University of Groningen

Accuracy of an Artificial Intelligence Deep Learning Algorithm Implementing a Recurrent Neural Network With Long Short-term Memory for the Automated Detection of Calcified Plaques From Coronary Computed Tomography Angiography

Fischer, Andreas M.; Eid, Marwen; De Cecco, Carlo N.; Gulsun, Mehmet A.; van Assen, Marly; Nance, John W.; Sahbaee, Pooyan; De Santis, Domenico; Bauer, Maximilian J.; Jacobs, Brian E.

Published in:
Journal of thoracic imaging

DOI:
[10.1097/RTI.0000000000000491](https://doi.org/10.1097/RTI.0000000000000491)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Fischer, A. M., Eid, M., De Cecco, C. N., Gulsun, M. A., van Assen, M., Nance, J. W., Sahbaee, P., De Santis, D., Bauer, M. J., Jacobs, B. E., Varga-Szemes, A., Kabakus, I. M., Sharma, P., Jackson, L. J., & Schoepf, U. J. (2020). Accuracy of an Artificial Intelligence Deep Learning Algorithm Implementing a Recurrent Neural Network With Long Short-term Memory for the Automated Detection of Calcified Plaques From Coronary Computed Tomography Angiography. *Journal of thoracic imaging*, 35, S49-S57. <https://doi.org/10.1097/RTI.0000000000000491>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Accuracy of an Artificial Intelligence Deep Learning Algorithm Implementing a Recurrent Neural Network With Long Short-term Memory for the Automated Detection of Calcified Plaques From Coronary Computed Tomography Angiography

Andreas M. Fischer, MD,* Marwen Eid, MD,* Carlo N. De Cecco, MD, PhD,*
 Mehmet A. Gulsun, MSc,† Marly van Assen, MSc,*‡ John W. Nance, MD,*
 Pooyan Sahbaee, PhD,§ Domenico De Santis, MD,*||
 Maximilian J. Bauer, BS,* Brian E. Jacobs, BS,*
 Akos Varga-Szemes, MD, PhD,* Ismail M. Kabakus, MD, PhD,*
 Puneet Sharma, PhD,† Logan J. Jackson, BS,* and U. Joseph Schoepf, MD*

Purpose: The purpose of this study was to evaluate the accuracy of a novel fully automated deep learning (DL) algorithm implementing a recurrent neural network (RNN) with long short-term memory (LSTM) for the detection of coronary artery calcium (CAC) from coronary computed tomography angiography (CCTA) data.

Materials and Methods: Under an IRB waiver and in HIPAA compliance, a total of 194 patients who had undergone CCTA were retrospectively included. Two observers independently evaluated the image quality and recorded the presence of CAC in the right (RCA), the combination of left main and left anterior descending (LM-LAD), and left circumflex (LCx) coronary arteries. Non-contrast CACS scans were allowed to be used in cases of uncertainty. Heart and coronary artery centerline detection and labeling were automatically performed. Presence of CAC was assessed by a RNN-LSTM. The algorithm's overall and per-vessel sensitivity, specificity, and diagnostic accuracy were calculated.

Results: CAC was absent in 84 and present in 110 patients. As regards CCTA, the median subjective image quality, signal-to-noise ratio, and contrast-to-noise ratio were 3.0, 13.0, and 11.4. A total of 565 vessels were evaluated. On a per-vessel basis, the algorithm achieved a sensitivity, specificity, and diagnostic accuracy of 93.1% (confidence interval [CI], 84.3%-96.7%), 82.76% (CI, 74.6%-89.4%), and 86.7% (CI, 76.8%-87.9%), respectively, for the RCA, 93.1% (CI, 86.4%-97.7%), 95.5% (CI, 88.7%-98.75%), and 94.2% (CI, 90.2%-94.6%), respectively, for the LM-LAD, and 89.9% (CI, 80.2%-95.8%), 90.0% (CI, 83.2%-94.7%), and

89.9% (CI, 85.0%-94.1%), respectively, for the LCx. The overall sensitivity, specificity, and diagnostic accuracy were 92.1% (CI, 92.1%-95.2%), 88.9% (CI, 84.9%-92.1%), and 90.3% (CI, 88.0%-90.0%), respectively. When accounting for image quality, the algorithm achieved a sensitivity, specificity, and diagnostic accuracy of 76.2%, 87.5%, and 82.2%, respectively, for poor-quality data sets and 93.3%, 89.2% and 90.9%, respectively, when data sets rated adequate or higher were combined.

Conclusion: The proposed RNN-LSTM demonstrated high diagnostic accuracy for the detection of CAC from CCTA.

Key Words: machine learning, coronary artery calcium score, coronary computed tomography angiography, recurrent neural network, convolutional neural network, long short-term memory

(*J Thorac Imaging* 2020;35:S49–S57)

IMPLICATIONS FOR PATIENT CARE

The ability to automatically detect calcium from contrast-enhanced coronary computed tomography angiography (CCTA) has the potential to reduce user variability and standardize results, help radiologists differentiate stent from calcification, and focus on more crucial tasks. Increasing utilization of various artificial intelligence techniques has the potential to improve workflow and/or diagnostic accuracy; identification of coronary artery calcium (CAC) is one step in achieving that goal. Automated CAC detection can help quickly focus diagnosticians on lesions of interest, help interpretation of the coronary artery lumen, and provide a safety net such that subtle lesions are not overlooked, which could be particularly valuable in trainees or in those with limited experience in interpretation of CCTA. Future deep learning (DL) algorithms with the ability to automatically derive calcium score from contrast-enhanced CCTA have the potential to reduce radiation exposure by omitting noncontrast calcium scoring.

IMPORTANT FINDINGS

In this study, we determined the improved accuracy of a novel DL method implementing recurrent neural networks with long short-term memory for the fully automated detection of CAC from contrast-enhanced CCTA data.

From the *Department of Radiology and Radiological Science, Division of Cardiovascular Imaging, Medical University of South Carolina, Charleston, SC; †Medical Imaging Technologies, Siemens Medical Solutions, Princeton, NJ; ‡Siemens Medical Solutions, Malvern, PA; §Department of Radiological Sciences, Oncology and Pathology, University of Rome "Sapienza," Rome, Italy; and ‡Center for Medical Imaging-North East Netherlands, University Medical Center Groningen, University of Groningen, The Netherlands.

U.J.S. has received institutional research support, consulting fees, and/or speaker honoraria from Bayer, Bracco, Elucid, GE, Guerbet, HeartFlow Inc., and Siemens. A.V.-S. receives institutional research support from Siemens. J.W.N. is a speaker for Genentech and Boehringer-Ingelheim. The remaining authors declare no conflicts of interest.

Correspondence to: U. Joseph Schoepf, MD, Department of Radiology and Radiological Science, Medical University of South Carolina, 25 Courtenay Drive, Charleston, SC 29425 (e-mail: schoepf@musc.edu). Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved. DOI: 10.1097/RTI.0000000000000491

INTRODUCTION

CAC is an independent and strong risk predictor of adverse cardiac events and mortality.^{1–5} Although studies have demonstrated the feasibility of accurately detecting calcified lesions and calculating the coronary artery calcium score (CACS) from CCTA acquisitions,^{6–8} a patient's CACS is routinely determined using traditional noncontrast scans.^{9–11} However, calculating CACS is a semiautomatic process that often requires manual intervention by the operator, and thus it is time-consuming because it consists of drawing contours to obtain the region of interest (ROI) or clicking inside all calcium objects.^{12–14} Fully automatic algorithms able to accurately detect calcified lesions and derive CACS from CCTA may prove to be more clinically relevant, as this might obviate the need for a dedicated noncontrast scan while reducing radiation dose, time, and variability of the process.^{15,16}

Machine learning (ML) is a field of computer science whereby computer algorithms are trained to recognize certain features and patterns from input data.¹⁷ In cardiovascular imaging, ML algorithms have been trained to automatically detect obstructive and nonobstructive lesions,^{18–20} calcified lesions,^{21–25} and calculate CACS from noncontrast scans.^{26–31} While effective in many applications, traditional neural networks are not well suited for tasks that require processing of sequential data, that is, where the data exhibit temporal or spatial dynamics such as image and text processing or speech recognition. For such tasks, recurrent neural networks (RNN) have been shown to be more effective. The RNN architecture contains cycles that allow the network to take inputs from both current and previously learned information. Consequently, information can persist in the neural network beyond its initial introduction. In addition, long short-term memory (LSTM), a type of RNN, has the capacity to learn long-term dependencies that may exist within the data.

Recently, a study using convolutional neural networks (CNN) for calcium detection without a need for coronary segmentation was able to achieve an overall sensitivity of 72% for CAC detection.²² However, the accuracy of a DL algorithm implementing RNN-LSTM has not been investigated to date.

Increasing utilization of these techniques for pathology detection has two potential implications in this clinical space: (1) Improved workflow via automated highlighting of atherosclerotic plaque for further analysis, and (2) providing an additional layer of safety to ensure pathology detection, which could be particularly relevant for trainees or for those with limited experience in interpretation of CCTA. It also helps radiologists to differentiate stent from calcification.

Thus, the purpose of this study was to evaluate the accuracy of a novel fully automated DL algorithm implementing a RNN with LSTM for the detection of CAC from CCTA data. Future DL algorithms with the ability to automatically derive calcium score from contrast-enhanced CCTA have the potential to reduce radiation exposure by omitting noncontrast calcium scoring.

MATERIALS AND METHODS

Patient Population

This retrospective single-center study was approved by our institutional review board with a waiver for informed consent and conducted in compliance with the Health Insurance Portability and Accountability Act. The study was carried out in an academic center. In the context of a

random series, a total of 200 patients who had undergone CCTA with noncontrast calcium scoring for either chest pain ($n=92$), suspected coronary artery disease ($n=99$), or preintervention evaluation ($n=9$) were included. Exclusion criteria included nondiagnostic image quality due to heavy motion or metal artifacts, prior history of coronary artery bypass surgery, and anatomic variants such as anomalous courses of the coronary arteries or congenital heart disease.

Image Acquisition

All scans were performed on either second (SOMATOM Definition Flash) or third generation dual-source computed tomography systems (SOMATOM Force, Siemens Healthineers, Forchheim, Germany). All patients underwent non-contrast-enhanced CACS studies followed by a contrast-enhanced CCTA acquisition. Sublingual nitroglycerin and betablocker (Metoprolol) was given before scanning to patients with no contraindication.

Contrast-enhanced CCTA

CCTA parameters on the second generation system included the following (and they are): a prospectively ECG-triggered sequential acquisition, tube current of 380 mAs with automated tube current modulation, temporal resolution of 75 ms, collimation of $2 \times 128 \times 0.6$ mm, automated tube voltage selection (CARE kV, Siemens) ranging from 70 to 120 kV in 10 kV increments, 0.28 s gantry rotation, and a pitch of 1.0. CCTA parameters on the third generation scanner included a prospectively ECG-triggered sequential acquisition, tube current of 300 mAs with automated tube current modulation, temporal resolution of 66 ms, collimation of $2 \times 192 \times 0.6$ mm, automated tube voltage selection ranging from 70 to 120 kV in 10 kV increments, 0.25 s gantry rotation, and a pitch of 1.0.

Contrast enhancement was achieved via Iopromide 370 mgI/mL (Ultravist, Bayer, Wayne, NJ) injection at 0.8 mL/kg body weight with a flow rate of 5.0 mL/s, followed by a 50 mL saline flush. An automated dual-syringe power injector (Stellant D, Medrad Inc., Warrendale, PA) was used for both contrast and saline administration. Weighted filtered back projection was applied for all reconstructions with a section thickness of 0.75 mm at 0.5 mm increments. A smooth convolutional kernel (B26f) was used for all reconstructions. Scan initiation was performed using the bolus-tracking technique (CARE Bolus, Siemens). An ROI was placed within the descending aorta at the level of the carina, with the scan being automatically triggered 4 seconds after a threshold of 100 HU was reached. All scan lengths ranged from the carina to the cardiac apex.

CACS

The corresponding images for calcium scoring were reconstructed with a slice thickness of 3 mm, which is an increment of 1.5 mm; the tube voltage was 120 kVp. Calcium score based on the Agatston method was defined as the presence of a lesion with an area $> 1 \text{ mm}^2$, and peak intensity $> 130 \text{ HU}$, which was automatically identified and marked with color by the software.

Image Quality and Manual Calcium Detection

Data sets were transferred to a dedicated postprocessing workstation for analysis (*syngo.via* VB10B, Siemens Healthineers, Forchheim, Germany). To determine objective image quality, a 5 mm^2 circular ROI was placed in the

mid left anterior descending coronary artery (LAD), left pectoralis major muscle, and subcutaneous fat. For each ROI, HU, and SD were recorded. Image noise was defined as the SD of subcutaneous fat. Each measurement was performed three times and subsequently averaged to ensure consistency of data. Contrast-to-noise ratio (CNR) was calculated using the following formula:

$$\text{CNR} = \frac{\text{HU}_{\text{LAD}} - \text{HU}_{\text{muscle}}}{\text{Image noise}}$$

To assess subjective image quality and calcium detection, the imaging studies were evaluated by 2 independent observers with 4 and 5 years of experience in cardiovascular imaging. Image quality was scored using a 4-point Likert scale with the following criteria: 1 = poor, reduced image quality limited by excessive image noise and/or poor vessel wall delineation and contrast enhancement; 2 = adequate, reduced image quality with high image noise and poor vessel wall delineation and limitations in contrast enhancement; 3 = good, impact of image noise, vessel wall delineation, and contrast condition are minimal; 4 = excellent, image noise not perceivable, optimal vessel wall delineation, and excellent vessel lumen enhancement. The presence of calcium in the right coronary artery (RCA), left main (LM), LAD, and left circumflex (LCx) arteries was recorded by both observers' visual assessment on the axial images with freely adjustable window/level settings. Discrepancies between the 2 observers were resolved by a third reader with 10 years of experience in cardiovascular imaging arbitrated. Non-contrast CACS scans were allowed to be used in cases of

uncertainty. Notably, the LM and LAD (LM_LAD) were merged for all analyses.

Automatic Calcium Detection

A DL algorithm was used for fully automatic coronary artery segmentation and calcified plaque detection. The algorithm first detects the pericardium, followed by the detection of the left and right coronary ostia. Next, it automatically extracts and labels the entire coronary artery centerline tree, as previously described.³² Once the centerline tree has been detected and labelled, a trained RNN with LSTM architecture is used for the automatic detection of calcified plaques (Fig. 1A). The proposed method samples 2D cross-sectional image patches along the coronary centerline, where each patch can be considered a temporal input of the LSTM network. These patches are then connected to LSTM layers through multiple CNN layers (Fig. 1B). The output of the neural network is a binary label denoting the presence or absence of calcification at a particular centerline point (Fig. 2). In addition, bi-directional LSTM was used to exploit image information both upstream and downstream of a particular centerline point.

Algorithm Training

For the initial training and testing phase of the algorithm, that is, the validation step, a separate annotated database of 232 CCTA studies was used; 164 scans were used to train the algorithm, and another 68 were subsequently used for initial performance testing. The database was representative of a diverse set of CCTA studies, all from the same site, however, with varying scanner models, image acquisition

FIGURE 1. A, Cross-sectional samples along the centerline and our neural network with coupled LSTM and CNN. B, CNN component which takes cross-sectional image patches as input and output features for LSTM layer. Network weights are shared among all CNNs in the sequence. ReLU indicates rectified linear unit. [full color online](#)

FIGURE 2. Case example of a CCTA study in a 76-year-old man presenting with chest pain, analyzed by the proposed RNN-LSTM. First, the algorithm automatically computes the centerline tree, and then samples a cross-section image along the tree. Each cross-section is then classified as either containing calcified plaque or not. For representation purposes, a colored triangle was manually placed on a cross section according to its classification by the algorithm.

protocols, and image quality. For each of these studies, coronary artery centerlines and calcified plaque makers (placed at the start and end locations for each calcified plaque along the vessel centerline) were annotated and used as ground-truth to train and test the DL algorithm (Fig. 3). The algorithm was trained and tested in a point-wise manner wherein each point on the centerline was compared with the equivalent ground-truth point with the label. The centerline points were sampled with 0.5 mm resolution. On the testing set, the proposed algorithm (LSTM+CNN) achieved a sensitivity and specificity of 97% and 98%, respectively. The sensitivity and specificity were measured on the basis of the presence of calcification on an overall and per-vessel basis (RCA, LM-LAD, and LCX). The diagnostic accuracy of the CNN-only algorithm (that was trained and tested on the same cases) was 12% lower than the proposed LSTM+CNN algorithm (Fig. 4).

Statistical Analysis

All statistical analyses were performed using statistical software (SPSS for Windows version 24.0, SPSS Inc., Chicago, IL). The normality of data distribution was assessed using the Kolmogorov-Smirnov test. Categorical data were compared using the χ^2 test and presented as frequencies and percentages. Normally distributed continuous data were compared with a 2-tailed student *t* test and presented as mean $\pm$ SD; non-normally distributed data were compared with the Mann-Whitney test and presented as medians with interquartile range (25th percentile to 75th percentile).

As regards the diagnostic performance of the algorithm, overall and per-vessel sensitivity, specificity, and diagnostic accuracy were calculated using the observers' assessment as the reference standard. Data from the LM and LAD were aggregated. To evaluate the accuracy according to image quality, data sets rated as adequate or higher were combined and compared

FIGURE 3. A, Annotated LCA centerlines for a sample case. B, Annotated calcification markers on LAD branch (right). full color online

FIGURE 4. A, Receiver operating curves for LSTM+CNN. B, Receiver operating curves for CNN only (right) methods from 164 training (blue) and 68 testing (red) cases. [full color online](#)

with those rated as poor. Cohen κ statistic was applied to evaluate interobserver subjective image quality agreement and the overall and per-vessel agreement between the algorithm and the observers with the following interpretation: $\kappa \leq 0.20$, poor agreement; $\kappa = 0.21$ to 0.40 , fair agreement; $\kappa = 0.41$ to 0.60 , moderate agreement; $\kappa = 0.61$ to 0.80 , good agreement; and $\kappa > 0.80$, excellent agreement.³³ A 2-tailed P -value < 0.05 was considered statistically significant for all analyses.

RESULTS

Patient Population

Of the 200 studies initially identified, 6 studies were excluded from the final analysis due to the presence of congenital heart disease (single ventricle, $n = 2$; aberrant origin of the LCx arising from the RCA; $n = 1$; dextrocardia, $n = 1$) or bypass surgery ($n = 2$). Patients with stents were included ($n = 2$, 1 in the LAD and 1 in the LCx). We kept the 2 patients with stents in order to evaluate the robustness of our artificial intelligence algorithm against stents. The proposed algorithm accounts for

the sequential image information embedded along the branches and therefore is capable of making decisions using a broader image context. This allows for differentiating calcifications from confounding factors such as stents. Of the 194 included scans, CAC was absent in 84 patients. CAC scans were only used to support the evaluation process performed by radiologists in case of uncertainties. There was no need for CACS scans for evaluating the 84 patients, which were therefore kept in our study cohort. The median subjective image quality was 3 (2 to 4) according to both readers, with excellent agreement ($\kappa = 0.88$) for the entire group. In total, 20 data sets were rated as having poor image quality, while 174 cases were rated as adequate or higher image quality. As regards CCTA, the median signal-to-noise ratio and CNR were 13.0 (9.3 to 20.4) and 11.4 (7.7 to 17.7), respectively.

Calcium Detection

Overall and per-vessel sensitivity, specificity, and diagnostic accuracy results are shown in Table 1. A remaining total of 194 patients and 565 vessels were analyzed. Overall,

TABLE 1. Sensitivity, Specificity, Diagnostic Accuracy, and Intraclass Correlation Coefficient of the Algorithm Results

	RCA (n = 187)	LM-LAD (n = 189)	LCx (n = 189)	Overall (n = 565)
Sensitivity	93.1	93.1	89.9	92.1
Specificity	82.8	95.5	90.0	88.9
Accuracy	86.7 (0.74)	94.2 (0.88)	89.9 (0.78)	90.3 (0.85)

Data are presented as percentage (intraclass correlation coefficient).

the sensitivity, specificity, and diagnostic accuracy achieved by the algorithm were 92.1% (confidence interval [CI], 92.1%-95.2%), 88.9% (CI, 84.9%-92.1%), and 90.3% (CI, 88.0%-90.0%), respectively, with a false-positive (FP) rate of 0.2 per scan. On a per-vessel basis, the algorithm reached a sensitivity, specificity, and diagnostic accuracy of 93.1% (CI, 84.3%-96.7%), 82.76% (CI, 74.6%-89.4%), and 86.7% (CI, 76.8%-87.9%), respectively, for the RCA, 93.1% (CI, 86.4%-97.7%), 95.5% (CI, 88.77%-98.75%), and 94.2% (CI, 90.2%-94.6%), respectively, for the LM-LAD, and 89.9% (CI, 80.2%-95.8%), 90.0% (CI, 83.2%-94.7%), and 89.9% (CI, 85.0%-94.1%), respectively, for the LCx. Overall, excellent agreement was achieved between the algorithm and observers $\kappa=0.85$ ($P \leq 0.001$). On a per-vessel basis, agreement between the algorithm and observers was ranked good for all 3 vessels (all $\kappa \geq 0.74$, all $P \leq 0.001$). When divided on the basis of image quality the algorithm achieved a sensitivity, specificity, and diagnostic accuracy of 76.2%, 87.5%, and 82.2%, respectively, for poor quality data sets and 93.3%, 89.2%, and 90.9%, respectively, when data sets were rated adequate or higher (Table 2). The average automated analysis time was 35 seconds; 20 seconds was needed for centerline extraction and 15 seconds for lesion detection. Analysis was performed on a 2.8 GHz machine with 32 GB of Random Access Memory. A comparison of our results with previous studies is summarized in Table 3.

DISCUSSION

The purpose of our study was to evaluate the accuracy of a DL RNN-LSTM algorithm to automatically detect coronary artery calcified plaques from contrast-enhanced CCTA. Our results, based on imaging studies with varying image quality, demonstrate that the proposed algorithm reaches high overall diagnostic performance in the detection of calcified lesions (overall sensitivity: 92.1%; specificity: 88.9%; diagnostic accuracy: 90.3%) while maintaining good to excellent agreement with expert observers.

Several investigators recently evaluated the feasibility of automated calcified plaque detection and CACS from contrast-enhanced CCTA.^{22,24,30,34,35,37} The previous studies used threshold-based algorithms for the detection of calcium.

Specifically, the algorithms were based on the following (and they are): HU-attenuation deviations within the extracted coronary lumen,³⁰ deviations from a noncalcified artery segmentation model,³⁵ patient-specific HU thresholds,³⁷ or fitting an adaptive intensity distribution model to the lumen attenuation profile.³⁶ While some of these studies focused on CACS without reporting lesion detection accuracy, Ebersberger and Teßmann and colleagues reported a slightly higher sensitivity of 94% with a higher FP rate of 0.9 per scan.^{36,37} Notably, these studies were not ML-based and were programmed to detect calcium on the basis of a fixed or derived adaptive threshold from lumen attenuation intensities. The main advantage of the ML-based models applied in our study is their ability to learn from experience without the need to explicitly program them. This leads to an independent, constantly improving performance in relation to the diagnostic accuracy of the algorithm.

Some studies have investigated the feasibility of fully automated CAC detection and CACS based on ML algorithms. Mittal et al previously investigated the accuracy of 2 combined supervised learning classifiers, a probabilistic boosting tree and a random forest, for the detection of calcified plaques based on 165 CCTA data sets.²⁴ They reported a true detection rate of 70% with 0.1 FP per scan and 81% with 0.3 FP per scan. In our study, our method reached a sensitivity and accuracy of 92.1% and 88.9%, respectively, with 0.2 FP rate per scan. More recently, Wolterink et al²² assessed the feasibility of supervised learning with paired CNNs for the detection of calcified plaques and CACS from CCTA on 200 data sets. The first CNN identified voxels that were likely to be calcium, while the second CNN discriminated between calcium and calcium-like regions. The authors reported a sensitivity of 72% for lesion detection with an FP rate of 0.48. Interestingly, our method achieved a substantially higher sensitivity for overall lesion detection while maintaining a lower FP rate (92% vs. 72% and 0.2 vs. 0.48). Moreover, while the previous method did not require the segmentation of the coronary artery tree, the end-result was a higher FP rate due to the identification of extracoronary calcifications. The method used in our study is able to reduce the rate of FP by automatically segmenting coronary artery centerlines and effectively excluding extracoronary calcium while reducing analysis time. In fact, average analysis time with our method was 35 seconds compared with 81 seconds in the aforementioned method, which required 7 seconds for Bounding Box extraction, 46 seconds for ConvNet1, and 28 seconds for ConvNet2. Furthermore, considering the self-learning capability of the ML algorithm, increasing the number of training data sets can continue improving the accuracy of the algorithm.

In addition, a major strength of this current study may be seen in the algorithm used, that is, an RNN with bi-directional LSTM, which has proven to be a powerful method in computer vision for prediction tasks based on sequential information.³⁸ However, to the best of our knowledge, this specific type of

TABLE 2. Performance of the RNN-LSTM According to Image Quality

		Poor				Adequate or Higher			
		RCA (n = 15)	LM-LAD (n = 15)	LCx (n = 15)	Overall (n = 45)	RCA (n = 153)	LM-LAD (n = 155)	LCx (n = 155)	Overall (n = 463)
Sensitivity	92.1	100	75.0	70.0	76.2	92.9	94.1	92.7	93.3
Specificity	88.9	77.8	100.0	87.5	87.5	82.5	95.8	91.0	89.2
Accuracy	90.3	86.7	86.7	77.8	82.2	86.3	94.8	91.6	90.9

TABLE 3. Comparison of Our Results and Previously Published Results on Automated CAC Detection

References	No. Scans (Training)	No. Scans (Testing)	Type of Algorithm	Sensitivity (%)	FP per Scan
Wolterink et al ²²	250	100	Paired CNN (ConvNets)	72	0.48
Schuhbaeck et al ³⁴	40	44	Threshold based	NA	NA
Ahmed et al ³⁰	NA	100	Threshold based	NA	NA
Eilot and Goldenberg ³⁵	NA	263	Threshold based	94	0.9
Ebersberger et al ³⁶	NA	127	Threshold based	NA	NA
Tefmann et al ³⁷	NA	53	Threshold based	94	0.9
Mittal et al ²⁴	355	165	PBT and RF	70	0.1
Current study	232	194	RNN with LSTM	92	0.2

NA indicates not applicable; PBT, probabilistic boosting tree; RF, random forests.

algorithm has not been investigated in cardiac imaging with the specific purpose of disease detection. Most of the previous work on automatic detection of vascular abnormalities relied heavily on local image features extracted along vascular branches. Thus, previous methods are only able to produce local predictions by disregarding the overall dependency between image features and predictions along the branch. Conversely, the proposed LSTM algorithm accounts for the sequential image information embedded along the entire length of the branches, enabling decision making using a broader image context. In addition, we used a bi-directional LSTM framework to incorporate both the forward and backward observations in the spatial sequence along the vessel centerline. This information is then used to classify the absence/presence of calcified plaque at each point along the detected centerline. During the training and testing phase of the algorithm, incorporating the bi-directional LSTM architecture resulted in a 12% improvement in specificity for detecting calcifications when compared with a traditional CNN-based architecture. Moreover, the algorithm implemented in our

method seems to perform well with confounding factors, particularly in the presence of stents (Fig. 5A), and on low image quality data sets (Fig. 5B). While the performance of the algorithm decreased slightly on image data rated as poor, the overall accuracy remained relatively high (82.2%). Interestingly, most studies do not score the image quality of their data sets or evaluate its impact on their methodology in patients with sub-optimal image quality or stents. To the best of our knowledge, this is the first study to utilize a DL-based method for the fully automated detection of CAC that included such confounding factors. This further emphasizes the strength of the algorithm used in our study. However, little information is available with regard to the performance of automated algorithms in the presence of confounding factors such as low image quality and stents. Thus, confounding factors remain a challenge for automated methods that will need to be addressed in future investigations.

Beyond its retrospective nature, this study has specific limitations that deserve special mention. First, we only

FIGURE 5. A, CCTA study in a 68-year-old man after stent placement for an acute coronary syndrome within the LCx Artery, analyzed by the proposed RNN-LSTM. This case illustrates the ability of the algorithm to differentiate calcification from the stent. B, CCTA in a 59-year-old woman presenting with chest pain. This case illustrates the accurate performance of the neural network in a study rated as low quality by the observers. For representation purposes, a colored triangle was manually placed on a cross section according to its classification by the algorithm.

assessed the overall ability of the algorithm to detect calcified plaque on CCTA. The algorithm's ability and accuracy in synthesizing the number and location of plaques remains unclear. Second, in this initial development of the algorithm, we opted to focus on the accuracy for detecting calcified lesions. Further studies are required to assess the reliability of the algorithm in calculating CACS in comparison with non-contrast-enhanced CACS. Visual evaluation of contrast-enhanced coronary CTA data sets is not a perfect reference standard for CAC detection. Lastly, while the proposed algorithm is capable of a fully automated coronary artery tree segmentation, the accuracy of the segmentation has not been assessed. We acknowledge that patients with congenital heart disease or bypass surgery were excluded, potentially impacting the accuracy of the segmentation process.

In conclusion, the proposed DL algorithm implementing RNN-LSTM demonstrated high diagnostic accuracy for the detection of coronary artery calcified plaques from CCTA. Furthermore, despite the impact of poor image quality on the algorithm's performance, it maintained a remarkable accuracy when applied to low image quality data. It is the initial successful step for future fully automated advanced DL algorithm, which will derive calcium score from CCTA.

REFERENCES

- Nance JW Jr, Bamberg F, Schoepf UJ, et al. Coronary atherosclerosis in African American and white patients with acute chest pain: characterization with coronary CT angiography. *Radiology*. 2011;260:373–380.
- Greenland P, Bonow RO, Brundage BH, et al. ACCF/AHA 2007 clinical expert consensus document on coronary artery calcium scoring by computed tomography in global cardiovascular risk assessment and in evaluation of patients with chest pain: a report of the American College of Cardiology Foundation Clinical Expert Consensus Task Force (ACCF/AHA Writing Committee to Update the 2000 Expert Consensus Document on Electron Beam Computed Tomography). *Circulation*. 2007;115:402–426.
- van Kempen BJ, Spronk S, Koller MT, et al. Comparative effectiveness and cost-effectiveness of computed tomography screening for coronary artery calcium in asymptomatic individuals. *J Am Coll Cardiol*. 2011;58:1690–1701.
- Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N Engl J Med*. 2008;358:1336–1345.
- Schlett CL, Nattenmuller J, Tsuchiya N, et al. Noncontrast chest computed tomographic imaging of obesity and the metabolic syndrome: part I cardiovascular findings. *J Thorac Imaging*. 2019;34:116–125.
- Hong C, Becker CR, Schoepf UJ, et al. Coronary artery calcium: absolute quantification in nonenhanced and contrast-enhanced multi-detector row CT studies. *Radiology*. 2002;223:474–480.
- Tabari A, Gullo RL, Murugan V, et al. Recent advances in computed tomographic technology: cardiopulmonary imaging applications. *J Thorac Imaging*. 2017;32:89–100.
- Caruso D, De Santis D, Biondi T, et al. Half-dose coronary artery calcium scoring: impact of iterative reconstruction. *J Thorac Imaging*. 2019;34:18–25.
- Conti CR. Coronary artery calcium scanning: past, present, and future. *JACC Cardiovasc Imaging*. 2016;9:330–331.
- Mylonas I, Alam M, Amily N, et al. Quantifying coronary artery calcification from a contrast-enhanced cardiac computed tomography angiography study. *Eur Heart J Cardiovasc Imaging*. 2014;15:210–215.
- Ohnesorge BM, Hofmann LK, Flohr TG, et al. CT for imaging coronary artery disease: defining the paradigm for its application. *Int J Cardiovasc Imaging*. 2005;21:85–104.
- Agatston AS, Janowitz WR, Hildner FJ, et al. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15:827–832.
- Hecht HS, Cronin P, Blaha MJ, et al. 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. *J Cardiovasc Comput Tomogr*. 2017;11:74–84.
- de Vos BD, Wolterink JM, Leiner T, et al. Direct automatic coronary calcium scoring in cardiac and chest CT. *IEEE Trans Med Imaging*. 2019;38:2127–2138.
- Opolski MP. Cardiac computed tomography for planning revascularization procedures. *J Thorac Imaging*. 2018;33:35–54.
- Maroules CD, Rajiah P, Bhasin M, et al. Current evidence in cardiothoracic imaging: growing evidence for coronary computed tomography angiography as a first-line test in stable chest pain. *J Thorac Imaging*. 2019;34:4–11.
- Morris MA, Saboury B, Burkett B, et al. Reinventing radiology: big data and the future of medical imaging. *J Thorac Imaging*. 2018;33:4–16.
- Kang D, Dey D, Slomka PJ, et al. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography. *J Med Imaging (Bellingham)*. 2015;2:014003.
- Kang K-W, Chang H-J, Shim H, et al. Feasibility of an automatic computer-assisted algorithm for the detection of significant coronary artery disease in patients presenting with acute chest pain. *Eur J Radiol*. 2012;81:e640–e646.
- Retsen TA, Besser AH, Sall S, et al. Machine learning and deep neural networks in thoracic and cardiovascular imaging. *J Thorac Imaging*. 2019;34:192–201.
- Wolterink JM, Leiner T, de Vos BD, et al. An evaluation of automatic coronary artery calcium scoring methods with cardiac CT using the orCaScore framework. *Med Phys*. 2016;43:2361–2373.
- Wolterink JM, Leiner T, de Vos BD, et al. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal*. 2016;34:123–136.
- Isgum I, Rutten A, Prokop M, et al. Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. *Med Phys*. 2007;34:1450–1461.
- Mittal S, Zheng Y, Georgescu B, et al. Fast automatic detection of calcified coronary lesions in 3D cardiac CT images. In: Wang F, Yan P, Suzuki K, Shen D, eds. *Machine Learning in Medical Imaging*. Berlin, Heidelberg: Springer; 2010:1–9.
- Brunner G, Chittajallu DR, Kurkure U, et al. Toward the automatic detection of coronary artery calcification in non-contrast computed tomography data. *Int J Cardiovasc Imaging*. 2010;26:829–838.
- Takx RA, de Jong PA, Leiner T, et al. Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability. *PLoS One*. 2014;9:e91239.
- Wolterink JM, Leiner T, Takx RA, et al. Automatic coronary calcium scoring in non-contrast-enhanced ECG-triggered cardiac CT with ambiguity detection. *IEEE Trans Med Imaging*. 2015;34:1867–1878.
- Shahzad R, van Walsum T, Schaap M, et al. Vessel specific coronary artery calcium scoring: an automatic system. *Acad Radiol*. 2013;20:1–9.
- Yang G, Chen Y, Ning X, et al. Automatic coronary calcium scoring using noncontrast and contrast CT images. *Med Phys*. 2016;43:2174–2186.
- Ahmed W, de Graaf MA, Broersen A, et al. Automatic detection and quantification of the Agatston coronary artery calcium score on contrast computed tomography angiography. *Int J Cardiovasc Imaging*. 2015;31:151–161.
- Heuvelmans MA, Vonder M, Rook M, et al. Screening for early lung cancer, chronic obstructive pulmonary disease, and cardiovascular disease (the Big-3) using low-dose chest computed tomography: current evidence and technical considerations. *J Thorac Imaging*. 2019;34:160–169.
- Gülsün MA, Funka-Lea G, Sharma P, et al. Coronary centerline extraction via optimal flow paths and CNN path pruning.

- International Conference on Medical Image Computing and Computer-Assisted Intervention. 2016. Springer.
33. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
 34. Schuhbaeck A, Otaki Y, Achenbach S, et al. Coronary calcium scoring from contrast coronary CT angiography using a semi-automated standardized method. *J Cardiovasc Computed Tomogr*. 2015;9:446–453.
 35. Eilert D, Goldenberg R. Fully automatic model-based calcium segmentation and scoring in coronary CT angiography. *Int J Computer Assist Radiol Surg*. 2014;9:595–608.
 36. Ebersberger U, Eilert D, Goldenberg R, et al. Fully automated derivation of coronary artery calcium scores and cardiovascular risk assessment from contrast medium-enhanced coronary CT angiography studies. *Eur Radiol*. 2013;23:650–657.
 37. Teßmann M, Vega-Higuera F, Bischoff B, et al. Automatic detection and quantification of coronary calcium on 3D CT angiography data. *Computer Sc Res Dev*. 2011;26:117–124.
 38. Donahue J, Anne Hendricks L, Guadarrama S, et al. Long-term recurrent convolutional networks for visual recognition and description. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.