all of these patients, FFR was higher than 0.8 (mean  $\pm$  SD;  $0.89 \pm 0.05$ ) and no high-grade stenosis on ICA was present. Fifteen patients (15/126, 12%) had  $\geq 70\%$  DS at CCTA, and in 14 of these patients (14/15, 93%), FFR was  $\leq 0.8$  or high-grade stenosis on ICA was present. In one case, FFR was 0.82, and no high-grade stenosis on ICA was present. The remaining 101 patients with intermediate DS at CCTA were subjected to LVM DL analysis. Classification based on the combination of DS from CCTA ( $\leq 24\%$  and  $\geq 70\%$ ) and LVM DL analysis on intermediate stenosis (25–69%) improved discrimination (AUC =  $0.76 \pm 0.02$ ) compared to classification based on DS from CCTA only (AUC =  $0.68$  [95%CI 0.59–0.76]) (Fig. 6). The combined method resulted in an increased specificity of  $48.4 \pm 0.04\%$  at the

expense of a small decrease in sensitivity ( $84.6 \pm 0.03\%$ ), compared to the specificity 31.1% (95%CI 18.2–46.6%) and sensitivity 92.6% (95%CI 84.6–97.2%) of DS from CCTA only ( $\geq 50\%$  DS to indicate functional significant stenosis) (Table 2). To evaluate the robustness of the DL method applied on different patient cohorts, the DL analysis was applied on the complete cohort ( $n = 126$ ) and compared to DL analysis of intermediate stenosis only ( $n = 101$ ) (supplement Fig. 2).

**Diagnostic performance in patients without prior MI, PCI, or CABG ( $n = 103$ )**

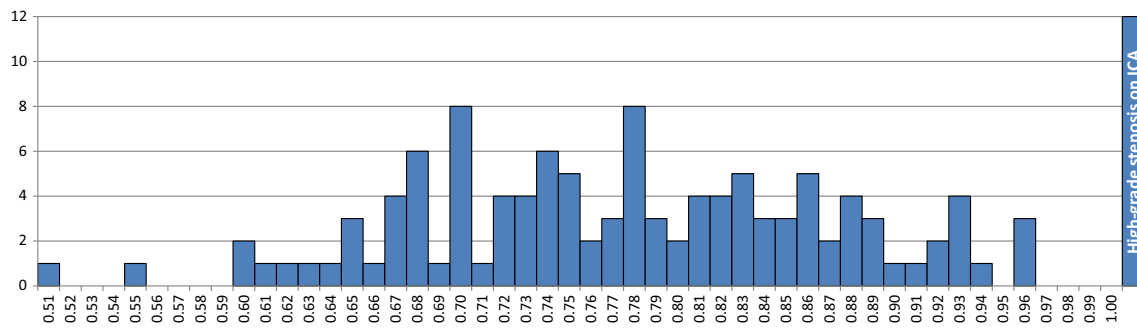
In eight patients (8/103, 8%), maximum DS at CCTA was  $\leq 24\%$ , and in 12 patients (12/103, 12%), maximum DS was

**Table 1** Patient characteristics

Characteristics	Total population ( $n = 126$ )	Patients with intermediate stenosis ( $n = 101$ )
Sex (men), $n$ (%)	97 (77)	77 (76)
Age (years), mean $\pm$ SD	$59 \pm 9$	$60 \pm 9$
Men	$58 \pm 9$	$58 \pm 9$
Women	$65 \pm 8$	$65 \pm 8$
Body mass index ( $\text{kg}/\text{m}^2$ ), mean $\pm$ SD	$27 \pm 4$	$28 \pm 4$
Cardiovascular risk factors, $n$ (%)		
Current smoker	31 (25)	23 (23)
Diabetes	24 (19)	21 (21)
Dyslipidemia	109 (87)	87 (86)
Hypertension	110 (87)	89 (88)
Previous myocardial infarction, $n$ (%)	14 (11)	11 (11)
Previous percutaneous coronary intervention, $n$ (%)	19 (15)	15 (15)
Previous coronary artery bypass grafting, $n$ (%)	2 (2)	2 (2)
Coronary calcium Agatston score, median (IQR)	464 (800) <sup>a</sup>	526 (782) <sup>a</sup>

IQR interquartile range,  $n$  number of patients, SD standard deviation

<sup>a</sup> in three cases coronary calcium scoring was not performed due to motion artifacts of the coronary arteries



**Fig. 5** Distribution of lowest FFR value per patient ( $n = 126$ ). Ischemia was evaluated on a patient basis, an  $\text{FFR} \leq 0.8$  or angiographic high-grade stenosis ( $\geq 90\%$  DS) was used to indicate the presence of a functionally significant stenosis. Per patient, the lowest FFR value is depicted, unless this concerned an  $\text{FFR} > 0.8$ , and a high-grade stenosis on ICA ( $\geq 90\%$

DS) was present in another vessel, then this was scored as “high-grade stenosis on ICA”, and FFR value of that patient is not depicted. DS, degree of stenosis; FFR, fractional flow reserve; ICA, invasive coronary angiography

$\geq 70\%$ . DL was applied to the intermediate stenosis ( $n = 83$ ). Also in this sub analysis, the combined method showed an improved discrimination ( $\text{AUC} = 0.77 \pm 0.02$ ) compared to evaluation of DS alone ( $\text{AUC} = 0.67$  [95%CI 0.57–0.76]) (Supplement Fig. 3). Other measures of diagnostic performance were also comparable to the results of the full sample (Supplement Table 1).

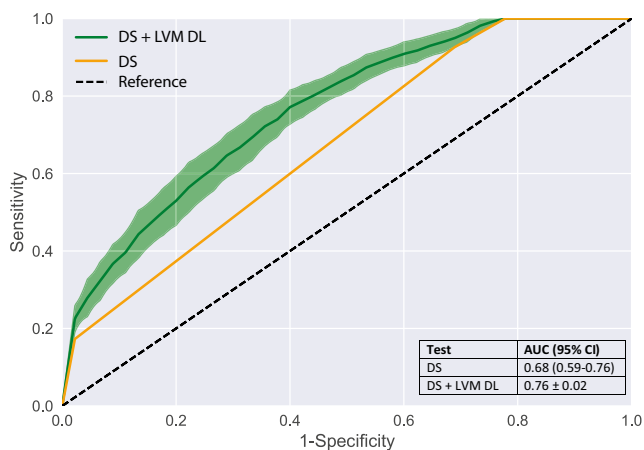
## Discussion

In the current study, we aimed to evaluate the added value of resting CCTA LVM deep learning analysis over coronary DS evaluation only, for identification of patients with functionally significant coronary artery stenosis. We demonstrated that the combination of DS with DL analysis of the LVM in

intermediate-degree coronary stenosis improves the specificity of CCTA in comparison to evaluation of the DS alone.

Visual determination of the DS from CCTA is a highly sensitive and established diagnostic tool for evaluation of patients with chest pain of suspected coronary origin. However, CCTA presently lacks specificity for identifying the functional significance of coronary stenosis [1, 2]. A tiered approach with subsequent functional testing (e.g., stress myocardial perfusion imaging) can reduce the number of patients undergoing ICA [26]. However, this approach mandates additional diagnostic testing before deciding on the necessity for PCI. Here, we propose combining evaluation of DS with LVM DL analysis to improve the specificity of CCTA. The combined evaluation has the potential to avoid additional examinations because analysis is performed on the already acquired CCTA images.

Analysis of the LVM on resting CCTA has been previously studied. Two studies evaluated resting dual-energy CT for the detection of perfusion defects confirmed by  $\geq 50\%$  stenosis at ICA [15, 16]. They found a slight decrease in sensitivity (79–90%) and increase in specificity (86–92%) compared to  $\geq 50\%$  DS on CCTA only (82–98% and 88–91%, respectively). Osawa et al [10] visually evaluated perfusion of the LVM in all cardiac phases from a retrospectively ECG-triggered scan at rest and compared their results with invasive FFR measurements. By combining this LVM analysis with DS, they found an incremental value ( $\text{AUC} = 0.82$ ) over CCTA alone ( $\text{AUC} = 0.71$ ), which is in line with our study. In two recent studies, machine learning–approaches using expert-designed features for classifying patients with functionally significant stenosis were described [8, 9]. Xiong et al [8] found a good discrimination for the detection of a perfusion defect (max.  $\text{AUC} = 0.73$ ). However, in this study, a stenosis degree of  $\geq 50\%$  DS on ICA was used as reference [8], and no FFR was performed, making a direct comparison with our results impossible. Han et al [9] used FFR as reference and described that their algorithm, combined with evaluation of DS, showed an added value ( $\text{AUC} = 0.75$ ) over DS alone ( $\text{AUC} = 0.68$ ) [9], which is closely in line with



**Fig. 6** Receiver operating characteristic curves. Diagnostic performance of DS and a combination of DL added to DS on CCTA for predicting functionally significant stenosis on a patient basis. For the combined method, ROC curves and AUC are depicted as average  $\pm$  SD of 50 cross-validation experiments. AUC, area under the receiver operating characteristic curve; CCTA, coronary computed tomography angiography; DL, deep learning; DS, degree of stenosis; LVM, left ventricular myocardium; ROC, receiver operating characteristic

**Table 2** Diagnostic performance

Method and threshold	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC
CCTA ≥ 25% DS	100.0 (81/81) [95.5–100.0]	22.2 (10/45) [11.2–37.1]	69.8 (81/116) [66.4–73.0]	100.0 (10/10) [100.0–100.0]	72.2 (91/126) [64.3–80.2]	0.68 [0.59–0.76]
CCTA ≥ 50% DS	92.6 (75/81) [84.6–97.2]	31.1 (14/45) [18.2–46.6]	70.8 (75/106) [66.3–74.8]	70.0 (14/20) [49.1–85.0]	70.6 (89/126) [62.6–78.7]	0.68 [0.59–0.76]
CCTA ≥ 70% DS	17.3 (14/81) [9.8–27.3]	97.8 (44/45) [88.2–99.9]	93.3 (14/15) [65.5–99.0]	39.6 (44/111) [37.1–42.3]	46.0 (58/126) [37.2–54.9]	0.68 [0.59–0.76]
CCTA DS + DL combined	84.6 ± 0.03	48.4 ± 0.04	74.7 ± 0.01	64.0 ± 0.05	71.7 ± 0.02	0.76 ± 0.02

Data is given in percentage, data in parentheses is raw data, and data in square brackets is 95% confidence interval. For the combined method, data is depicted as average ± SD of 50 cross-validation experiments. AUC area under the receiver operating characteristic curve, CCTA coronary computed tomography angiography, DL deep learning, DS degree of stenosis, NPV negative predictive value, PPV positive predictive value, SD standard deviation

the findings in the present study. However, patients with intermediate DS comprised a minority of the subjects studied by Han et al [9] (33%, 82/252), whereas the present study consisted of a majority of intermediate stenosis (80%, 101/126) and was exclusively focused on the added value of DL in this most challenging patient category.

FFRct is another technique performed on resting CCTA that has shown promising results (AUC = 0.79–0.93) for the evaluation of functionally significant stenosis [5–7]. However, FFRct depends on lumen segmentations which can be challenging or impossible in patients with high-density calcified plaque, motion or misalignment artifacts, and/or prior CABG or PCI [5–7]. Our proposed method may be less affected by these challenges as the analysis of intermediate stenosis is performed only on the LV myocardium. Results of the present study underscore the need to look beyond the coronary arteries in the quest to improve specificity of CCTA. A recent systematic review by Cook et al [27] found that FFRct has a lower diagnostic performance with FFRct values around the cut point (0.7–0.8). In addition, patients with physiologically intermediate lesions (invasive FFR 0.7–0.8) comprised only a small minority of all the patients studied (12.8%) with a median FFR of 0.88, indicating a focus on patients with milder disease. In the current study, nearly three times as many patients (36.5%, 46/126) had physiologically intermediate lesions, and median FFR was 0.78, indicating a more diseased population. This supports the notion that the value of additional analysis of the myocardium is likely to be highest in patients with intermediate stenosis. However, a limitation of the present study is that results are only available on a patient basis; therefore, no indication could be given as to which stenosis was functionally significant, while FFRct has the ability to evaluate functional significance of a stenosis on a vessel basis. It is likely that a combined approach using both techniques will lead to further increase in specificity for identifying flow-limiting lesions in patients with intermediate-degree stenosis at CCTA.

DL algorithms are able to learn new complex patterns and relationships directly from the images without prior specification of which image features represent presence of disease. Therefore, DL may be more sensitive to subtle changes in the LVM caused by functionally significant stenosis, which can be difficult to detect by a human observer [19, 20]. In spite of excellent performance of deep learning techniques demonstrated in many medical image analysis tasks, their interpretability is very limited, and we do not yet fully understand their inner working [19, 28]. Also in this work, even though the extracted encodings are relevant in representing LVM (Supplement Fig. 1), they are not readily interpretable. As distinct combinations of encodings represent different LVM appearances (e.g., normal vs. thin), specific encodings do not correspond to specific physical appearances. This prevented us from interpreting, visualizing, and localizing differences within the LVM for patients with significant stenosis. Future work might address these limitations [19, 28].



This study has limitations. First, the retrospective single-center study design has to be taken into account. This could have induced selection bias as CT could have been the reason for referral to invasive FFR. In addition, because of this bias, the study group consisted of patients with high Agatston scores and a high prevalence of stenosis and extensive CAD with mainly intermediate stenosis ( $n = 101/126$ ). Due to blooming, calcified plaques are known to cause overestimation of DS and thereby decrease specificity [29]. In addition, as reported in prior studies [2, 3], DS on CCTA showed high diagnostic accuracy for low- and high-grade stenosis ( $\leq 24\%$  and  $\geq 70\%$ ), which decreased when evaluating intermediate-degree stenosis (25–69%). This is reflected in the low specificity for DS only on CCTA found in the current study (31.1%) compared to reported in literature (40–83%) [1–3]. Second, maximum DS for each coronary artery was visually categorized, and no continuous quantitative measurements were performed. Although continuous quantitative measurements would allow for a more accurate way to evaluate diagnostic accuracy of DS, this is not regularly performed in clinical practice and was therefore not performed. Third, in the current study, no comparison with another type of functional assessment was performed (e.g., FFRct or visual assessment of myocardial perfusion), thereby limiting the evaluation of the added value of the current method compared to other available methods. Fourth, we used equipment of one single vendor and results may therefore be limited to this vendor. A final limitation is the relative small patient cohort and unbalanced dataset, which can introduce bias in performance. Although we performed 50 repetitions of 10-fold cross-validation experiments with randomization after each repetition, results may still be partially affected by coincidental findings. Analysis on a separate patient cohort as well as prospective studies needs to be performed to assess whether the found correlation also implies causation. Future work will address these limitations.

In conclusion, combining assessment of degree of stenosis with DL analysis of the LVM may result in improved diagnostic performance for identification of patients with functionally significant coronary artery stenosis. Future research is warranted to evaluate this approach on patients with low pre-test probability of obstructive coronary disease, more clinically encountered at CCTA.

**Funding** This work was financially supported by the project FSCAD, funded by the Netherlands Organization for Health Research and Development (ZonMw) with participation of Pie Medical imaging BV in the framework of the research program IMDI (Innovative Medical Devices Initiative), project 104003009. The University Medical Center Utrecht, Department of Radiology, receives research support from Philips Healthcare.

## Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Prof. Dr. Tim Leiner.

**Conflict of interest** Robbert W. van Hamersvelt, Majd Zreik, and Michiel Voskuil have nothing to disclose. Max A. Viergever and Ivana Išgum received Research grants from the Netherlands Organization for Scientific Research (NWO)/ Foundation for Technological Sciences (number 12726) with industrial participation (Pie Medical Imaging, 3Mensio Medical Imaging). Max A. Viergever, Ivana Išgum, and Tim Leiner received research grants from the Netherlands Organization for Health Research and Development (FSCAD, number 104003009); Research grants from Netherlands Organization for Scientific Research (NWO)/ Foundation for Technological Sciences (number P15–26) with industrial participation (Pie Medical Imaging, Philips Healthcare); and Research grants Pie Medical Imaging. Tim Leiner received research grants from Philips Healthcare and Research grants Bayer.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Informed consent** Written informed consent was waived by the Institutional Review Board.

**Ethical approval** Institutional Review Board approval was obtained by our local institutional review board at the University Medical Center Utrecht in the Netherlands (protocol number 15/608).

**Study subjects or cohorts overlap** Some study subjects or cohorts have been previously reported in “Zreik M, Lessmann N, van Hamersvelt RW, et al (2018) Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.*” which was focused on the development of the DL algorithm. The current clinical study expands and validates on the previous study by applying a combined method of visual stenosis grading on CCTA and only applying the deep learning-based analysis to the intermediate-degree stenosis.

## Methodology

- Retrospective
- Diagnostic study
- Performed at one institution

**Disclaimer** The sponsors had no role in the design and conduct of the study, in the collection, analysis, interpretation of the data, and in the preparation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Budoff MJ, Dowe D, Jollis JG et al (2008) Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease. *J Am Coll Cardiol* 52:1724–1732. <https://doi.org/10.1016/j.jacc.2008.07.031>
2. Meijboom WB, Van Mieghem CA, van Pelt N et al (2008) Comprehensive assessment of coronary artery stenoses: computed tomography coronary angiography versus conventional coronary angiography and correlation with fractional flow reserve in patients

- with stable angina. *J Am Coll Cardiol* 52:636–643. <https://doi.org/10.1016/j.jacc.2008.05.024>
3. Ko BS, Wong DT, Cameron JD et al (2014) 320-row CT coronary angiography predicts freedom from revascularisation and acts as a gatekeeper to defer invasive angiography in stable coronary artery disease: a fractional flow reserve-correlated study. *Eur Radiol* 24:738–747. <https://doi.org/10.1007/s00330-013-3059-8>
  4. Tonino PA, De Bruyne B, Pijls NH et al (2009) Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med* 360:213–224. <https://doi.org/10.1056/NEJMoa0807611>
  5. Tesche C, De Cecco CN, Albrecht MH et al (2017) Coronary CT angiography–derived fractional flow reserve. *Radiology* 285:17–33. <https://doi.org/10.1148/radiol.2017162641>
  6. Norgaard BL, Leipsic J, Gaur S et al (2014) Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease the NXT trial (analysis of coronary blood flow using CT angiography: next steps). *J Am Coll Cardiol* 63:1145–1155. <https://doi.org/10.1016/j.jacc.2013.11.043>
  7. Koo BK, Erglis A, Doh JH et al (2011) Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. Results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study. *J Am Coll Cardiol* 58:1989–1997. <https://doi.org/10.1016/j.jacc.2011.06.066>
  8. Xiong G, Kola D, Heo R, Elmore K, Cho I, Min JK (2015) Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. *Med Image Anal* 24:77–89. <https://doi.org/10.1016/j.media.2015.05.010>
  9. Han D, Lee JH, Rizvi A, et al (2017) Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: a machine learning approach. *J Nucl Cardiol* 25:223–233. <https://doi.org/10.1007/s12350-017-0834-y>
  10. Osawa K, Miyoshi T, Miki T et al (2016) Diagnostic performance of first-pass myocardial perfusion imaging without stress with computed tomography (CT) compared with coronary CT angiography alone, with fractional flow reserve as the reference standard. *PLoS One* 11:e0149170. <https://doi.org/10.1371/journal.pone.0149170>
  11. Branch KR, Busey J, Mitsumori LM et al (2013) Diagnostic performance of resting CT myocardial perfusion in patients with possible acute coronary syndrome. *AJR Am J Roentgenol* 200:W450–W457. <https://doi.org/10.2214/AJR.12.8934>
  12. Pursnani A, Lee AM, Mayrhofer T et al (2015) Early resting myocardial computed tomography perfusion for the detection of acute coronary syndrome in patients with coronary artery disease. *Circ Cardiovasc Imaging* 8:e002404–e002404. <https://doi.org/10.1161/CIRCIMAGING.114.002404>
  13. Zreik M, Lessmann N, van Hamersvelt RW et al (2018) Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal* 44:72–85. <https://doi.org/10.1016/j.media.2017.11.008>
  14. Mannil M, von Spiczak J, Manka R, Alkadhi H (2018) Texture analysis and machine learning for detecting myocardial infarction in noncontrast low-dose computed tomography. *Invest Radiol* 53:338–343. <https://doi.org/10.1097/RLI.0000000000000448>
  15. Ruzsics B, Lee H, Zwerner PL, Gebregziabher M, Costello P, Schoepf UJ (2008) Dual-energy CT of the heart for diagnosing coronary artery stenosis and myocardial ischemia-initial experience. *Eur Radiol* 18:2414–2424. <https://doi.org/10.1007/s00330-008-1022-x>
  16. Wang R, Yu W, Wang Y et al (2011) Incremental value of dual-energy CT to coronary CT angiography for the detection of significant coronary stenosis: comparison with quantitative coronary angiography and single photon emission computed tomography. *Int J Cardiovasc Imaging* 27:647–656. <https://doi.org/10.1007/s10554-011-9881-7>
  17. Ko BS, Cameron JD, Meredith IT et al (2012) Computed tomography stress myocardial perfusion imaging in patients considered for revascularization: a comparison with fractional flow reserve. *Eur Heart J* 33:67–77. <https://doi.org/10.1093/eurheartj/ehr268>
  18. Ko BS, Cameron JD, Leung M et al (2012) Combined CT coronary angiography and stress myocardial perfusion imaging for hemodynamically significant stenoses in patients with suspected coronary artery disease. *JACC Cardiovasc Imaging* 5:1097–1111. <https://doi.org/10.1016/j.jcmg.2012.09.004>
  19. Shen D, Wu G, Suk HI (2017) Deep learning in medical image analysis. *Annu Rev Biomed Eng* 19:221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
  20. Shah SJ, Katz DH, Selvaraj S et al (2015) Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 131:269–279. <https://doi.org/10.1161/CIRCULATIONAHA.114.010637>
  21. Abadi S, Roguin A, Engel A, Lessick J (2010) Feasibility of automatic assessment of four-chamber cardiac function with MDCT: initial clinical application and validation. *Eur J Radiol* 74:175–181. <https://doi.org/10.1016/j.ejrad.2009.01.035>
  22. Abbara S, Blanke P, Maroules CD et al (2016) SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: a report of the society of Cardiovascular Computed Tomography Guidelines Committee: endorsed by the North American Society for Cardiovascular Imaging (NASCI). *J Cardiovasc Comput Tomogr* 10:435–449. <https://doi.org/10.1016/j.jcct.2016.10.002>
  23. Leipsic J, Abbara S, Achenbach S et al (2014) SCCT guidelines for the interpretation and reporting of coronary CT angiography: a report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *J Cardiovasc Comput Tomogr* 8:342–358. <https://doi.org/10.1016/j.jcct.2014.07.003>
  24. Cury RC, Abbara S, Achenbach S et al (2016) CAD-RADS™ Coronary Artery Disease – Reporting and Data System. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT), the American College of Radiology (ACR) and the North American Society for Cardiovascular Imaging (NASCI). Endorsed by the American College of Cardiology. *J Cardiovasc Comput Tomogr* 10:269–281. <https://doi.org/10.1016/j.jcct.2016.04.005>
  25. Windecker S, Kolh P, Alfonso F et al (2014) 2014 ESC/EACTS guidelines on myocardial revascularization. *Eur Heart J* 35:2541–2619. <https://doi.org/10.1093/eurheartj/ehu278>
  26. Takx RA, Blomberg BA, El Aidi H et al (2015) Diagnostic accuracy of stress myocardial perfusion imaging compared to invasive coronary angiography with fractional flow reserve meta-analysis. *Circ Cardiovasc Imaging* 8:e002666–e002666. <https://doi.org/10.1161/CIRCIMAGING.114.002666>
  27. Cook CM, Petraco R, Shun-Shin MJ et al (2017) Diagnostic accuracy of computed tomography–derived fractional flow reserve: a systematic review. *JAMA Cardiol* 2:803–810. <https://doi.org/10.1001/jamacardio.2017.1314>
  28. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) *Comput. Vis. – ECCV 2014 13th Eur. Conf. Zurich, Switzerland, Sept. 6–12, 2014, Proceedings, Part I*. Springer International Publishing, Cham, pp 818–833
  29. Sarwar A, Rieber J, Mooyaart EA et al (2008) Calcified plaque: measurement of area at thin-section flat-panel CT and 64-section multidetector CT and comparison with histopathologic findings. *Radiology* 249:301–306. <https://doi.org/10.1148/radiol.2483072003>