View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Groningen

University of Groningen

Quantitative cardiac dual source CT; from morphology to function

Assen, van, Marly

DOI:
10.33612/diss.93012859

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:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Assen, van, M. (2019). Quantitative cardiac dual source CT; from morphology to function. [Groningen]: Rijksuniversiteit Groningen. https://doi.org/10.33612/diss.93012859

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).

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.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 13-11-2019

2

Machine Learning in Cardiac CT: Basic Concepts and Contemporary Data

Gurpreet Singh, Subhi J. Al'Aref, Marly Van Assen, Timothy Suyong Kim, Alexander van Rosendael, Kranthi K. Kolli, Aeshita Dwivedi, Gabriel Maliakal, Mohit Pandey, Jing Wang, Virginie Do, Manasa Gummalla, Carlo De Cecco, James K. Min

Dr. Singh and Dr. Al'Aref have equally contributed to the manuscript.

Published JCCT 2018

ABSTRACT

Propelled by the synergy of the groundbreaking advancements in the ability to analyze high-dimensional datasets and the increasing availability of imaging and clinical data, machine learning (ML) is poised to transform the practice of cardiovascular medicine. Owing to the growing body of literature validating both the diagnostic performance as well as the prognostic implications of anatomic and physiologic findings, coronary computed tomography angiography (CCTA) is now a well-established non-invasive modality for the assessment of cardiovascular disease. ML has been increasingly utilized to optimize performance as well as extract data from CCTA as well as non-contrast enhanced cardiac CT scans. The purpose of this review is to describe the contemporary state of ML based algorithms applied to cardiac CT, as well as to provide clinicians with an understanding of its benefits and associated limitations.

INTRODUCTION

The term "machine learning" defines computer-based algorithms that can effectively learn from data to make predictions on future observations, without being explicitly programmed for a specific task or following pre-specified rules. In this era of "big data", the ability to analyze large datasets and the element of "learning from experience" using data in lieu of a defined rule-based system is making these machine learning (ML) algorithms increasingly useful and popular in various domains. Integration of ML-based predictive analytics within clinical imaging is a natural order of progression wherein developments in cardiovascular imaging now provide high-fidelity datasets that possess more data than those acquired from prior generation scanners. The amalgamation of ML-based algorithms with clinical imaging holds the promise to automate redundant tasks and improve disease diagnoses and prognostication, as well as offer the potential to provide new insights into novel biomarkers associated with specific disease processes.

Cardiovascular disease (CVD) is the leading cause of death with a worldwide estimated mortality rate of 31% in 2015 (1). For assessment of cardiovascular health, cardiac computed tomography angiography (CCTA) is a well-established non-invasive modality. The increasing integration of CCTA in clinical practice can be attributed to a growing body of evidence validating both its efficacy and effectiveness in the assessment and support of decisions related to diagnosis and treatment of coronary artery disease (CAD). In particular, CCTA's diagnostic performance demonstrates a pooled sensitivity and specificity of 98% and 89%, respectively, with a negative predictive value approximating 100%; thus indicating that CCTA can safely exclude obstructive CAD (2). Consequently, CCTA has been successfully implemented in the noninvasive diagnostic workup of patients with suspected CAD in multiple clinical settings (3,4).

Non-contrast coronary artery calcium (CAC) scoring by CT is another method for determining the presence and extent of atherosclerotic cardiovascular disease. CAC has proven to be a robust parameter for cardiovascular risk assessment in landmark trials and as such societal guidelines recommend CAC scoring in asymptomatic patients at low to intermediate risk (5–8). In contrast to CAC, CCTA enables description of the entire atherosclerotic plaque phenotype, including for different types of non-calcified plaque, such as necrotic core, fibro-fatty and fibrous plaque. Recent technological advances also enable extraction of functional information beyond atherosclerotic plaque characterization provided with CCTA. For instance, CT-myocardial perfusion techniques and non-invasive CT-based fractional flow computed (CT-FFR) have been compared with traditional functional imaging techniques such as cardiac magnetic

resonance (CMR) imaging, single photon emission tomography (SPECT), and invasively FFR, illustrating its ability to detect flow-limiting CAD (9–16). In this regard, cardiac CT enables a non-invasive approach to comprehensive evaluation of CAD—from anatomical characterization of atherosclerotic plaque to functional characterization of coronary lesions.

Consequently, the role of cardiac CT imaging in clinical practice is expected to continue to grow following these impressive technological advancements, and current professional societal guidance documents support for CT as a first-line test for patients with suspected CAD (17,18). In only the past year, strong interest has arisen within the cardiovascular imaging community to couple the increasing imaging and clinical data associated with cardiac CT with ML algorithms to determine their potential utility for enhanced assessment of CAD. The introduction of these algorithms in the clinical workflow hold promise for automating cardiac CT across the gamut of its implementation, from optimizing day-to-day workflow to supporting data-informed decisions (**Figure 1**). In this manuscript, we review the current literature on the role of cardiac CT and application of ML-based approaches in CAD.

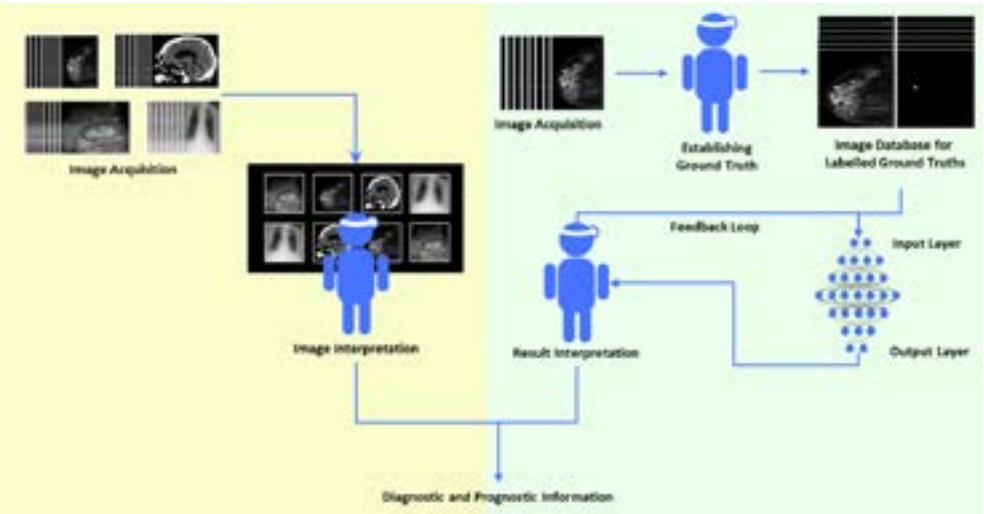

Figure 1: A graphical comparison between the workflow of the traditional and machine learning-based approach for disease diagnosis. The introduction of artificial intelligence-based models could improve diagnostic accuracy and help automate redundant tasks associated with radiologic image interpretation.

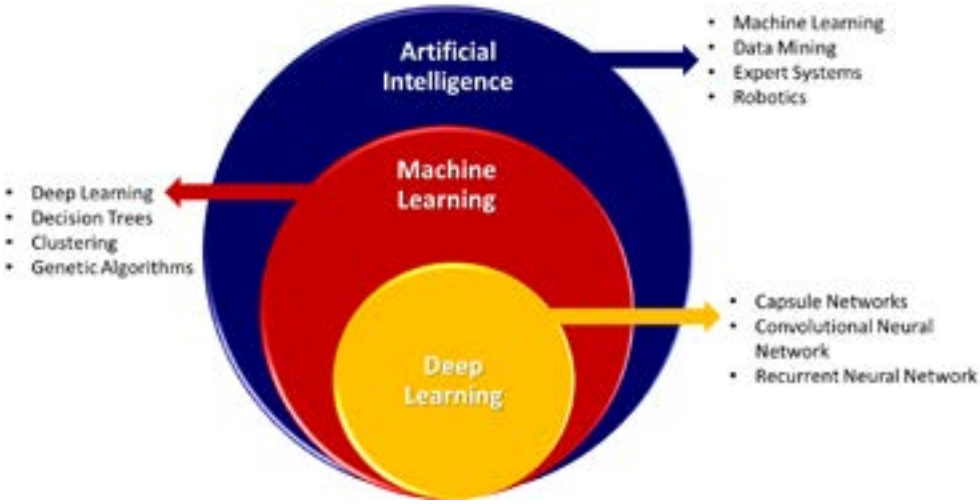

Figure 2: Hierarchy and subfields of artificial intelligence.

An overview of Machine Learning (ML)

ML is a subfield of artificial intelligence (AI) with a primary focus on developing predictive algorithms through unbiased identification of patterns within large datasets and without being explicitly programmed for a particular task. **Figure 2** shows the hierarchy and subfields of artificial intelligence. Based on the task, ML models can be broadly categorized as (19):

A. **Supervised learning:** For supervised learning-based tasks, the model is presented with a labeled dataset also known as feature vectors (i.e. the dataset that contains the examples of observations), as well as their corresponding expected output labels. The goal of such models is to generate an inferred function that maps the feature vectors to the output labels. Some of the most notable supervised learning-based approaches are Support Vector Machines, Linear Regression, Random Forest, Decision Trees, and Convolutional Neural Networks.

B. **Unsupervised learning:** For unsupervised learning, the dataset does not contain information about the output labels. Instead, the goal of these models is to derive the relationship between the observations and/or reveal the latent variables. Some of the most notable unsupervised learning-based approaches are k-means, Self-Organizing Maps, and Generalized Adversarial Networks (GANs).

Traditional ML-based approaches typically require feature extraction to select relevant representation of features before a model for the specific task can be developed, such as Support Vector Machines, Logistic Regression etc. Selecting features is at the heart of developing better models since they directly influence their performance. However, sometimes the relationship between the combinations of the features might be highly

multi-dimensional, non-linear, and/or difficult to comprehend in its entirety. This limits the performance and application of these models.

Recently, methods based upon deep learning, a subfield of ML, have gained much attention owing to the ability of these model architectures to extract features and predict an outcome using raw data. Neural networks (NN) and in particular convolutional neural networks (CNN), a type of deep learning model architecture, are specifically suited for image analysis. NN are inspired by biological neural networks. They are able to model complex relationships between input, output, and pattern recognition, rendering it both germane and useful for image analytics. While these model architectures have been around since the early 20th century, their modern-day resurgence is mostly attributed to the development of LeNet for optical character recognition and then AlexNet in 2012 that won the ImageNet challenge by a large margin (20). The evidence supporting the strength of CNN for image-related tasks kicked off a new era for computer vision that continues to have a positive impact on every aspect of our daily lives.

As an example, deep learning for image analysis has shown great potential and lends itself to implementation in large-scale commercial application. This potential has attracted the financial interest of many private companies that are seeking to implement ML based image analysis into proprietary software. Examples of CNN implementations are traffic sign recognition, vehicle classification, and face recognition (21–23). CNNs also are playing an increasingly large role in medical image analysis. In particular, image segmentation is a field of interest that can be applied to the precise isolation of organs on images—including lungs, brain, bones—as well as pathologic abnormalities within them, such as tumors (24–29). Beyond segmentation alone, ML based classification algorithms are also being applied in medical image analysis. Some examples of this application include the identification, detection, and diagnosis of tumors in different parts of the body such as breast, lung, brain and colon cancer and (early) diagnosis of Alzheimer's disease (30–34).

One very interesting potential application of deep learning-based models is to expand the potential of extracting more knowledge from radiological imaging datasets. This high-throughput extraction of imaging features and the use of this information for precision medicine is known as radiomics. Radiomics based analysis can use the full scale of the imaging knowledge derived from multiple patients imaged at different time points by treating the images as data points. This analysis may be synergistic with other genomics, proteomics, and other clinical findings that may improve decision making and provide individualized therapies (35–37). This application of deep learning algorithms to medical images should be considered decidedly nascent in its development, and there is still a need for establishing standardized evaluation and

reporting guidelines for radiomics. Nevertheless, ML-based algorithms are gradually integrating into clinical practice, including radiologic image interpretation, and an understanding of the evaluation metrics is of paramount importance.

Performance metrics

The effectiveness of a ML model is fundamentally dependent upon the choice of the performance metric. There exist different performance metrics for different ML tasks such as classification, regression etc. Classification is a task of predicting discrete prediction labels (class) given an input data. Classification tasks can be either binary-class (two labels) or multi-class (more than two labels) and the performance metrics for a binary class can be extended to a multi-class task. Typically, in a classification problem, the same performance metric is applied both during the training phase and the testing phase. The performance metric in the training phase is generally used to optimize the classifier (classification algorithm) for an accurate prediction of the future observations. However, in the testing phase, the performance metric is used as a measure of the effectiveness of classifier when tested on unseen test data. The commonly used classification metrics are described below:

1. *Accuracy*: Accuracy is one of the most widely used performance metrics in ML applications. It is defined as the proportion of correct predictions the classifier makes relative to the total size of the dataset. Additionally, error rate (misclassification rate) is a complementary metric of accuracy that evaluates the classifier by its percentage of incorrect predictions relative to the size of the dataset.
2. *Confusion matrix*: Although accuracy is a straightforward metric, it makes no distinction between the classes, i.e., in a binary class problem the correct predictions for both the classes are treated equally. Thus, in the case of unbalanced datasets, relying solely on accuracy could be misleading. For example, for the task of binary classification, with the ratio of the number of samples for the two classes as 9:1, even if a classifier is biased (overfitted) towards the class with larger number samples, it will still have an accuracy of 90% even if it wrongly predicts all the samples of the other class. A confusion matrix (or confusion table) addresses this issue by displaying a more detailed breakdown of correct and incorrect classifications for each class. It is a two by two table that contains four outcomes produced by a binary classifier. The rows of the matrix correspond to ground truth labels, and the columns represent the prediction. Moreover, various other measures like error-rate, accuracy, specificity, sensitivity, and precision can be derived from the confusion matrix.
3. *Log-Loss*: Logarithmic loss (Log-loss) is a performance metric that is applicable when the output of a classifier is a numeric probability instead of class labels. Log-loss is the cross-entropy between the distribution of the true labels and the predictions. Entropy is a measure of unpredictability. Cross-entropy incorporates

the entropy of the true distribution with the additional unpredictability when one assumes a different distribution than the true distribution. Thus, log-loss can also be interpreted as an information-theory based measure to gauge the "extra noise" that comes from using a predictor as opposed to the true labels. Hence, by minimizing the cross entropy, one maximizes the accuracy of the classifier.

4. *AUC*: The area under the Receiver operating curve (AUC-ROC) shows the sensitivity of the classifier by plotting the rate of true positives to the rate of false positives. It shows how many correct positive classifications can be gained as one allows for more and more false positives. The perfect classifier that makes no mistakes would hit a true positive rate of 100% immediately, without incurring any false positives—this almost never happens in practice.
5. *Precision and recall:* Precision is the fraction of examples predicted as positive that are positive. Recall is the fraction of the true positives that are predicted as positives. These measures are trivially maximized by not predicting anything, or predicting everything, respectively, as positive.
6. *DICE Coefficient:* This coefficient measures the degree of similarity between two sets. It is typically used for evaluating the tasks of image segmentation. It is a pixel-wise measure of the degree of similarity between the predicted mask and the labeled ground truth. Mathematically it is represented as follows

$$\text{DICE Score} = \frac{Correctly\ predicted\ pixels}{Total\ pixels\ in\ the\ ground\ truth + Incorrectly\ predicted\ pixels}$$

For a regression task, the model learns to predict numeric scores. The most commonly used metric for regression tasks is root-mean-square-error (RMSE) which is defined as the square root of the average squared distance between the actual and predicted score. RMSE is the most common metric for regression, but since it is based on an average it is always sensitive to large outliers. Thus, if the regressor performs badly on a single data point, the average error could be very big. In such situations, however, quartiles are much more robust as they are not affected at all by the outliers. Median absolute percentage error (MAPE) is one such metric that is a relative measure of error. These metrics (or similar) are typically found in studies reporting ML based analysis. In the following section, we will discuss the present applications of ML in cardiac CT imaging (See **Table 1**).

APPLICATIONS OF ML FOR CARDIAC CT IMAGING ANALYSIS

Automation of coronary artery calcium score measurement

Coronary artery calcium (CAC) scoring is an independent measure as well as a strong risk predictor of adverse cardiac events and mortality (38–43). The amount of CAC in a particular individual can be quantified using the Agatston scoring method, applied to low dose ECG-gated coronary computed tomography (CT) images. The CAC Agatston score increases when either the calcification volume or density increases (44). CAC score has demonstrated strong predictive value for the occurrence of future cardiovascular events, independent from traditional cardiovascular risk factors (45). In addition, a CAC score of 0 is associated with excellent outcomes at very long-term follow up (46). Especially among patients at intermediate cardiovascular risk, the CAC score significantly improves risk stratification and is generally used to tailor medical therapy (45). Based on the CAC scores, patients are assigned to different cardiovascular risk categories and corresponding treatment plans (42,43,47). Measurement of CAC score requires manual placement of regions of interest around all coronary plaques, for every CT slice that covers the coronary vasculature. Manual CAC score measurement is time-consuming, especially when artifacts, image noise, and numerous calcifications are present. Further, this process is often sensitive to interrater variability due to required, time-consuming manual adjustments. Furthermore, separating coronary calcium from adjacent calcified structures (for instance mitral annular calcification and calcification in the left circumflex coronary artery (LCx)) can be challenging when using non-contrast enhanced CT images. In this regard, an automated CAC quantification would be valuable, especially in a large volume screening settings. Using ML to fully automate this task may reduce the time and variability of the process, ultimately improving the clinical workflow and accuracy.

The feasibility of a supervised ML approach to automatically identify and quantify coronary calcifications was demonstrated by Wolterink et al. using 914 scans (48). Patient-specific centerlines of the three coronary arteries were estimated using 10 manually annotated contrast-enhanced CT as the gold standard. Subsequently, 'candidate' calcifications were created based on size, shape, intensity and location characteristics. For instance, candidate calcifications were defined to be between 1.5 and 1500 mm^3. Finally, a classification algorithm allocated candidate calcifications to the specific coronary artery. Lesions that could not be classified with high certainty were presented for expert review. High intra-class correlation coefficients were achieved between expert assessment of CAC volume and the automatic algorithm: 0.95 for all coronary arteries and 0.98, 0.69 and 0.95 for left anterior descending (LAD), LCx and right coronary artery (RCA) respectively. Išgum et al. showed an automated method for coronary calcium detection for the automated risk assessment of CAD on

non-contrast, ECG-gated CT scans of the heart. They reported a detection rate of 73.8% of coronary calcification. After a calcium score was calculated, 93.4% of patients were classified into the correct risk category (40). In another study, they also used ML to measure aortic calcification (compared versus manual assessment) and reported a very high correlation coefficient of 0.960, which was similar to the correlation between two expert observers (R = 0.961) (49). Brunner at al. used a coronary artery region (CAR) model for the detection of CAC, which automatically identifies coronary artery zones and sections (50). The proposed CAR models detected CAC with a sensitivity, specificity, and accuracy of 86%, 94%, and 85%, respectively, compared to manual detection.

Although previous studies used dedicated calcium scoring scans, calculation of CACS has proven to be feasible in non-cardiac scans as well, e.g. non-gated chest CT acquisitions. One example by Takx et al. applied a machine learning approach that identified coronary calcifications and calculated the Agatston score using a supervised pattern recognition system with k-nearest neighbor and support vector machine classifiers in low-dose, non-contrast enhanced, non-ECG-gated chest CT within a lung cancer screening setting (51). In this study, the authors demonstrated the ability of ML to quantify CAC from lower quality images than a dedicated CAC score scan. For instance, among 1793 chest CT scans, the median difference between expert assessment and the automated CAC measurement was 2.5 (interquartile range (IQR): 0-53.2) for Agatston CAC score and 7.6 (IQR: 0-94.4) for CAC volume. When dividing the CAC score into conventional risk groups (0, 1-10, 11-100, 101-400 and >400) the proportion of agreement was 79.2%. They found that the fully automated CAC scoring was feasible with acceptable reliability and agreement; however, the amount of calcium was underestimated when compared to reference scores determined from dedicated CAC score acquisitions (51).

Several studies have demonstrated the feasibility of detecting calcification on CCTA acquisitions (52,53). The use of CCTA images could eliminate the need for a dedicated non-contrast scan, thereby reducing radiation dose. For example, Mittal et al. detected calcified plaques on CCTA images using two combined supervised learning classifiers, a probabilistic boosting tree, and random forest. They reported a true detection rate of calcium volume of 70% with a 0.1 false positive detection per scan, and 81% with a 0.3 false positive detection per scan. However, they only performed calcium detection in the main coronary vessels, not taking into account any side branches (54). Using a combination of contrast-enhanced and non-contrast images, Yang et al. were able to detect CAC with a sensitivity of 99%. The contrast-enhanced scans were used to determine the region of interest for the support vector machine in the detection of calcification on the non-contrast images, thereby excluding calcifications from the surrounding areas (55). Utilizing CCTA images only, Wolterink et al. showed that

coronary calcium can be automatically identified with accuracy and quantified using a ML approach with paired convolutional neural networks (56). Excellent agreement was achieved between CCTA and non-contrast enhanced acquisitions; 83% of patients were assigned to the correct risk category. Analysis of CAC scores performed on non-contrast acquisitions and CCTA images showed similar detection rates and sensitivity; however, the wide range of accuracy parameters makes direct comparison difficult.

Miscellaneous applications

Beyond simple coronary calcium scoring, recent investigations have attempted to evaluate the feasibility of deriving additional coronary artery disease measures from non-contrast CT. As an example, Mannil et al., in a proof-of-concept retrospective study, combined texture analysis and ML to detect myocardial infarction (MI) on non-contrast enhanced low dose cardiac CT images (60). The study included a total of 87 patients, of which 27 patients had acute MI, 30 patients had chronic MI and 30 patients had no cardiac abnormality (controls). A total of 308 texture analysis (TA) features were extracted for each free-hand region of interest (ROI). Feature selection was performed on all the TA features using intra-class correlation coefficient (ICC). Texture features were classified using 6 different classifiers in two approaches: (i) Multi-class model (I): acute MI vs. chronic MI vs. controls and (ii) Binary class model (II): cases (acute and chronic) vs. controls. This proof-of-concept study indicates that certain TA features combined with ML algorithms enable the differentiation between controls and patients with acute or chronic MI on non-contrast-enhanced low radiation dose cardiac CT images.

Quantification of epicardial and thoracic adipose tissue

The amount of fat surrounding the heart has been proven to correlate with an increased cardiovascular risk (66). An automated approach for the quantification of epicardial fat could help assess cardiovascular risk while reducing the time of manual measurements, thereby increasing the clinical applicability. Rodrigues et al. proposed a methodology in which features related to pixels and their surrounding area is extracted from standard CAC scoring acquisitions, and a data mining classification algorithm is applied to segment the different fat types. In this study, the mean accuracy for the epicardial and mediastinal fat was 98.4%, with a mean true positive rate of 96.2% and a DICE similarity index of 96.8% (67). In a previous publication, several classification algorithms, including NN, probabilistic models, and decision tree algorithms, were evaluated for automated fat quantification. They found that decision tree algorithms provided much better performance over NN, function-based classification algorithms and probabilistic models with a DICE similarity index equal to 97.7% (68).

Table 1: Key publications highlighting the application of ML in cardiac CT imaging

Authors	Objectives and Key Findings	Algorithm/Tool Used	Number
Motwani et al. (57)	Prediction of 5-year all-cause mortality (ACM) from clinical and CT variables **ML-AUC =0.79 vs. SSS-AUC=0.64**	LogitBoost	10,030 Patients
Otaki et al. (58)	Prediction of impaired myocardial blood flow from clinical and imaging data (EFV) **ML-AUC=0.73 vs. EFV-AUC=0.67**	LogitBoost	85 Patients
Han et al. (59)	Prediction of ischemia from CCTA variables and CT perfusion **AUC (CTP+CT stenosis) =0.75 vs. AUC (CT stenosis) = 0.67**	Gradient Boosting Classifier	250 Patients
Mannil et al.(60)	Detection of myocardial infarction (MI) on low dose cardiac CT images using texture analysis and machine learning **Model I: KNN-AUC =0.77. Model II: LWL-AUC =0.78**	Decision tree, KNN, Random Forest, ANN, locally weighted learning (LWL) and sequential minimal optimization (SMO)	87 Patients
Freiman et al. (61)	Improving hemodynamic assessment of stenosis by accounting for PVE in lumen segmentation **AUC-segmentation with PVE=0.80. AUC-segmentation without PVE=0.76**	ML based graph-cut segmentation	115 Patients
Wolterink et al. (56)	Automating CAC Scoring. Best performing ConvPair took was 46 s for ConvNet1 and 28 s for ConvNet2 to predict CAC Scores **MSCCTA vs. MSCSCT: Pearson Correlation = 0.950 and ICC = 0.944**	Convolutional Neural Networks	100 Scans
Xiong et al. (62)	Prediction of coronary stenosis with reference to an invasive QCA ground truth standard. **AdaBoost: accuracy= 0.70, sensitivity = 0.79, and specificity = 0.64**	Principal Component Analysis, AdaBoost, Naive Bayes, Random Forest	140 Images
Dey et. al. (63)	Integrated ML ischemia risk score for the prediction of lesion-specific ischemia by invasive FFR **AUC (CT+ Stenosis+LD-NCP+Total Plaque Volume) = 0.84 vs. AUC (stenosis)= 0.76 vs. AUC (LD-NCP volume)= 0.77 vs. AUC (total plaque volume) =0.74**	LogitBoost	80 Patients
Kang et. al. (64)	Detection of nonobstructive and obstructive coronary plaque lesions **Accuracy = 0.94, sensitivity = 0.93, specificity = 0.95, AUC = 0.94**	Support Vector Machines	42 Patients
Commandeur et. al. (65)	Quantification of Epicardial and Thoracic Adipose Tissue from Non-Contrast CT. **Mean DSC: 0.823 (IQR): 0.779-0.860) and 0.905 (IQR: 0.862-0.928) for Epicardial adipose tissue (EAT) and Thoracic adipose tissue (TAT), respectively**.	Convolutional Neural Networks	250 Patients

Abbreviations: ML: machine learning, AUC: area under the curve, SSS: segment stenosis score, EFV: epicardial fat volume, CTP: CT perfusion, KNN: K nearest neighbor, PVE: partial volume effect, CAC: coronary artery calcium, MSCCTA: multislice CCTA, CSCT: cardiac calcium scoring CT, ICC: intraclass correlation, QCA: quantitative coronary angiography, FFR: fractional flow reserve, LD-NCP: low density non-calcified plaque, DSC: Dice score coefficients, IQR: interquartile range.

Similar results were reported for a different method using a CNN approach for fully automated quantification of epicardial and thoracic fat volumes from non-contract CT acquisitions. Strong agreement between automatic and expert manual quantification was shown for both epicardial and thoracic fat volumes with a DICE similarity index of 82% and 91%, respectively; along with excellent correlations of 0.924 and 0.945 to the manual measurements for epicardial and thoracic fat volumes (65).

In another study, Otaki et al. combined clinical and imaging data to explore the relationship between epicardial fat volume (EFV) from non-contrast CT and impaired myocardial blood flow reserve (MFR) from PET imaging (58). The study population comprised of 85 consecutive patients without a previous history of CAD who underwent rest-stress Rb-82 positron emission tomography (PET) and subsequently referred to invasive coronary angiography (ICA). A boosted-ensemble algorithm was used to develop a ML based composite risk score that encompassed variables like age, gender, cardiovascular risk factors, hypercholesterolemia, family history, CAC score, and EFV indexed to body surface area to predict impaired global MFR by PET. In the evaluated risk factors, using multivariate logistic regression, the authors' report that only EFV indexed to body surface was shown to be an independent predictor of impaired MFR. The ML based composite risk score was found to significantly improve risk reclassification (AUC = 0.73) of impaired MFR when compared to multivariate logistical regression analysis of risk factors (AUC = 0.67 for EFV, 0.66 for CAC score). This study thus showed that a combination of risk factors and non-invasive CT-based measures including EFV could be used to predict impaired MFR by PET.

In summary, for non-contrast CT, ML approaches for detection and quantification of CAC scores have been thoroughly investigated. Given the prognostic value of the CAC score, accurate identification of coronary calcification from gated and non-gated chest CT (not specifically performed to assess coronary calcium) is important (6). Additionally, accurate epicardial fat quantification is achievable and could represent a new quantitative parameter that can potentially be implemented in patient risk assessment, similar to CAC score. Automated ML can maximize information extraction from chest CT scans and may eventually improve cardiovascular risk assessment and subsequently patient's outcome.

Coronary Computed Tomographic Angiography (CCTA)

Often obtained in tandem with the CAC score, CCTA has been established as a reliable imaging modality in patients with stable or atypical symptoms requiring noninvasive assessment of the coronary arteries (10,69,70). CCTA allows direct evaluation of the entire coronary artery tree for the presence, distribution, and extent of atherosclerotic plaque. Finer atherosclerotic plaque analyses have expanded to include atherosclerotic

plaque characterization, ranging from the determination of calcification extent (i.e. presence of non-calcified (NCP), partially calcified (PCP) or calcified plaque (CP)) to the presence of CCTA-features that have been associated with the presence of high-risk plaque (i.e. napkin ring sign, low attenuation plaque, spotty calcification and positive remodeling) (71–73). However, such measurements require subjective visual interpretation of images and are thus subject to high inter-observer variability and a high rate of false-positive findings, which can lead to unnecessary downstream testing and increased overall costs (74).

As such, ML has been extensively used for the optimization of information extraction from CCTA, specifically to generate algorithms that can perform plaque analyses in an automated, accurate, and objective manner. Utilizing a two-step ML algorithm which incorporated support vector machine, Kang and colleagues were able to automatically detect non-obstructive and obstructive CAD on CCTA with an accuracy of 94% and an AUC of 0.94 (75). Utilizing a combined segmentation-classification approach, Dey et al. developed an automated algorithm (AUTOPLAQ) for the accurate volumetric quantification of NCP and CP from CCTA (76). Only requiring as input a region of interest in the aorta defining the "normal blood pool", their software was able to automatically extract coronary arteries and generate NCP and CP volumes correlating highly with manual measurements obtained from the same images (R = 0.94 and R = 0.88, respectively).

Plaque Segmentation for Physiologic Characterization of CAD

Hell et al. utilized AUTOPLAQ to derive the contrast density difference (CDD), defined as the maximum percent difference of contrast densities within an individual lesion, which they hypothesized could help predict the hemodynamic relevance of a given coronary artery lesion (77). They found that CDD was significantly increased in hemodynamically relevant lesions (26.0% vs. 16.6%; $p = 0.013$) and at a threshold of ≥ 24% predicted hemodynamically significant lesions with a specificity of 75% and negative predictive value of 73%, as compared to invasive FFR. In a multicenter study of 254 patients with CCTA, Dey et al. inputted a number of AUTOPLAQ-derived image features into a LogitBoost algorithm to generate an integrated ischemia risk score and predict the probability of low value by invasive FFR (63). ML exhibited higher AUC (0.84) compared with any individual CCTA image measurement, including stenosis severity (0.76), low-density NCP (0.77), and total plaque volume (0.74). ML has thus demonstrated value in aiding classification of atherosclerotic lesions identified by rapid non-invasive CCTA imaging analysis as functionally significant (low invasive FFR).

Han et al. reported a different ML approach that aimed at improved prediction of ischemia through an analysis of CCTA variables integrated with CT Myocardial Perfusion Imaging

(CTP) (59). The study population comprised of 252 stable patients with suspected CAD from the DeFACTO study, who underwent clinically indicated CCTA and ICA (78). Using a previously validated custom software (SmartHeart; Weill Cornell Medicine, New York, USA), the myocardium was mapped and subdivided into 17-segment AHA model (62,79). A total of 51 features were extracted per heart, with three features for each of the 17 segments: normalized perfusion intensity (NPI), transmural perfusion intensity ratio (TPI), and myocardial wall thickness (MWT) (59). CCTA-based stenosis characterization, location, and quality were combined with perfusion mapping model variables to demonstrate ischemia (validated by invasive FFR). The results suggest that the addition of CTP data to CCTA stenosis characterization increased the predictive ability to detect ischemia over each set of variables alone.

CT-FFR enables the evaluation of the hemodynamic significance of coronary artery lesions using a non-invasive approach. There are two main approaches to calculate CT-FFR: one uses computational fluid dynamics while the other uses a ML approach (14,80,81). The ML approach that has been tested in clinical practice uses a multilayer NN, trained to comprehend the relationship between coronary anatomy and coronary hemodynamics. The training set for this algorithm consists of a large database of synthetically generated coronary trees, and the hemodynamic parameters are calculated using computational fluid dynamics. The algorithm uses the learned relationship to calculate the ML–based CT-FFR values. In a retrospective analysis, Renker et al. evaluated CT-FFR on a per-lesion and per-patient basis, resulting in the following outcomes: a sensitivity of 85% and 94%, a specificity of 85% and 84%, a positive predictive value of 71% and 71%, and a negative predictive value of 93% and 97%, respectively, with an AUC of 0.92 (82). Coenen et al. reported similar diagnostic performance in two prospective studies with a sensitivity of 82-88%, specificity of 60-65%, and an accuracy of 70-75% compared to invasive FFR (83,84). Similarly, Yang et al. showed per-vessel sensitivity and specificity of 87% and 77%, respectively, with an AUC of 0.89(64); and Kruk at al. showed a per vessel AUC of 0.84 with corresponding sensitivity and specificity of 76% and 72%, respectively (85).

Coronary Plaque Characterization by Machine Learning and Prognostication of Outcomes

ML has also shown promise in its ability to prognosticate cardiovascular outcomes with the combination of clinical and imaging data. Hell et. al performed a case-control study investigating AUTOPLAQ-derived quantitative plaque characteristics for the prediction of incident cardiac mortality during a 5-year period following CCTA (86). The authors found that higher per-patient NCP, low-density NCP, total plaque volumes, and CDD were associated with increased risk of death, even after adjustment with segment involvement score (SIS). Motwani et al. recently utilized raw data from

the Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter (CONFIRM) registry, comprising 10,030 patients with suspected CAD and 5-year follow-up, to investigate the feasibility and accuracy of ML to predict 5-year all-cause mortality (ACM) in patients undergoing CCTA (57). Beginning with more than 60 clinical and CCTA parameters available for each patient, the authors utilized automated feature selection to ensure only parameters with appreciable information gain (information gain > 0) were used for model building. These selected parameters were subsequently inputted into an iterative Logit-Boost algorithm to generate a regression model capable of calculating a patient's 5-year risk of ACM. ML exhibited higher AUC (0.79) compared with the Framingham Risk Score (0.61) or CCTA severity scores alone (segment stenosis score: 0.64, SIS: 0.64, modified Duke Index: 0.62; $p < 0.001$) in the prediction of 5-year ACM. This study elegantly captures the power of ML to not only analyze vast amounts of data, which easily exceeds the analytic capacity of the human brain, but also use this ability to produce clinically meaningful predictive models which may outperform those in current use.

DISCUSSION

Recent Advances in ML Application in Cardiovascular Imaging

ML in medical imaging is considered by many to represent one of the most promising areas of research and development (87). There are numerous recent publications utilizing ML algorithms that either automate the processes or improve diagnostic performance of cardiovascular imaging. The ability of a ML based system to analyze high-dimensional raw images and produce valuable clinical information without human input holds a tremendous potential in clinical practice. Freiman et al.(61), in an attempt to automate coronary measurement using ML, employed an existing coronary lumen segmentation algorithm to account for the partial volume effects (PVE) in the hemodynamic assessment of coronary stenosis. Lumen segmentation was initially automatically evaluated and then corrected by an expert observer. A K-nearest neighbor (KNN) algorithm was used for ML based likelihood estimation within a graph min-cut framework for coronary artery lumen segmentation. The algorithm was also given an additional input in the form of an intensity profile from the PVE evaluation using an ML graph min-cut segmentation algorithm. This enhancement of accounting for PVE improved the AUC for detection of hemodynamically significant stenosis from 0.76 to 0.80. However, it should be noted that the improvement in AUC did not reach statistical significance ($p=0.22$).

In line with the above, more recently, Yi et al. proposed a sharpness-aware generative adversarial network (SAGAN) for low-dose CT de-noising (88). The proposed ML/deep learning network is based on the Generative Adversarial network theory proposed by Goodfellow et al. where a generative model (G) tries to generate real-world images by employing min-max optimization framework and are pitted against a discriminator (D) that distinguishes between real and generated images (89) . While utilizing SAGAN, the authors used three networks namely the generator (G), the discriminator (D) and sharpness detection network (S). The generator (G) utilized a U-net style segmentation network, the discriminator differentiated patches in the image rather than the full image itself and the sharpness detection network used local binary pattern to quantize local sharpness (sharpness loss) in low-contrast regions . The authors report that the newly proposed SAGAN network achieves improved performance in the quantitative assessment of low-dose CT images. Although ML can be a powerful tool for image analysis, it is also subject to some limitations as discussed in the following sections.

Pitfalls / Limitations of ML

The accuracy of ML algorithms is highly dependent on the amount and quality of the input data. For example, with a multi-layer approach using many parameters for image analysis, CNNs need large amounts of data to make accurate predictions. The use of different acquisition protocols for the accurate training of the algorithm will also increase the needed number of cases. A second issue that comes into play concerns the use of large imaging databases, assuring the quality and consistency of the data, and the corresponding output label given as the reference standard. In the field of medical imaging, inter and intra rater variability plays an important role and can represent a significant source of biases. Thus, to increase the accuracy of ML algorithms, there should be a focus on creating a consistent and reliable ground truth to train the algorithm to take into consideration different experts' opinions on what constitutes the ground truth. Every algorithm is limited by the quality of the ground truth that is being used to train and test the algorithm. This can cause problems for accurate training especially when the ground truth is subjective, e.g. expert opinion, or is subjective to high interrater variability (90). To mitigate the need for labeling a large number of images, some of the new CNN architectures, such as U-Net, have been built that train well on a low number of images (91). **Figure 3** shows a simplified representation of CNN-based architecture for image segmentation. Furthermore, another recent deep learning architecture called GANs has recently been proposed to mimic the distribution of data (89). These networks are an active area of research and in the near future could altogether eliminate the need for manual image annotation.

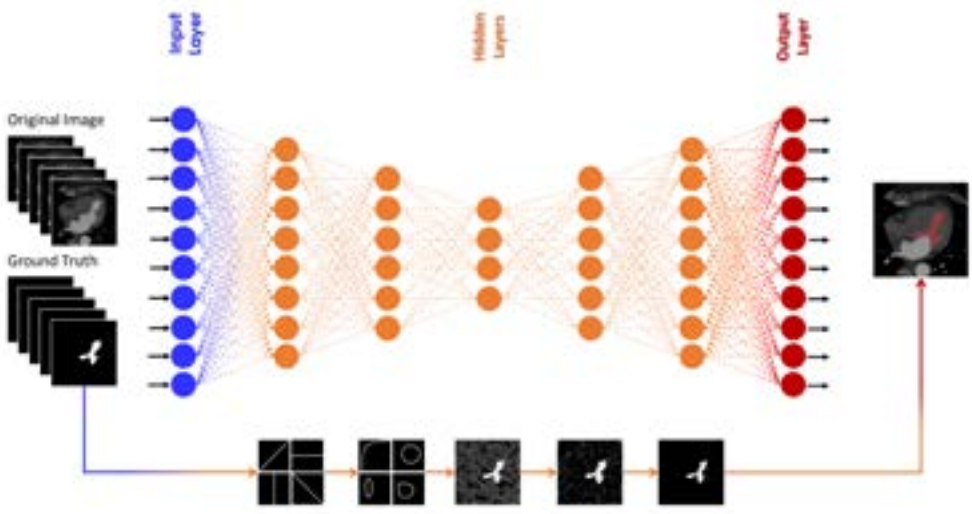

Figure 3: A simplified representation of convolutional neural network-based architecture for image segmentation. Typically, these models have three distinct regions: (1) Input layer, (2) hidden layer and (3) output layer. This figure shows a representation of the features learned in consecutive layers of the model.

Underfitting and Overfitting of ML models

Underfitting is the term used when a ML algorithm is unable to capture the underlying trend of the data. In such an instance, the algorithm does not fit the data well and is often caused by an excessively simple model. Underfitting usually results in a poor accuracy of the model. Poor accuracy is mainly attributed to a small sample size with the model incorrectly assuming a relationship among the data; for example, trying to predict a linear relationship on non-linear data. In these cases, the model will underestimate the complexity of the data and make the wrong prediction, thus decreasing its accuracy. Specifically, underfitting can be recognized if the model shows low variance but high bias. The methodologies most commonly used to avoid this situation ensure an adequate sample size and apply feature selection to reduce the number of features used in the model. **Figure 4A** shows an example of underfitting in a classification problem.

Opposite of underfitting is overfitting. This is a more frequent problem that occurs when a ML algorithm captures not only the data but also the noise of the data and inaccurate data entries. This often happens on a large dataset with a high number of features, resulting in an excessively complicated model. Specifically, overfitting can be recognized if the algorithm shows low bias but high variance. Non-parametric and non-linear methods are more sensitive to overfitting because of the higher degree of freedom they have in constructing the model. Overfitting can be avoided with several methodologies; the most commonly used being cross-validation, pruning, early

stopping and regularization. **Figure 4B**, shows an example of overfitting, while **Figure 4C** shows appropriate fitting in a classification problem.

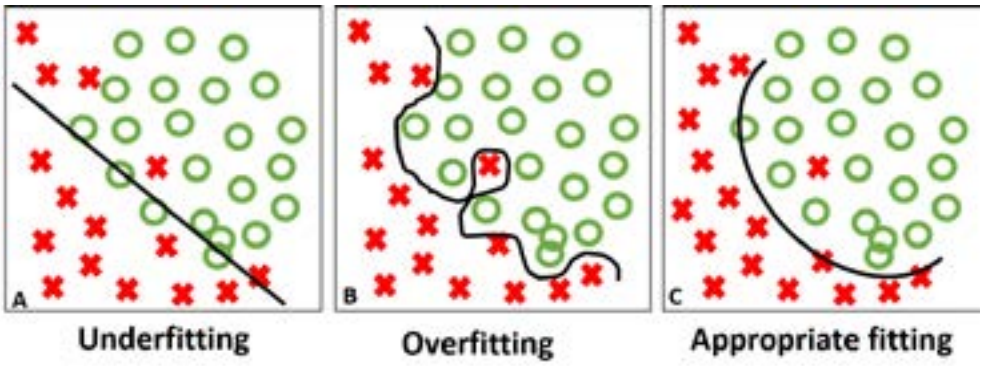

Figure 4. Variance-bias spectrum is machine learning models. (A) An underfitted model that weakly captures the dataset characteristics is considered a poorly performing model. (B) An overfitted model captures almost all the individual characteristics of the dataset, even the noise. An overfitted model is generally not useful as it does not generalize beyond the training data. (C) An appropriately fitting model might not correctly classify every single observation due to presence of noise in that dataset. However, such models will generalize well to the data beyond the training examples.

Interpretability

One of the advantages of ML applications is the ability to assess large amounts of data and find patterns that are invisible to the human eye. Herein also lies one of the bottlenecks of the implementation of ML algorithms in clinical practice, the so-called black box nature of ML algorithms. While ML algorithms are capable of accurately predicting an outcome, computers are not able, or not programmed, to logically and comprehensively translate the complex and often abstract calculations leading to the prediction back to its user. The use of these complex systems makes it difficult to explain the origin and logic behind the predictions that are made. The inability to comprehend the logic behind these predictions can cause issues for the clinician interpreting and using these predictions in clinical practice (92–95). An example in which a ML algorithm gave technically sound results but lacked clinical logic is a study investigating the risk prediction in pneumonia patients (15,96). The goal of this study was to predict the probability of death for patients with pneumonia in order to admit high-risk patients to the hospital, while low-risk patients can be treated without hospital admission (15). A multitask NN model was considered the most accurate model; however, it predicted that patients with pneumonia and asthma have a lower risk of death than patients with pneumonia without asthma. Although this result reflected the data accurately, it is a counterintuitive observation. Patients with asthma were admitted directly to the intensive care unit (ICU) as a precautious measure,

resulting in a lowered risk of dying from pneumonia; where non-asthma patients did not receive this precautious measure, thus demonstrating higher mortality. This caused the model to train on the effect of the increased intensity of treatment caused by ICU admission instead of the presence of asthma (15). When ML algorithm results are applied without trying to understand the logic behind the predictions, it can lead to false assumptions and poor clinical care. In this regard, it is important to not confuse accuracy with competence. A ML algorithm can accurately tell if a CT image depicts the heart, the lungs or the abdomen, but algorithms do not have any conception about what a heart or lung is. Although they are able to recognize the objects in an image, ML algorithms are not able to tell how a heart looks or how it works. ML can use a multitude of features from an image to predict the outcomes, but the algorithm has no notion of the actual meaning or content of those features. Therefore, the algorithms will not generalize to answer any question other than the one they were trained to answer and cannot provide context like humans can. The lack of interpretability of ML algorithms makes it difficult to link features with physical phenomena. This explains why the developments of current ML application in cardiac imaging are mainly focused on supporting human readers rather than replacing them. For example, automated CAC scores or left ventricular functional analysis using a ML approach holds the promise to significantly reduce the workload and time used to read images. These ML approaches are easy to check for outliers and have a direct translation to the manual task, leaving the task of interpretation to the assigned physician. For more complex analyses performed by ML algorithms which cannot be directly checked by a human operator, such as disease outcomes or prognostication, the black box conundrum needs to be addressed (97). Towards this, better visualization techniques should be used to shed light on the black box conundrum. Informative plots such as, **t**-distributed stochastic neighbor embedding (T-SNE) should be used to show the clustering of the data points in a trained model (98,99). Reporting of ML results along with these informative plots should become a norm. Similarly, for CNN models, reporting the findings with heat maps on the input images could help in delineating the inner workings of these algorithms (100)(101).

Conclusion

In summary, ML has been widely applied in cardiovascular imaging in order to improve diagnostic performance, for a specific outcome, as well as maximization of gain of new information to understand etiology of diseases. ML algorithms have also been employed for the detection and quantification of anatomic and physiologic atherosclerotic features detected on cardiovascular CT imaging. The continued expansion of ML applications coupled with deeper appreciation of its capabilities, as well as limitations, will enable healthcare to make the leapfrog into an era of individualized and precise healthcare administration. It will also provide the ability to investigate the effect, and prognostic significance, of phenotypic features seen on non-invasively acquired imaging studies.

REFERENCES

1. WHO. World Health Organization (2017) Key facts Cardiovascular Diseases (CVDs) [Internet]. [cited 2019 Mar 18]. Available from: https://www.who.int/cardiovascular_diseases/en/

2. Ballmoos MW Von, Haring B, Juillerat P, Alkadhi H. Annals of Internal Medicine Review Meta-analysis : Diagnostic Performance of Low-Radiation-Dose. 2011;

3. Wintersperger BJ, Bamberg F, De Cecco CN. Cardiovascular imaging: The past and the future, perspectives in computed tomography and magnetic resonance imaging. Invest Radiol. 2015;50(9):557–70.

4. De Cecco CN, Meinel FG, Chiaramida SA, Costello P, Bamberg F, Schoepf UJ. Coronary artery computed tomography scanning. Circulation. 2014;129(12):1341–5.

5. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J. 2016;37(29):2315–81.

6. Taylor AJ, Cerqueira M, Hodgson JMB, Mark D, Min J, O'Gara P, et al. ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 appropriate use criteria for cardiac computed tomography. J Am Coll Cardiol [Internet]. 2010;56(22):1864–94. Available from: http://dx.doi.org/10.1016/j.jacc.2010.07.005

7. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the American college of cardiology/American heart association task force on practice guidelines. J Am Coll Cardiol. 2014;63(25 PART B):2889–934.

8. Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, et al. 2013 ESC guidelines on the management of stable coronary artery disease. Eur Heart J. 2013;34(38):2949–3003.

9. Koo BK, Erglis A, Doh JH, Daniels D V., Jegere S, Kim HS, et al. 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 Noni. J Am Coll Cardiol [Internet]. 2011;58(19):1989–97. Available from: http://dx.doi.org/10.1016/j.jacc.2011.06.066

10. Min JK, Leipsic J, Pencina MJ, Berman DS, Koo B-K, van Mieghem C, et al. Diagnostic Accuracy of Fractional Flow Reserve From Anatomic CT Angiography. Jama [Internet]. 2012;308(12):1237. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/2012.jama.11274

11. Rocha-Filho JA, Blankstein R, Shturman LD, Bezerra HG, Okada DR, Rogers IS, et al. Incremental Value of Adenosine-induced Stress Myocardial Perfusion Imaging with Dual-Source CT at Cardiac CT Angiography. Radiology [Internet]. 2010;254(2):410–9. Available from: http://pubs.rsna.org/doi/10.1148/radiol.09091014

12. Rochitte CE, George RT, Chen MY, Arbab-Zadeh A, Dewey M, Miller JM, et al. Computed tomography angiography and perfusion to assess coronary artery stenosis causing perfusion defects by single photon emission computed tomography: The CORE320 study. Eur Heart J. 2014;35(17):1120–30.

13. De Cecco CN, Harris BS, Schoepf UJ, Silverman JR, McWhite CB, Krazinski AW, et al. Incremental value of pharmacological stress cardiac dual-energy CT over coronary CT angiography alone for the assessment of coronary artery disease in a high-risk population. Am J Roentgenol. 2014;203(1):70–7.

14. Nørgaard BL, Leipsic J, Gaur S, Seneviratne S, Ko BS, Ito H, et al. Diagnostic Performance of Noninvasive Fractional Flow Reserve Derived From Coronary Computed Tomography Angiography in Suspected Coronary Artery Disease. J Am Coll Cardiol [Internet]. 2014;63(12):1145–55. Available from: http://linkinghub.elsevier.com/retrieve/pii/S073510971400165X

15. Cooper GF, Abraham V, Aliferis CF, Aronis JM, Buchanan BG, Caruana R, et al. Predicting dire outcomes of patients with community acquired pneumonia. J Biomed Inform. 2005;38(5):347–66.

16. Cury RC, Magalhães TA, Borges AC, Shiozaki AA, Lemos PA, Soares J, et al. Dipyridamole stress and rest myocardial perfusion by 64-detector row computed tomography in patients with suspected coronary artery disease. Am J Cardiol [Internet]. 2010;106(3):310–5. Available from: http://dx.doi.org/10.1016/j.amjcard.2010.03.025

17. Williams MC, Shambrook J, Nicol ED. Assessment of patients with stable chest pain. Heart. 2018;

18. Hendel RC, Lindsay BD, Allen JM, Brindis RG, Patel MR, White L, et al. ACC Appropriate Use Criteria Methodology: 2018 Update. J Am Coll Cardiol [Internet]. 2018;71(8):935–48. Available from: http://dx.doi.org/10.1016/j.jacc.2018.01.007

19. Miles N. Wernick, Yang Y, Brankov JG, Yourganov G, Strother SC. Machine Learning in Medical Imaging. IEEE Signal Process Mag. 2014;27(4):25–38.

20. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: ImageNet Classification with Deep Convolutional Neural Networks. 2012.

21. Wen X, Shao L, Xue Y, Fang W. A rapid learning algorithm for vehicle classification. Inf Sci (Ny) [Internet]. 2015;295:395–406. Available from: http://dx.doi.org/10.1016/j.ins.2014.10.040

22. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans Pattern Anal Mach Intell. 2018;40(4):834–48.

23. Cire\csan D, Meier U, Masci J, Schmidhuber J. Multi-column deep neural network for traffic sign classification. Neural Networks. 2012;32(January 23, 2012):333–8.

24. Cernazanu-glavan C, Holban S. Segmentation of bone structure in X-ray images using convolutional neural network. 2013;xx(x).

25. Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJNL, Išgum I. Automatic segmentation of MR brain images with a convolutional neural network. 2017;1–11. Available from: http://arxiv.org/abs/1704.03295%0Ahttp://dx.doi.org/10.1109/TMI.2016.2548501

26. Işin A, Direkoğlu C, Şah M. Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods. Procedia Comput Sci. 2016;102(August):317–24.

27. Middleton I, Damper RI. Segmentation of magnetic resonance images using a combination of neural networks and active contour models. Med Eng Phys. 2004;26(1):71–86.

28. Ibragimov B, Pernus F, Strojan P, Xing L. TH-CD-206-05: Machine-Learning Based Segmentation of Organs at Risks for Head and Neck Radiotherapy Planning. Med Phys [Internet]. 43(6Part46):3883. Available from: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4958186

29. Sjogren AR, Leo MM, Feldman J, Gwin JT. Image segmentation and machine learning for detection of abdominal free fluid in focused assessment with sonography for trauma examinations: A pilot study. J Ultrasound Med. 2016;35(11):2501–9.

30. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep Learning for Identifying Metastatic Breast Cancer. 2016;(June). Available from: http://arxiv.org/abs/1606.05718

31. Yu-Jen Chen Y-J, Hua K-L, Hsu C-H, Cheng W-H, Hidayati SC. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. Onco Targets Ther [Internet]. 2015;2015. Available from: http://www.dovepress.com/computer-aided-classification-of-lung-nodules-on-computed-tomography-i-peer-reviewed-article-OTT

32. Poulton R, Caspi A, Milne BJ, Thomson WM, Taylor A, Sears MR, et al. Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis. 2013;360(9346):1640–5.

33. Cheng J-Z, Ni D, Chou Y-H, Qin J, Tiu C-M, Chang Y-C, et al. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. Sci Rep [Internet]. 2016;6(1):24454. Available from: http://www.nature.com/articles/srep24454

34. Bychkov D, Turkki R, Haglund C, Linder N, Lundin J. Abstract 5718: Outcome prediction in colorectal cancer using digitized tumor samples and machine learning. Cancer Res [Internet]. 2017 Jul 1;77(13 Supplement):5718 LP-5718. Available from: http://cancerres.aacrjournals.org/content/77/13_Supplement/5718.abstract

35. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016;

36. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, De Jong EEC, Van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology. 2017.

37. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep. 2015;

38. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, Folsom AR, et al. Coronary Calcium as a Predictor of Coronary Events in Four Racial or Ethnic Groups. N Engl J Med. 2008;358(13):1336–45.

39. Greenland P, Bonow RO, Brundage BH, Budoff MJ, Eisenberg MJ, Grundy SM, 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 Cl. J Am Coll Cardiol. 2007;49(3):378–402.

40. Išgum I, Rutten A, Prokop M, Van Ginneken B. Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. Med Phys. 2007;34(4):1450–61.

41. McEvoy JW, Blaha MJ, Nasir K, Blumenthal RS, Jones SR. Potential use of coronary artery calcium progression to guide the management of patients at risk for coronary artery disease events. Curr Treat Options Cardiovasc Med. 2012;14(1):69–80.

42. Yeboah J, McClelland RL, Polonsky TS, Burke GL, Sibley CT, O'Leary D, et al. Comparison of Novel Risk Markers for Improvement in Cardiovascular Risk Assessment in Intermediate-Risk Individuals. Jama [Internet]. 2012;308(8):788. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2012.9624

43. Arad Y, Goodman KJ, Roth M, Newstein D, Guerci AD. Coronary calcification, coronary disease risk factors, C-reactive protein, and atherosclerotic cardiovascular disease events: The St. Francis heart study. J Am Coll Cardiol [Internet]. 2005;46(1):158–65. Available from: http://dx.doi.org/10.1016/j.jacc.2005.02.088

44. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol [Internet]. 1990;15(4):827–32. Available from: http://dx.doi.org/10.1016/0735-1097(90)90282-T

45. Polonsky TS, Mcclelland RL, Jorgensen NW, Bild DE, Burke GL, Guerci AD, et al. Coronary Artery Calcium Score and Risk Classification for Coronary Heart Disease Prediction. Jama. 2010;303(16):1610–6.

46. B.T. H, V. V, I. C, J. S-M, H. G, J. K, et al. 15-Year prognostic utility of coronary artery calcium scoring for all-cause mortality in the elderly. Atherosclerosis. 2016;

47. Ma S, Liu A, Carr J, Post W, Kronmal R. Statistical modeling of Agatston score in multi-ethnic study of atherosclerosis (MESA). PLoS One. 2010;5(8).

48. Wolterink JM, Leiner T, Takx RAP, Viergever MA, Isgum I. Automatic Coronary Calcium Scoring in Non-Contrast-Enhanced ECG-Triggered Cardiac CT With Ambiguity Detection. IEEE Trans Med Imaging. 2015 Sep;34(9):1867–78.

49. Išgum I, Rutten A, Prokop M, Staring M, Klein S, Pluim JPW, et al. Automated aortic calcium scoring on low-dose chest computed tomography. Med Phys. 2010;

50. Brunner G, Chittajallu DR, Kurkure U, Kakadiaris IA. Toward the automatic detection of coronary artery calcification in non-contrast computed tomography data. Int J Cardiovasc Imaging [Internet]. 2010;26(7):829–38. Available from: http://link.springer.com/10.1007/s10554-010-9608-1

51. Takx RAP, De Jong PA, Leiner T, Oudkerk M, De Koning HJ, Mol CP, et al. Automated coronary artery calcification scoring in non-gated chest CT: Agreement and reliability. PLoS One. 2014;9(3).

52. Mylonas I, Alam M, Amily N, Small G, Chen L, Yam Y, et al. Quantifying coronary artery calcification from a contrast-enhanced cardiac computed tomography angiography study. Eur Heart J Cardiovasc Imaging. 2014;15(2):210–5.

53. Hecht HS. Coronary artery calcium scanning: Past, present, and future. JACC Cardiovasc Imaging [Internet]. 2015;8(5):579–96. Available from: http://dx.doi.org/10.1016/j.jcmg.2015.02.006

54. Mittal S, Zheng Y, Georgescu B, Vega-Higuera F, Zhou SK, Meer P, et al. Fast Automatic Detection of Calcified Coronary Lesions in 3D Cardiac CT Images. In: Wang F, Yan P, Suzuki K, Shen D, editors. Machine Learning in Medical Imaging. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 1–9.

55. Yang G, Chen Y, Ning X, Sun Q, Shu H, Coatrieux J. Automatic coronary calcium scoring using noncontrast and contrast CT images Automatic coronary calcium scoring using noncontrast and contrast CT images. Med Phys [Internet]. 2016;2174. Available from: http://dx.doi.org/10.1118/1.4945045

56. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal [Internet]. 2016;34:123–36. Available from: http://dx.doi.org/10.1016/j.media.2016.04.004

57. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. Eur Heart J. 2017;38(7):500–7.

58. Otaki Y, Hell M, Slomka PJ, Schuhbaeck A, Gransar H, Huber B, et al. Relationship of epicardial fat volume from noncontrast CT with impaired myocardial flow reserve by positron emission tomography. J Cardiovasc Comput Tomogr. 2015;

59. Han D, Lee JH, Rizvi A, Gransar H, Baskaran L, Schulman-Marcus J, et al. Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: A machine learning approach. J Nucl Cardiol. 2017;1–11.

60. Mannil M, Von Spiczak J, Manka R, Alkadhi H. Texture Analysis and Machine Learning for Detecting Myocardial Infarction in Noncontrast Low-Dose Computed Tomography: Unveiling the Invisible. Invest Radiol. 2018;

61. Freiman M, Nickisch H, Prevrhal S, Schmitt H, Vembar M, Maurovich-Horvat P, et al. Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation. Med Phys. 2017 Mar;44(3):1040–9.

62. Xiong G, Kola D, Heo R, Elmore K, Cho I, Min JK. Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. Med Image Anal. 2015;

63. Dey D, Gaur S, Ovrehus KA, Slomka PJ, Betancur J, Goeller M, et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: A multicentre study. Eur Radiol. 2018;

64. Yang DH, Kim Y-HH, Roh JH, Kang J-WW, Ahn J-MM, Kweon J, et al. Diagnostic performance of on-site CT-derived fractional flow reserve versus CT perfusion. Eur Heart J Cardiovasc Imaging [Internet]. 2017;18(4):432–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27354345

65. Commandeur F, Goeller M, Betancur J, Cadet S, Doris M, Chen X, et al. Deep Learning for Quantification of Epicardial and Thoracic Adipose Tissue from Non-Contrast CT. IEEE Trans Med Imaging. 2018;(99).

66. Mahabadi AA, Berg MH, Lehmann N, Kälsch H, Bauer M, Kara K, et al. Association of epicardial fat with cardiovascular risk factors and incident myocardial infarction in the general population: The Heinz Nixdorf recall study. J Am Coll Cardiol [Internet]. 2013;61(13):1388–95. Available from: http://dx.doi.org/10.1016/j.jacc.2012.11.062

67. Rodrigues O, Morais FFC, Morais NAOS, Conci LS, Neto L V., Conci A. A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography. Comput Methods Programs Biomed. 2016;123:109–28.

68. Rodrigues EO, Conci A, Morais FFC, Perez MG. Towards the Automated Segmentation of Epicardial and Mediastinal Fats. 2015 Ieee Int Conf Ind Technol. 2015;1779–85.

69. Douglas PS, Hoffmann U, Patel MR, Mark DB, Al-Khalidi HR, Cavanaugh B, et al. Outcomes of Anatomical versus Functional Testing for Coronary Artery Disease. N Engl J Med [Internet]. 2015;372(14):1291–300. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1415516

70. Newby D. CT coronary angiography in patients with suspected angina due to coronary heart disease (SCOT-HEART): An open-label, parallel-group, multicentre trial. Lancet. 2015;

71. Motoyama S, Sarai M, Harigaya H, Anno H, Inoue K, Hara T, et al. Computed Tomographic Angiography Characteristics of Atherosclerotic Plaques Subsequently Resulting in Acute Coronary Syndrome. J Am Coll Cardiol [Internet]. 2009;54(1):49–57. Available from: http://dx.doi.org/10.1016/j.jacc.2009.02.068

72. Motoyama S, Ito H, Sarai M, Kondo T, Kawai H, Nagahara Y, et al. Plaque characterization by coronary computed tomography angiography and the likelihood of acute coronary events in mid-term follow-up. J Am Coll Cardiol. 2015;66(4):337–46.

73. Motoyama S, Sarai M, Narula J, Ozaki Y. Coronary CT angiography and high-risk plaque morphology. Cardiovascular Intervention and Therapeutics. 2013.

74. Min JK, Shaw LJ, Berman DS. The Present State of Coronary Computed Tomography Angiography. A Process in Evolution. Journal of the American College of Cardiology. 2010.

75. Kang D, Dey D, Slomka PJ, Arsanjani R, Nakazato R, Ko H, et al. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography. J Med Imaging. 2015;

76. Dey D, Schepis T, Marwan M, Slomka PJ, Berman DS, Achenbach S. Automated Three-dimensional Quantification of Noncalcified Coronary Plaque from Coronary CT Angiography: Comparison with Intravascular US. Radiology. 2010;

77. Hell MM, Dey D, Marwan M, Achenbach S, Schmid J, Schuhbaeck A. Non-invasive prediction of hemodynamically significant coronary artery stenoses by contrast density difference in coronary CT angiography. Eur J Radiol. 2015;

78. Min JK, Berman DS, Budoff MJ, Jaffer FA, Leipsic J, Leon MB, et al. Rationale and design of the DeFACTO (Determination of Fractional Flow Reserve by Anatomic Computed Tomographic AngiOgraphy) study. J Cardiovasc Comput Tomogr. 2011;

79. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, et al. Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart. Circulation. 2002;539–42.

80. Tesche C, De Cecco CN, Albrecht MH, Duguay TM, Bayer RR, Litwin SE, et al. Coronary CT Angiography–derived Fractional Flow Reserve. Radiology [Internet]. 2017;285(1):17–33. Available from: http://pubs.rsna.org/doi/10.1148/radiol.2017162641

81. Tesche C, Cecco CN De, Baumann S, Renker M, McLaurin TW, Duguay TM, et al. Coronary CT Angiography–derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling. Radiology [Internet]. 0(0):171291. Available from: https://doi.org/10.1148/radiol.2018171291

82. Renker M, Schoepf UJ, Wang R, Meinel FG, Rier JD, Bayer RR, et al. Comparison of diagnostic value of a novel noninvasive coronary computed tomography angiography method versus standard coronary angiography for assessing fractional flow reserve. Am J Cardiol [Internet]. 2014;114(9):1303–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25205628

83. Coenen A, Lubbers MM, Kurata A, Kono A, Dedic A, Chelu RG, et al. Fractional Flow Reserve Computed from Noninvasive CT Angiography Data: Diagnostic Performance of an On-Site Clinician-operated Computational Fluid Dynamics Algorithm. Radiology [Internet]. 2015;274(3):674–83. Available from: http://pubs.rsna.org/doi/10.1148/radiol.14140992

84. Coenen A, Rossi A, Lubbers MM, Kurata A, Kono AK, Chelu RG, et al. Integrating CT Myocardial Perfusion and CT-FFR in the Work-Up of Coronary Artery Disease. JACC Cardiovasc Imaging. 2016;(2014).

85. Kruk M, Wardziak Ł, Demkow M, Pleban W, Pręgowski J, Dzielińska Z, et al. Workstation-Based Calculation of CTA-Based FFR for Intermediate Stenosis. JACC Cardiovasc Imaging. 2016;9(6):690–9.

86. Hell MM, Motwani M, Otaki Y, Cadet S, Gransar H, Miranda-Peats R, et al. Quantitative global plaque characteristics from coronary computed tomography angiography for the prediction of future cardiac mortality during long-term follow-up. Eur Heart J Cardiovasc Imaging. 2017;

87. Nensa F, Forsting M, Wetter A. [The future of radiology : What can we expect within the next 10 years?]. Urol A. 2016;

88. Yi X, Babyn P. Sharpness-Aware Low-Dose CT Denoising Using Conditional Generative Adversarial Network. J Digit Imaging. 2018;

89. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets (NIPS version). Adv Neural Inf Process Syst 27. 2014;

90. Krig S. Ground Truth Data, Content, Metrics, and Analysis. Comput Vis Metrics [Internet]. 2014;283–311. Available from: http://link.springer.com/10.1007/978-1-4302-5930-5_7

91. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2015.

92. Cabitza F, Rasoini R, Gensini GF. Unintended Consequences of Machine Learning in Medicine. Jama [Internet]. 2017;2017–8. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.7797

93. Lipton ZC. The Mythos of Model Interpretability. 2016;(Whi). Available from: http://arxiv.org/abs/1606.03490

94. Krause J, Perer A, Ng K. Interacting with Predictions: Visual Inspection of Black-box Machine Learning Models. ACM Conf Hum Factors Comput Syst [Internet]. 2016;5686–97. Available from: http://dl.acm.org/citation.cfm?doid=2858036.2858529

95. Lee J-G, Jun S, Cho Y-W, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. Korean J Radiol [Internet]. 2017;18(4):570. Available from: https://synapse.koreamed.org/DOIx.php?id=10.3348/kjr.2017.18.4.570

96. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible Models for HealthCare. Proc 21th ACM SIGKDD Int Conf Knowl Discov Data Min - KDD '15 [Internet]. 2015;1721–30. Available from: http://dl.acm.org/citation.cfm?doid=2783258.2788613

97. Ponte P, Melko RG. Kernel methods for interpretable machine learning of order parameters. Phys Rev B [Internet]. 2017 Nov;96(20):205146. Available from: https://link.aps.org/doi/10.1103/PhysRevB.96.205146

98. Maaten L van der, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res. 2008;

99. van der Maaten L. Accelerating t-SNE using Tree-Based Algorithms. J Mach Learn Res. 2014;

100. Samek W, Binder A, Montavon G, Lapuschkin S, Müller KR. Evaluating the visualization of what a deep neural network has learned. IEEE Trans Neural Networks Learn Syst. 2017;

101. Erhan D, Bengio Y, Courville A, Vincent P. Visualizing higher-layer features of a deep network. Univ Montréal, Tech Rep. 2009;