Journal of Cardiovascular Computed Tomography 13 (2019) 274–280

Contents lists available at ScienceDirect

Journal of Cardiovascular Computed Tomography

journal homepage: www.elsevier.com/locate/jcct

Invited review

Advanced atherosclerosis imaging by CT: Radiomics, machine learning and deep learning

Márton Kolossváry[a,*], Carlo N. De Cecco[b], Gudrun Feuchtner[c], Pál Maurovich-Horvat[a]

[a] Cardiovascular Imaging Research Group, Heart and Vascular Center, Semmelweis University, Budapest, Hungary
[b] Division of Cardiothoracic Imaging, Department of Radiology and Imaging Science, Emory University, Atlanta, GA, USA
[c] Department of Radiology, Innsbruck Medical University, Innsbruck, Austria

ARTICLE INFO

Keywords:
Atherosclerosis
Coronary CT angiography
Radiomics
Machine learning
Deep learning

ABSTRACT

In the last decade, technical advances in the field of medical imaging significantly improved and broadened the application of coronary CT angiography (CCTA) for the non-invasive assessment of coronary artery disease. Recently, similar breakthroughs are happening in the post-processing, analysis and interpretation of radiological images. Technologies such as radiomics allow to extract significantly more information from scans than what human visual assessment is capable of. This allows the precision phenotyping of diseases based on medical images. The increased amount of information can then be analyzed using novel data analytic techniques such as machine learning (ML) and deep learning (DL), which utilize the power of big data to build predictive models, which seek to mimic human intelligence, artificially. Thanks to big data availability and increased computational power, these novel analytic methods are outperforming conventional statistical techniques. In this current overview we describe the basics of radiomics, ML and DL, highlighting similarities, differences, limitations and potential pitfalls of these techniques. In addition, we provide a brief overview of recently published results on the applications of the aforementioned techniques for the non-invasive assessment of coronary atherosclerosis using CCTA.

1. Introduction

Coronary CT angiography (CCTA) plays a pivotal role in the non-invasive evaluation of coronary artery disease (CAD) and has gained worldwide clinical acceptance.[1–3] As a non-invasive and cost-effective imaging tool, it has a great potential in reducing the global socio-economic burden of CAD.[4–6] The PROMISE trial has shown equal outcomes considering the composite endpoint of death, myocardial infarction, hospitalization for unstable angina and major procedural complication as compared to the established modality SPECT at after a median follow-up of 25 months (3.3% vs. 3.0%, $p = 0.75$). Furthermore, the SCOT-HEART trial demonstrated lower major adverse cardiac event (MACE) rates (2.3 vs 3.9%, $p < 0.001$) for CCTA as compared to standard of care after 4.8 years.[1,2]

The strength of CCTA lies in its ability to reliably exclude coronary stenosis and, even more importantly, to directly visualize the vessel wall and plaque morphology.[7] This unique ability to characterize coronary atherosclerosis is a major advantage of CCTA over competitive diagnostic strategies.[8] Hence, plaque characterization by CCTA has become a major research focus over the past years. The detection of high-risk plaque (HRP) markers allows for a highly specific labeling of patients at increased risk for MACE.[9,10] The four HRP markers (low attenuation, positive remodeling, spotty calcification and the napkin-ring sign) have been shown to positively correlate with adverse outcomes and to predict ischemia even in non-obstructive lesions.[11–15]

CT technology has developed at unprecedented speeds over the past years, meanwhile data-analysis and image interpretation has evolved at a slower pace. CCTA interpretation relies on subjective image assessment. However, the possibility to extract quantitative information may provide a unique opportunity to overcome limitations of subjective visual assessment. For example, the inter-reader reproducibility of HRP features is poor (κ range: 0.15–0.34) even among experienced readers.[16] Furthermore, visual assessment is dependent on a priori knowledge and has limited capabilities to discern subtle differences and

Abreviations: AI, Artificial intelligence; AUC, Area under the curve; CAD, Coronary artery disease; CCTA, Coronary CT angiography; DL, Deep learning; DNN, Deep neural network; FFR, Fractional flow reserve; GLCM, Gray level co-occurrence matrix; GLRLM, Gray level run length matrix; HRP, High-risk plaque; HU, Hounsfield units; IG, Information gain; MACE, Major adverse cardiac event; ML, Machine learning

* Corresponding author. 68. Varosmajor street, 1122, Budapest, Hungary.
E-mail address: marton.kolossvary@cirg.hu (M. Kolossváry).

https://doi.org/10.1016/j.jcct.2019.04.007
Received 20 December 2018; Received in revised form 24 March 2019; Accepted 15 April 2019
Available online 21 April 2019

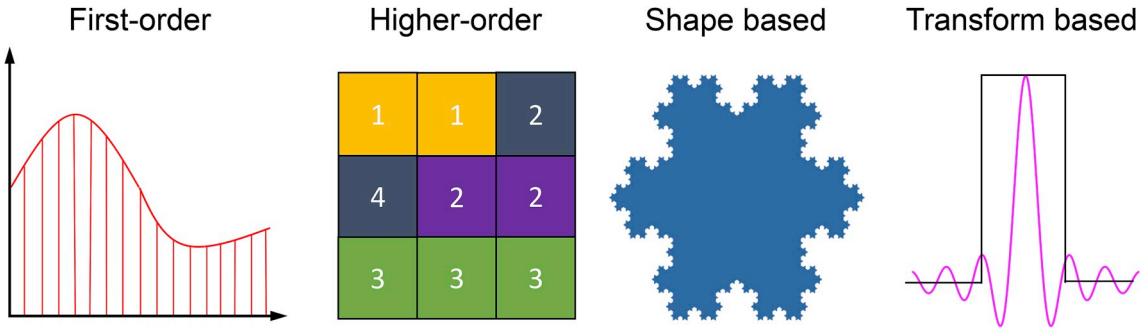

Fig. 1. Pictograms of classes of radiomic parameters.
First-order parameters calculate statistics based on the Hounsfield distribution of the voxel values and discard all spatial information. *Higher order* metrics calculate statistics based on the spatial co-occurrence of specific voxel values. *Shape based* statistics quantify the spatial complexity and self-symmetry of the lesion. *Transform based* methods convert the spatial information into another domain, such as the frequency domain to calculate radiomic statistics.

to define new imaging markers.[17] Novel image and data analytic techniques such as radiomics, machine learning (ML) and deep learning (DL) may decrease inter-reader variations, increase the amount of quantitative information and improve the diagnostic and prognostic accuracy, while reducing subjectivity and biases. In this overview, we provide a broad overview on the application of radiomics, ML and DL for the non-invasive assessment of coronary atherosclerosis using CCTA.

2. General overview of radiomics, machine and deep learning

2.1. Radiomics

Radiomics is the process of extracting numerous quantitative parameters from radiological images to describe the texture and spatial complexity of lesions.[18] Radiomics provides a tool for precision phenotyping of abnormalities based-on radiological images.[19] Its name refers to its similarities with other “omics” techniques which create big data on biological systems such as the genome to assist precision medicine.[20] Radiomics basically transforms images into data.[17]

2.1.1. First-order radiomic parameters

Radiomic features can be grouped into four categories.[18] *First-order parameters* use the Hounsfield units (HU) of the lesions to derive statistics describing the distribution of the HU values. These features such as the mean HU, minimum HU or standard deviation of the HU discard all spatial information and only rely on the frequency of specific HU values. This results in significant information loss with regards to the spatial distribution of HU values, since very different plaques can have very similar first-order parameter values.

2.1.2. Higher-order textural features

This resulted in the development of *higher-order features* which aim to describe not the HU distribution but the spatial distribution of the voxels basically characterizing the texture and heterogeneity of the lesion. These parameters originate from the field of computer vision, when researchers were confronted in the 1970's with satellite images and had to decide which areas are urban and which are industrial.[21,22] The basic concept of these parameters is to quantify the spatial co-occurrence of given voxel values. For example, the gray level co-occurrence matrix (GLCM) describes the frequency by which given value voxel pairs co-occur next to each other.[21] From this probability matrix different metrics are derived, which quantify for example the degree of heterogeneity in the image, which is defined as how often dissimilar value voxels co-occur next to each other. Another common matrix is the gray level run length matrix (GLRLM) which quantifies the frequency of identical voxel values occurring next to each other repeatedly in a given direction.[22] Based on these, features can be derived describing the texture of an abnormality. Several other matrices exist, which look at different attributes of the voxels and their surroundings.[23–25]

2.1.3. Shaped-based radiomic parameters

Shape based parameters quantify the spatial complexity of the lesion. These parameters range from commonly known parameters such as the surface or volume of a lesion or lesion component (e.g. low attenuation plaque volume) to more complex entities such as compactness which quantifies how the surface and the volume relate to each other. Another class of geometrical parameters are fractal dimensions which enumerate self-symmetry.[26] Self-repeating patterns can be observed in our environment, for example how the trunk of a tree divides can also be appreciated in how the branches divide and also in how the veins of a leaf divide. Fractal dimensions enumerate this self-symmetry and offer a quantitative approach to measure spatial complexity.

2.1.4. Transform based radiomic features

Medical images are basically intensity values in the spatial domain. *Transform based metrics* transfer this spatial information into another representation for example the frequency domain. The resulting new image or data then can be analyzed using previously discussed methods, or used to filter out specific information from the image. Illustrative pictograms of the classes of radiomic statistics can be found in Fig. 1.

2.2. Machine learning

The amount of available medical information is increasing at exponential speeds.[27] Currently the difficulty is not how to get medical big data, but how to organize, analyze and clinically utilize the data collected in biobanks and repositories. Conventional statistical methods utilize probability theory to create mathematical formulas which describe the relationship between variables. This approach is usually acceptable for population-based analysis, but in the era of precision medicine and big data, new methods are needed which can model complex non-linear relationships and infer results specific to each case rather than being generally true to the population.

2.2.1. Concept of artificial intelligence and machine learning

Humans are skilled at identifying unique patterns and inferring complex connections between data. However, this *natural intelligence* is not based on mathematical equations, but on observations and experience. *Artificial intelligence (AI)* tries to create models which think and act humanly and rationally.[28] To achieve this, first inputs are needed from the environment. Then this information and the previous observations need to be stored and analyzed, which can be performed using ML. ML is an analytic method a sub-division of AI, which uses computer algorithms with the ability to learn from data, without being

explicitly programmed.[29] These algorithms are similar to the human learning process, in the sense that more data they are trained on, the better they perform. In the medical field the main goal of ML techniques is to harvest the potential of big data to discover new relationships in the data that conventional statistical methods might not be able to. While conventional statistical approaches can provide a clear mathematical formula regarding the relationship of the variables, not all methods of ML are capable of describing the connection between parameters through mathematical equations. Instead they build their predictive models based-on patterns in the data experienced through training, and make prediction by comparing a new instance to previous similar occurrences.[30] To better understand how ML works, we describe examples of ML algorithms to help the readers have a general understanding of these methods.

2.2.2. Supervised learning algorithms

Supervised learning algorithms try to predict to which class a new instance belongs to (e.g. disease +/−) or what continuous value the instance has (e.g. Agatston score). To achieve this, supervised learning algorithms require labeled data to be trained on. One of the most intuitive algorithms is the *k-nearest neighbor*, which is based on the intuition that cases that are close to each other usually have much in common. In case of a classification problem, it evaluates which class the neighbors belong to and assigns the class that is most frequent among its k number of neighbors. In case of regression, it assigns the average value of its k number of neighbors. As one can appreciate, the value for *k* is chosen arbitrarily. This is referred to as a *hyperparameter*. As opposed to conventional linear regression equation, which only has one solution, ML models can have infinite number of solutions depending on the values of these hyperparameters. The optimal values for these hyperparameters need to explored balancing between overfitting and underfitting the ML model. Another commonly used method is based on the human decision-making process and is called *decision trees*. For example, if the goal is to assess CAD risk, the first step is to gather risk factors and based on one of the parameters (for example gender) the data is split into two groups. Then based on the next parameter the data is split further (for example age ranges) to achieve more homogeneous sub-groups. This methodology can be observed in guideline flow diagrams and also in risk calculation charts. An important difference is that decision trees select which parameters to use and where to divide them based on information theory and not on common knowledge. This data driven approach quantifies for each parameter the amount of randomness (entropy) that is reduced if the data is divided on a given metric. This so-called *information gain* (IG) can also be used to quantify the relative importance of a parameter with respect to a given outcome. However, this is a univariate analytic method, and therefore does not consider possible collinearities in the data.[31] Finally, as a last example for supervised learning, we would like to discuss *neural networks,* which are mathematical models inspired by the biological brain of how learning and decision making happens or could happen in living organisms.[32] The basic concept of neural networks is that the biological brain provides an archetype of an intelligent system capable of understanding its inputs and making rational decisions based-on previous experiences. Therefore, a logical idea is to reverse engineer the structure of the biological brain and to build computational models, so called neural networks mimicking its behavior. The building blocks of neural networks are single *perceptrons* or *neurons*. The perceptron receives inputs which are multiplied by weights, similarly to a neuron, which receives inputs through its dendrites. Based on an activation function (e.g. logistic function) it outputs either a continuous or binary value, similarly to a neuron through its axon. The simple perceptron is similar to a regression equation. Several perceptron can be put next to each other to create a *single layer perceptron*, which is capable of solving complex problems. Furthermore, by adding new layers of perceptrons, *multi-layer perceptrons* can be created. The perceptron layers which are neither connected to the inputs nor give the final output are called *hidden layers*. Neural networks which have two or more hidden layers are called *deep neural networks (DNN)*. These networks are basically many interconnected regression equations, where the output of one equation is the input to another equation.

2.2.3. Unsupervised learning algorithms

Unsupervised learning algorithms on the other hand use unlabeled data. The intention is not to predict the value of a new instance, but for example to find clusters in the data and assign the cases to these clusters. Then these clusters can be compared and further analyzed to infer new relationships in the data. The most common method is *k-means clustering*. This algorithm groups the data into *k* clusters by minimizing the distances of the instances to the center of the clusters. Several other clustering algorithms exist. However, evaluating which clustering techniques performs the best is not trivial and can lead to significantly different results.

2.2.4. Semi-supervised learning algorithms

Semi-supervised learning algorithms utilize data that are both labeled and unlabeled. It is used in cases when there is a large amount of data, however labeling all cases is not feasible. By adding some labeled cases to the model, the semi-supervised models usually outperform unsupervised learning since more information is present regarding how many clusters there may be, or what the distribution of the parameters are in each group.

Pictorial examples of ML models can be seen in Fig. 2.

2.3. Deep learning

DL refers to a sub-division of ML, which expands on the basics of DNN to create complex neural architectures to solve difficult problems, which using conventional programing based on mathematical logic would be impossible. Some complex architectures are inspired by nature. For example, instead of fully connecting all the perceptrons in a neural network as a simple DNN, skip connection similar to pyramidal cells in the cortex of the brain, skips some hidden layers to activate a neuron deeper in the network.[33] There is no exact boundary from which we would consider neural networks DL. It is rather a concept describing the use of complex neural networks for tasks that are commonly needed in AI, such as image recognition, natural language processing or classification and clustering. In addition to the interpretation and processing of data, complex neural architectures can be helpful to find efficient representations of the data using auto-encoders or restricted Boltzmann machines by reducing the dimensionality of the dataset.[34] Furthermore, these complex neural models can also be used to generate data. So called generative adversarial networks are capable of creating images of human faces which look similar to real faces.[35]

2.3.1. Convolutional neural networks

Convolutional neural networks commonly used for image recognition, are inspired by how the visual cortex is structured.[36] Instead of feeding all the inputs (pixels or voxels) to each neuron in the initial layer, only those pixels or voxels that are in the receptor field of the neuron are used. Similar to the visual cortex this allows identification of simple structures, such as lines. However, as we go deeper into the neural network, the neurons are capable of identifying more complex structures and eventually are able to classify the images, similar to as how our brain would.[37] Convolutional neural networks are extensively used in radiology for image segmentation and classification and have the potential to transform the field of radiology.[38]

2.4. Similarities and differences between radiomics, machine and deep learning

Radiomics is a feature generation approach. Based on human expert opinions, mathematical models are built to describe complex ideas such

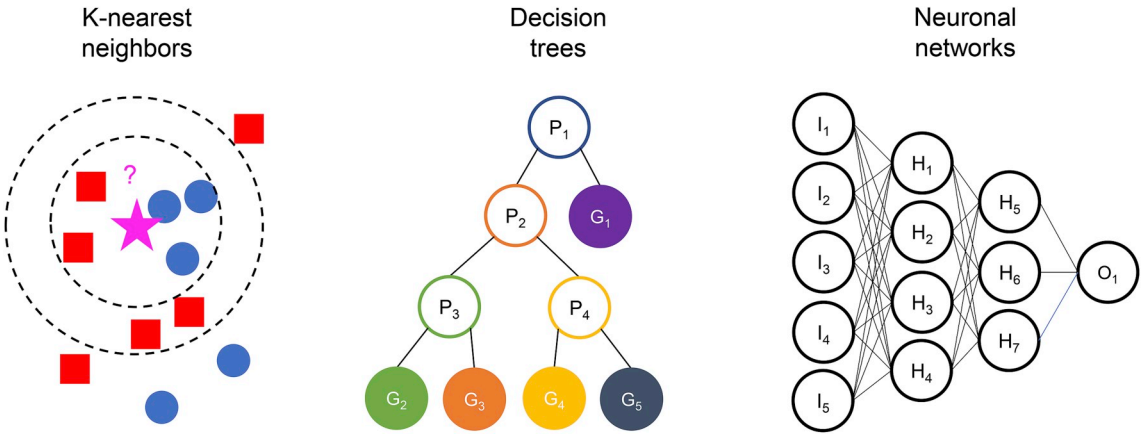

Fig. 2. Examples of machine learning algorithms.
K nearest neighbors algorithm searches for a new instance's nearest neighbors. Based on how many neighbors you consider, or the distance from the new case, the decision to which group to classify the given instance can be different. *Decision trees* classify instances based on rules. Splitting the cases based on a parameter (P_1), we might be able to differentiate a sub-group of cases which belong to the same class (G_1). The remaining instances need to be split based on different parameters ($P_{2\text{-}4}$) to receive reasonable classification accuracy ($G_{2\text{-}5}$). *Neural networks* are based on the architecture of our brains. The inputs (statistical parameters, or raw data: voxels) make up the input layer ($I_{1\text{-}5}$), which in case of a fully connected neural network are connected to the hidden layers ($H_{1\text{-}7}$). Then the last hidden layer is connected to the output (O_1) which produces the result of the model.

as texture or spatial complexity. Hundreds of different parameters can be calculated, all emphasizing different aspects of these concepts. The resulting big data databases then can be analyzed using conventional statistical methods, methods used in genomics or can be inputs to ML models. On the other hand, DL uses the raw data, the pixel or voxel values themselves. Using convolution, imaging features are automatically defined in the network, rather than defined beforehand based on expert knowledge. While we know which neurons are activated and know all the parameters of the fitted neural network, we have very limited information as to what it is actually seeing and what features are identified. Several research efforts are currently being conducted to improve the interpretability of the models and to provide solutions for the medical community to better understand how these models work. For example, by highlighting the pixels from which the model came to its conclusion, the medical professionals can better understand and interpret the outputs of these neural networks.[39,40]

On the other hand, radiomics can be seen as a precision phenotyping method for medical images, where the imaging features are carefully defined beforehand based-on expert opinions. It creates a phenogram, for example an atherosclerotic plaque fingerprint, which describes that specific lesion. The parameters, which produced the results can be interpreted in a straight forward manner as opposed to the DL outputs. Similarities and differences in radiomics, ML and DL with respect to AI can be appreciated in Fig. 3.

3. Limitations and potential pitfalls of radiomics, machine and deep learning

3.1. Radiomics

Radiomic parameters are calculated from the HU values of voxels, and therefore all possible effects, which influence HU values (CT hardware, patient characteristics, tube voltage and current, kernel, reconstruction algorithm, etc.) may influence radiomic statistics. There is only limited data on how these potential influencing factors may affect radiomics statistics derived from CCTA images.[41] Furthermore, as the metrics are calculated from a volume of interest (e.g. the coronary plaque) segmentations may also affect the parameter values. Nonetheless, radiomic parameters are constructs created by humans. Potentially infinite number of parameters could be defined; however, it remains a question what is the optimal number and type of radiomic parameters. Furthermore, with increasing number of parameters the possibility of overfitting also increases. Therefore, careful evaluation is needed in all cases to preserve generalizability.

3.2. Machine learning

ML techniques are superior to conventional statistical methods as they are able to model unique patterns in the data. However, similar to radiomics, consider: large number of parameters are used, thus they are prone to overfitting. Running a ML model on the whole population, and evaluating its diagnostic accuracy on the same population is incorrect. This can lead to biased estimates and result in diagnostic accuracies around 100% and AUC values of 1.0. The gold standard of evaluating a model is by evaluating its diagnostic accuracy on an external validation cohort. However, in many cases this is not feasible, and therefore a *test-set* is created from the data by randomly setting aside 20–40% of the data, which is only used for evaluating the fitted model which is optimized on the *training set*. Setting aside a test-set is not reasonable in some cases, when there is only a limited number of data. In these cases, n-fold cross-validation may be used. The data is partitioned into n equally sized subsamples. The ML model is trained on n-1 subsamples, and is evaluated on the remaining subsample. This is done n times so that each instance in the dataset was part of the evaluation once. The accuracies from each fold are averaged to receive an overall measure of diagnostic accuracy. To receive a better estimate, this can be repeated several thousands of times with random sampling to get a better estimate of accuracy. This usually still leads to overestimation of the diagnostic accuracy, but is better than modeling and validating on the whole dataset. Furthermore, with all statistical methods, ML is prone to training biases. As all other statistical techniques, it makes its predictions based on the data it was trained on. Therefore, applying it to datasets which the training data does not represent well, can lead to biased results and inaccurate predictions. In addition, models learn from the training dataset and therefore are inherently limited to the quality of the data.

3.3. Deep learning

Neural network models have hundreds or even thousands of perceptrons, which require optimization to determine their value in the models. Therefore, several thousands of images are needed for the network to converge to a solution in this n-dimensional hyperspace. However, having such a large number of labeled datasets in medicine is

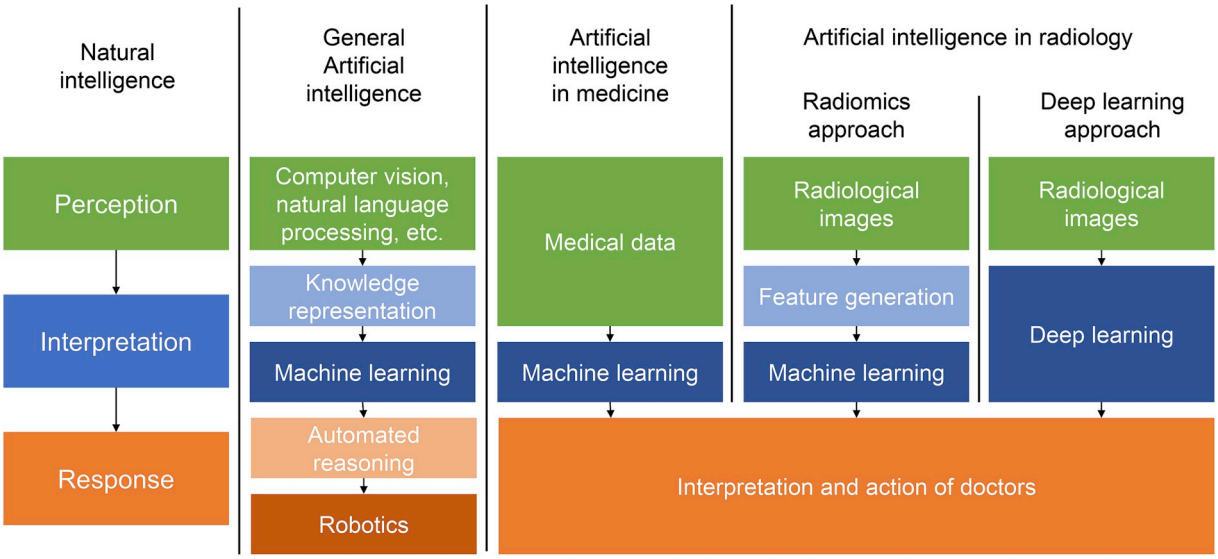

Fig. 3. Flow diagram showing the possible implementations of artificial intelligence to medical data and showing the similarities and differences between radiomics, machine learning and deep learning.
Artificial intelligence tries to mimic natural intelligence through automating processes that are needed for an intelligent system to perceive, interpret and respond to its surroundings. Medicine can utilize the benefits of machine learning to help interpret the large amounts of data currently available in medicine. In case of radiological images, radiomics can be used to extract vast amounts of information, which can be inputs to machine learning. On the other hand, deep learning automatically identifies imaging markers in the neural network while training, rather than defining them beforehand.

difficult. Furthermore, as these models learn from cases they were taught on, they will have difficulties to identify unusual cases of which only a few cases are available. While radiologist can identify rare cases after seeing a few instances, these models require many more cases to be able to recognize these instances with acceptable accuracy. In addition, DL networks can be tricked. So called adversarial images can be used, which can be added to images as noise which the human eye cannot see, but alters the output of the neural network to be completely wrong.[42–44]

4. Current results of implementing radiomics, machine and deep learning to cardiac CT

4.1. Radiomics

Among HRP features, the napkin-ring sign is the only qualitative feature, solely depending on the interpretation of the medical professional.[45] As its reproducibility varies widely (κ: 0.15–0.86) more quantitative methods, less dependent on the knowledge of medical professions would be warranted.[15,16] An initial study investigating the potential of radiomics in atherosclerosis imaging, selected 30 napkin-ring sign plaques and matched them with 30 non napkin-ring plaques which were however similar in degree of calcification and stenosis, plaque localization, tube voltage, and image reconstruction.[46] Among the calculated 4440 radiomics parameters, 20.6% of the parameters was consider: significantly different between the two groups at a corrected p level of 0.0012, while none of the conventional quantitative parameters showed any difference. The best radiomics parameter (short-run low-gray-level emphasis) significantly outperformed the best conventional parameter (mean plaque attenuation) to identify the napkin-ring sign (cross-validated AUC: 0.89 vs. 0.75; respectively). Furthermore, cluster analysis revealed 44 information clusters among the radiomic parameters indicating that there might be more types of plaque morphology as opposed to the four HRP which are known from the literature.

CCTA is considered an anatomical imaging modality. However, implementing radiomic analysis could potentially extract more information from images than visually possible, potentially identifying not only morphologic characteristics but also radionuclide activity. When applying radiomic analysis to CCTA, the presence of sodium fluoride positron emission tomography activity could be predicted with excellent diagnostic accuracy as compared to any of the conventional qualitative or quantitative CCTA parameters (cross-validated AUC: 0.87 vs. 0.65; respectively).[47] Furthermore, even though CCTA has limited spatial resolution to identify thin-cap fibroatheroma as opposed to intravascular imaging techniques, it seems that implementing radiomics can increase the diagnostic accuracy of CCTA to identify intravascular ultrasound attenuated plaque and optical coherence tomography thin-cap atheroma as compared to currently used qualitative and quantitative parameters (cross-validated AUC: 0.72 vs. 0.59; 0.80 vs. 0.66; respectively).[47]

4.2. Machine learning

Applying ML to the CONFIRM registry revealed that using clinical CTA and anthropometric parameters the diagnostic accuracy to predict 5-year all-cause mortality could be increased as compared to the Framingham risk score and CTA based risk scores.[31] The authors also reported the relative importance of parameters using IG in a univariate fashion which indicated that several CTA parameters are linked to all-cause mortality. A similar investigation looked at the composite endpoint of myocardial infarction and death at a mean follow-up time of 4.6 years. As opposed to the previous investigation the authors trained a ML model incorporating segmental stenosis and calcification information without any anthropometric data. The authors of the latter manuscript split their data into a training set and a separate test set which was not used for training, only the evaluation of the models, as opposed to the previous investigation. This method is superior to cross-validation which may significantly overestimate the diagnostic accuracy.[48,49]

Not only clinical CCTA and anthropometric information can be used as inputs to ML, but also vessel geometry which can be used to predict CT derived fractional flow reserve (FFR).[50] In this CT-FFR algorithm, the ML model which used DNN was trained on 12 000 synthetic coronary vessel models. However, the model was trained using computational fluid dynamic results, and therefore could not increase the

diagnostic accuracy of ML based FFR-CT as compared to the computational fluid dynamics model (AUC: 0.89 vs. 0.89, p = 0.41; respectively). This shows a training bias, which is a limitation of ML and DL as the model accuracy influenced by the training dataset.[51]

ML can also be used for improving the tissue characterization capability of CT. In a single-center investigation using intravascular ultrasound as a reference, histogram parameters of plaques on CTA were used to teach a ML model to differentiate fatty from fibro-fatty plaques. The ML model outperformed simple thresholding (cross-validated AUC: 0.92 vs. 0.83, p = 0.001; respectively).[52]

4.3. Deep learning

Convolutional neural networks have been used to provide a semi-automatic method for coronary tree extraction. With DL the authors were able to achieve 93.7% overlap with manually annotated reference centerlines from an external validation cohort.[53] These results show that DL may be able provide automated workflows and assist evaluation of radiological images. DL could also increase the diagnostic accuracy of our imaging modalities. Implementing DL for the characterization of the myocardium from CCTA images could help to identify hemodynamically significant stenoses. In a cohort of 126 patients, the researchers found that combining diameter stenosis with DL evaluation of the myocardium significantly improved the diagnostic accuracy to identify hemodynamically significant stenoses as compared to just diameter stenosis (cross validated AUC: 0.76 vs. 0.68; respectively).[54,55]

5. Conclusions

Similar to how the advancements in medical imaging hardware had a significant impact on cardiac CT imaging and moved this anatomical imaging test towards the direction of comprehensive CAD assessment; new image analytic techniques such as radiomics, ML and DL will reshape the field of cardiac imaging by increasing the amount of quantitative information extracted from CT datasets. This has the potential to further our understanding of CAD and provide a more precise diagnostics and prognostication. However, these techniques are still in their infancy and the way to routine clinical application needs to overcome many potential pitfalls and will require rigorous testing and the development of a dedicated legislative frameworks. Nevertheless, as fast as AI is transforming our everyday lives, these changes may come sooner than later.

Conflicts of interest

NA.

Acknowledgements

Márton Kolossváry received support from the ÚNKP-18-3-I-SE-1 new national excellence program of the Ministry of Human Capacities of Hungary and also from the National Research, Development and Innovation Office of Hungary (NKFIA; NVKP-16-1-2016-0017).Carlo N. De Cecco received support from Siemens Healthcare.Pál Maurovich-Horvat received support from the National Research, Development and Innovation Office of Hungary (NKFIA; NVKP-16-1-2016-0017).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jcct.2019.04.007.

References

1. Douglas PS, Hoffmann U, Patel MR, et al. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med.* 2015;372:1291–1300.
2. Investigators S-H, Newby DE, Adamson PD, et al. Coronary CT angiography and 5-year risk of myocardial infarction. *N Engl J Med.* 2018;379:924–933.
3. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease: results from the prospective multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) trial. *J Am Coll Cardiol.* 2008;52:1724–1732.
4. Mark DB, Federspiel JJ, Cowper PA, et al. Economic outcomes with anatomical versus functional diagnostic testing for coronary artery disease. *Ann Intern Med.* 2016;165:94–102.
5. Lee SP, Jang EJ, Kim YJ, et al. Cost-effectiveness of coronary CT angiography in patients with chest pain: comparison with myocardial single photon emission tomography. *J Cardiovasc Comput Tomogr.* 2015;9:428–437.
6. Min JK, Gilmore A, Budoff MJ, Berman DS, O'Day K. Cost-effectiveness of coronary CT angiography versus myocardial perfusion SPECT for evaluation of patients with chest pain and no known coronary artery disease. *Radiology.* 2010;254:801–808.
7. Arbab-Zadeh A, Miller JM, Rochitte CE, et al. Diagnostic accuracy of computed tomography coronary angiography according to pre-test probability of coronary artery disease and severity of coronary arterial calcification. The CORE-64 (Coronary Artery Evaluation Using 64-Row Multidetector Computed Tomography Angiography) International Multicenter Study. *J Am Coll Cardiol.* 2012;59:379–387.
8. Leber AW, Knez A, Becker A, et al. Accuracy of multidetector spiral computed tomography in identifying and differentiating the composition of coronary atherosclerotic plaques: a comparative study with intracoronary ultrasound. *J Am Coll Cardiol.* 2004;43:1241–1247.
9. Maurovich-Horvat P, Ferencik M, Voros S, Merkely B, Hoffmann U. Comprehensive plaque assessment by coronary CT angiography. *Nat Rev Cardiol.* 2014;11:390–402.
10. Kolossvary M, Szilveszter B, Merkely B, Maurovich-Horvat P. Plaque imaging with CT-a comprehensive review on coronary CT angiography based risk assessment. *Cardiovasc Diagn Ther.* 2017;7:489–506.
11. Feuchtner G, Kerber J, Burghard P, et al. The high-risk criteria low-attenuation plaque < 60 HU and the napkin-ring sign are the most powerful predictors of MACE: a long-term follow-up study. *Eur Heart J Cardiovasc Imaging.* 2017;18:772–779.
12. Conte E, Annoni A, Pontone G, et al. Evaluation of coronary plaque characteristics with coronary computed tomography angiography in patients with non-obstructive coronary artery disease: a long-term follow-up study. *Eur Heart J Cardiovasc Imaging.* 2017;18:1170–1178.
13. Otsuka K, Fukuda S, Tanaka A, et al. Napkin-ring sign on coronary CT angiography for the prediction of acute coronary syndrome. *JACC Cardiovasc Imaging.* 2013;6:448–457.
14. Ahmadi A, Leipsic J, Ovrehus KA, et al. Lesion-specific and vessel-related determinants of fractional flow reserve beyond coronary artery stenosis. *JACC Cardiovasc Imaging.* 2018;11:521–530.
15. Maurovich-Horvat P, Schlett CL, Alkadhi H, et al. The napkin-ring sign indicates advanced atherosclerotic lesions in coronary CT angiography. *JACC Cardiovasc Imaging.* 2012;5:1243–1252.
16. Maroules CD, Hamilton-Craig C, Branch K, et al. Coronary artery disease reporting and data system (CAD-RADS(TM)): inter-observer agreement for assessment categories and modifiers. *J Cardiovasc Comput Tomogr.* 2018;12:125–130.
17. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* 2016;278:563–577.
18. Kolossvary M, Kellermayer M, Merkely B, Maurovich-Horvat P. Cardiac computed tomography radiomics: a comprehensive review on radiomic techniques. *J Thorac Imaging.* 2018;33:26–34.
19. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics.* 2017;37:1483–1503.
20. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14:749–762.
21. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern.* 1973;3:610–621.
22. Galloway MM. Texture analysis using gray level run lengths. *Comput Graph Image Process.* 1975;4:172–179.
23. Thibault G, Angulo J, Meyer F. Advanced statistical matrices for texture characterization: application to cell classification. *IEEE Trans Biomed Eng.* 2014;61:630–637.
24. Sun C, Wee WG. Neighboring gray level dependence matrix for texture classification. *Comput Vis Graph Image Process.* 1983;23:341–352.
25. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern.* 1989;19:1264–1274.
26. Chen C, DaPonte JS, Fox MD. Fractal feature analysis and classification in medical imaging. *IEEE Trans Med Imaging.* 1989;8:133–142.
27. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol.* 2016;13:350–359.
28. Russell SJ, Norvig P. *Artificial Intelligence : A Modern Approach.* third ed. Upper Saddle River, N.J.: Prentice Hall; 2010.
29. Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev.* 1959;3:210–229.
30. Singh G, Al'Aref SJ, Van Assen M, et al. Machine learning in cardiac CT: basic concepts and contemporary data. *J Cardiovasc Comput Tomogr.* 2018;12:192–201.
31. Motwani M, Dey D, Berman DS, 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;38:500–507.
32. Goodfellow I, Bengio Y, Courville A. *Deep Learning.* Cambridge, Massachusetts: The MIT Press; 2016.
33. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *IEEE*

Conference On Computer Vision And Pattern Recognition (CVPR)2016. 2016; 2016:770–778.
34. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science.* 2006;313:504–507.
35. Karras T, Laine S, Aila T. *A Style-Based Generator Architecture for Generative Adversarial Networks.* 2018; 2018 arXiv e-prints.
36. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE.* 1998;86:2278–2324.
37. Hubel DH. Single unit activity in striate cortex of unrestrained cats. *J Physiol.* 1959;147:226–238.
38. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology.* 2019;290:590–606.
39. Shrikumar A, Greenside P, Kundaje A. *Learning Important Features Through Propagating Activation Differences.* 2017; 2017 arXiv e-prints.
40. Lundberg SM, Erion GG, Lee S-I. *Consistent Individualized Feature Attribution for Tree Ensembles.* 2018; 2018 arXiv e-prints.
41. Kolossváry M, Szilveszter B, Karády J, Drobni ZD, Merkely B, Maurovich-Horvat P. Effect of image reconstruction algorithms on volumetric and radiomic parameters of coronary plaques. J Cardiovasc Comput Tomogr. https://doi.org/10.1016/j.jcct.2018.11.004.
42. Goodfellow IJ, Shlens J, Szegedy C. *Explaining and harnessing adversarial examples.* 2014; 2014 ArXiv e-prints.
43. Moosavi-Dezfooli S-M, Fawzi A, Fawzi O, Frossard P. *Universal adversarial perturbations.* 2016; 2016 ArXiv e-prints.
44. Su J, Vasconcellos Vargas D, Kouichi S. *One pixel attack for fooling deep neural networks.* 2017; 2017 ArXiv e-prints.
45. Maurovich-Horvat P, Hoffmann U, Vorpahl M, Nakano M, Virmani R, Alkadhi H. The napkin-ring sign: CT signature of high-risk coronary plaques? *JACC Cardiovasc Imaging.* 2010;3:440–444.
46. Kolossvary M, Karady J, Szilveszter B, et al. *Radiomic features are superior to conventional quantitative computed tomographic metrics to identify coronary plaques with napkin-ring sign.* Circ Cardiovasc Imaging; 2017.
47. Kolossvary M, Park J, Bang JI, et al. Identification of invasive and radionuclide imaging markers of coronary plaque vulnerability using radiomic analysis of coronary computed tomography angiography. *Eur Heart J Cardiovasc Imaging.* 2019https://doi.org/10.1093/ehjci/jez033.
48. van Rosendael AR, Maliakal G, Kolli KK, et al. Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry. *J Cardiovasc Comput Tomogr.* 2018;12:204–209.
49. Géron A. Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems. In: Tache N, ed. first ed. ed O'Reilly Media, Inc.; 2017:541.
50. Tesche C, De Cecco CN, Baumann S, et al. Coronary CT angiography-derived fractional flow reserve: machine learning algorithm versus computational fluid dynamics modeling. *Radiology.* 2018;288:64–72.
51. Coenen A, Kim YH, Kruk M, et al. Diagnostic accuracy of a machine-learning approach to coronary computed tomographic angiography-based fractional flow reserve: result from the MACHINE consortium. *Circ Cardiovasc Imaging.* 2018;11:e007217.
52. Masuda T, Nakaura T, Funama Y, et al. Machine-learning integration of CT histogram analysis to evaluate the composition of atherosclerotic plaques: validation with IB-IVUS. *J Cardiovasc Comput Tomogr.* 2018.
53. Wolterink JM, van Hamersvelt RW, Viergever MA, Leiner T, Isgum I. Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier. *Med Image Anal.* 2018;51:46–60.
54. van Hamersvelt RW, Zreik M, Voskuil M, Viergever MA, Isgum I, Leiner T. Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis. *Eur Radiol.* May 2019;29(5):2350–2359.
55. Zreik M, Lessmann N, van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* 2018;44:72–85.