Beyond the Artificial Intelligence Hype What Lies Behind the Algorithms and What We Can Achieve

van Assen, Marly; Banerjee, Imon; De Cecco, Carlo N.

Published in:
Journal of thoracic imaging

DOI:
10.1097/RTI.0000000000000485

Publication date:
2020

Beyond the Artificial Intelligence Hype
What Lies Behind the Algorithms and What We Can Achieve

Marly van Assen, MSc, PhD,† Imon Banerjee, MSc, PhD,‡§ and Carlo N. De Cecco, MD, PhD†*

Abstract: The field of artificial intelligence (AI) is currently experiencing a period of extensive growth in a wide variety of fields, medicine not being the exception. The base of AI is mathematics and computer science, and the current fame of AI in industry and research stands on 3 pillars: big data, high performance computing infrastructure, and algorithms. In the current digital era, increased storage capabilities and data collection systems, lead to a massive influx of data for AI algorithm. The size and quality of data are 2 major factors influencing performance of AI applications. However, it is highly dependent on the type of task at hand and algorithm chosen to perform this task. AI may potentially automate several tedious tasks in radiology, particularly in cardiothoracic imaging, by pre-readings for the detection of abnormalities, accurate quantifications, for example, oncologic volume lesion tracking and cardiac volume and image optimization. Although AI-based applications offer great opportunity to improve radiology workflow, several challenges need to be addressed starting from image standardization, sophisticated algorithm development, and large-scale evaluation. Integration of AI into the clinical workflow also needs to address legal barriers related to security and protection of patient-sensitive data and liability before AI will reach its full potential in cardiothoracic imaging.

Key Words: artificial intelligence, cardiac imaging, thoracic imaging, radiology

(*J Thorac Imaging* 2020;35:S3–S10)

ARTIFICIAL INTELLIGENCE (AI) AND ITS JOURNEY

According to the Britannica definition, AI is the ability of computer-controlled agents (eg, software, robots) to perform cognitive tasks that broadly cover 4 core behaviors—learning, reasoning, perception, and action. The era of AI was originally initiated in the 1950s with a simple question posed by mathematician Alan Turin—"Can a machine think?"[1] This definition of AI has given rise to many debates, and, even after 80 years, no singular definition of the field is universally accepted. Although at its core AI is a branch of computer science, it is becoming more and more interdisciplinary, including fields ranging from material science to aeronautics; circuit design to greenhouse technology; and precision health to car technology.[2]

From the *University Medical Center Groningen, Faculty of Medicine, University of Groningen, Groningen, The Netherlands; †Department of Radiology and Imaging Sciences, Divisions of Cardiothoracic Imaging, Nuclear Medicine and Molecular Imaging; §Department of Radiology and Imaging Sciences, Emory University Hospital; and ‡Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA.
C.N.D.C. receives institutional funding from Siemens. The remaining authors declare no conflicts of interest.
Correspondence to: Marly van Assen, MSc, PhD, Emory University School of Medicine, Atlanta, GA 30322 (e-mail: marly.v.assen@gmail.com).

DOI: 10.1097/RTI.0000000000000485

In 1954, AI originally started with complex mathematical problem solving, proving logical theorems, and decryption of coded messages with the demonstration of machine translation in the Georgetown-IBM experiment, and stayed within a narrow range of usability in defense-related applications for a significant time-period. AI encountered a few years of "AI-winter" beginning in 1973, which was triggered by the report "Artificial Intelligence: A General Survey" authored by Professor Sir James Lighthill to the British Science Research Council where he expressed great disappointment in AI by stating that the existing AI techniques worked well only in a research environment, but were inadequate in a real-world setting. This report was followed by pessimism in the press, leading to a severe cutback in funding, resulting in the termination of AI research.

What changed that made AI so popular again in the 90s?—In the 90s, AI was introduced in larger industrial expert-systems[3] and achieved heavy commercial successes, particularly in machine translation, data mining, search engines, and robotics.

Although commercially successful, AI started with large industrial automation, and, in the current time, AI has penetrated our daily life with several successful products and applications, from virtual digital assistants such as Siri (Apple, Cupertino, CA) and Alexa (Amazon, Seattle, WA), which help find information and schedule appointments, to self-driving cars (Tesla, Palo Alto, CA; Uber, San Francisco, CA; Waymo [Google spin-off], Mountain View, CA). Amazon's transactional AI shopping engine is another successful example that's been in existence for quite some time and resulted in an astronomical increase in profit. The base of AI is mathematics and computer science, but the current fame of AI in industry and research stands on the following 3 pillars (Fig. 1):

(1) Big data: data sets are growing rapidly, as they are now gathered by cheap and numerous information-sensing devices such as mobile devices, aerial cameras, microphones, radio frequency readers, and wireless sensor networks. Every human being is roughly contributing 2 MB data per second,[4] which can potentially be used to train AI algorithms. Adequate AI modeling of a complex task with deep learning techniques[5] needs large heterogenous training data sets to be able to utilize the large number of free trainable parameters in the model, which can run in millions. *A rule of thumb* is that the AI models need 10 times the amount of data to be trained in relation to the number of parameters. Thus AI and big data are now seemingly inseparable.
(2) High performance computing (HPC) infrastructure: local computing infrastructure with limited computing resources and memory is not capable to deal with such big data sets. AI algorithms require immensely powerful processes across computer, networking, and storage, with highest benefit when computational resources are closest to the origin of data. This was a limiting factor of

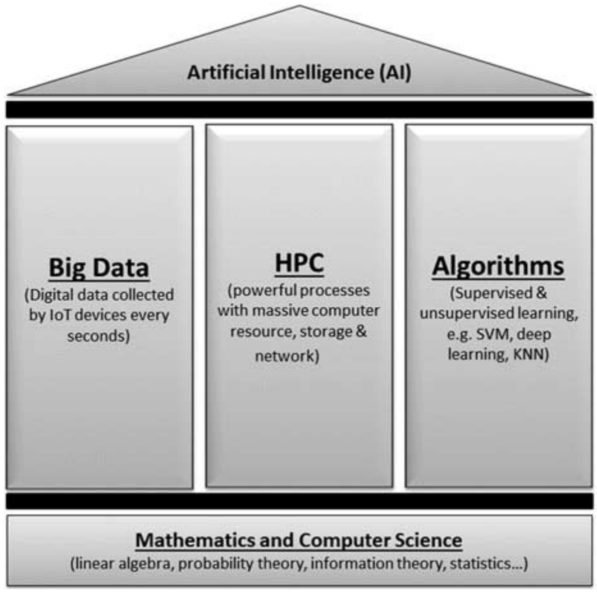

FIGURE 1. Three pillars of AI: big data, high performance computing, and algorithm.

AI utilization in the early days when AI was captured only by a few resourceful people. With the availability of HPC cloud computing by the major companies, training AI algorithms with extremely large data sets has now become achievable for a wide range of people from their home environment. HPC proved to be a game changer for AI adaptation and scientific discoveries in this field.

(3) Algorithms: the true AI's capabilities come from the machine learning (ML) algorithms, which can broadly be categorized as supervised and unsupervised learning algorithms. Supervised learning refers to the process of learning associations from training data sets wherein the algorithm learns a mapping function ($y = f(x)$), which maps the input variables (x) to an output variable (Y). Models such as support vector machine (SVM) and logistic regression fall into the category of supervised learning. In unsupervised learning, given the input variables (x) only, the algorithm explores the internal structure of the data. K-Nearest Neighbors, and K-means can be categorized as unsupervised learning algorithms when the model learns the structure of the data space without learning an association with targeted outcome. With the recent advancement in big data and HPC, ML algorithms are getting more and more complex, with millions of densely connected processing nodes (neurons), which attempt to replicate the functionality of the human brain categorized as "Deep learning" with multiple layers of processing. Deep learning (DL) algorithms are basically a class of ML algorithms with more processing power to perceive nonlinear structure within the data.[5]

AI IS HERE TO HELP: APPLICATION IN RADIOLOGY

In the digital era, volume of radiologic imaging examinations per day at any standard care center is growing enormously, while the number of available trained readers has stayed constant. Recently, this created a huge bottleneck in clinical practice and dramatically increased the workload for radiologists.[6] The primary goal behind the emergence of AI in radiology has always been the desire to assist the radiologist with the increasing workload by creating an automatic agent that can do the tedious tasks (eg, volumetric analysis, outline, measurement, reporting) and help with initial image interpretation, thereby reducing the workload.

The current AI applications are designed in such a way that they are integrated within the current radiology workflow, with the main goal being to assist radiologists with their day-to-day tasks by increasing efficiency, reducing errors, and achieving objectives with minimal manual input provided by trained radiologists. In addition, substantial efforts and policies are being put forward to facilitate technological advances related to AI in medical imaging. In the following subsection, we are highlighting a few types of AI tasks in medical imaging that are either the current focus of AI research or already have been successfully tackled by the AI components, followed by some core ingredients of success.

AI TASKS AND ALGORITHMS

Classification

One of the most popular uses of AI algorithms is for radiologic image classification tasks, in which the algorithm takes an image (or volume) as input and assigns a predetermined category (ie, differential diagnosis, clinical risk) to each data set, as seen in Figure 2. Classification models (algorithms) can be broadly categorized into the following (and they are):

(1) Linear classifier and classifier with kernel—a linear classifier without kernel computes the classification decision on the basis of the value of a linear combination of the characteristics derived from the images (ie, image features—intensity, shape, texture). Logistic regression, SVMs, and Perceptron all fall into this category. Linear classifiers are useful when the relationship between features and targeted outcome can be described with a linear function. In return, it provides nice probabilistic interpretation and can be regularized to avoid overfitting[7] and thus can be trained with limited data. However, these types of algorithms are not equipped to capture naturally complex relationships (nonlinear decision boundaries), which are often encountered in medical image classification tasks. In order to model such complex relationships, SVMs can use a trick called "kernel" to project the feature in a multidimensional space where it is linearly separable. Thus, SVM had been quite a popular solution in the medical image analysis field (magnetic resonance [MR],[8] computed tomography [CT],[9] Mammograms[10]) until the recent deep learning success.

(2) Tree-based methods and boosting —the core idea behind these methods is to stratify or segment the data space into regions on the basis of a set of splitting rules, which can be represented as a tree, for example, decision tree method. The output of tree-based algorithms are easy to interpret, as it is clarified by the rules; however, a single tree is often not capable of identifying complex boundaries between the regions. Boosting (XGBoost) and bagging (RandomForest) constructs multiple trees, which are then combined to yield a single consensus prediction. Such methods have successfully been applied in several challenging tasks, for example, classification

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

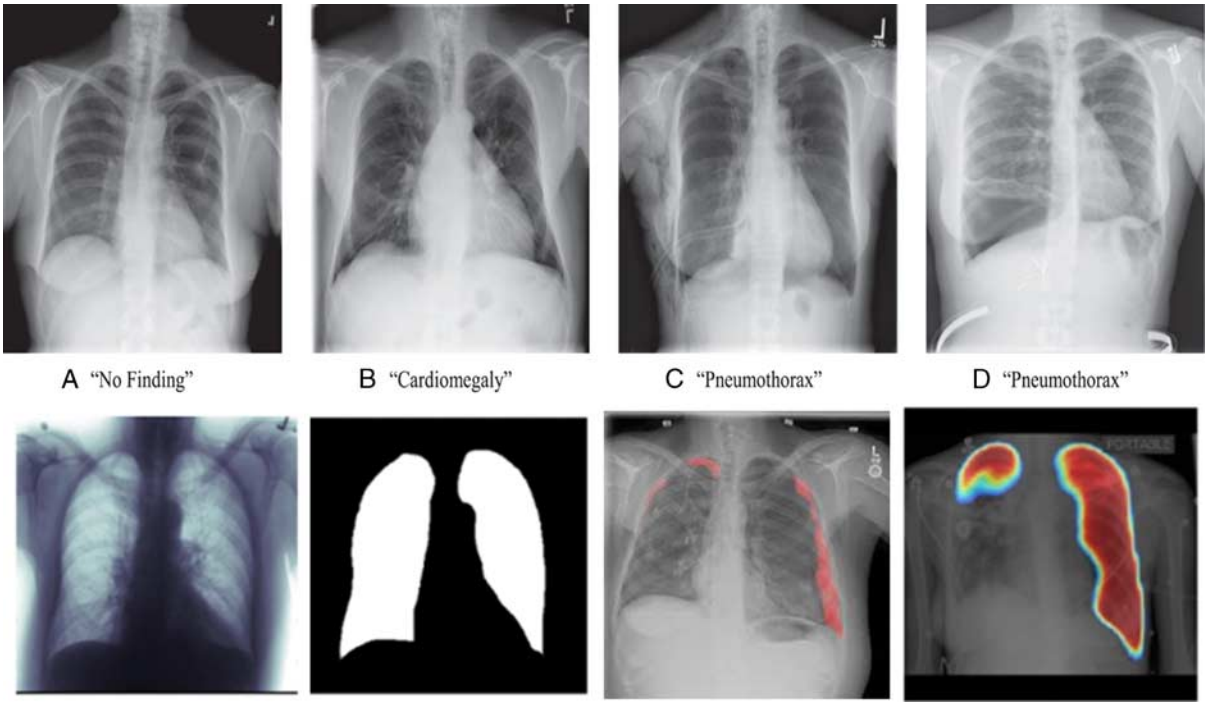

FIGURE 2. Classification (top) and segmentation (bottom) of chest radiograph.

of Alzheimer from MR volume, lung nodule classification, schizophrenia classification.[11–13]

(3) Deep neural network (DNN)—SVM, Random Forest, and XGBoost classification usually do not read the full images (or volume) as input; instead, they use handcrafted features acquired by applying several image feature extraction algorithms (eg, texture—Gabor, Riesz). Image feature engineering needs to be heavily supervised by humans and often requires extreme expertise in finding the right trade-off between accuracy and computational efficiency.[14] In the case of DNNs, the features are learned automatically from the raw images and are represented hierarchically in multiple levels; however, this comes with the cost of needing large amounts of data and computational power.

With the increased availability of open-source labeled x-ray image archives (eg, NIH ChestX-ray14, Open-i), several high-performance DNN classification algorithms have been developed for the detection of common chest diseases. Yao et al[15] presented a combination of 2 deep learning architectures—convolutional neural network (CNN) and a recurrent neural network, to exploit label dependencies in the chest data set. Rajpurkar et al[16] proposed transfer-learning with fine tuning, using a DenseNet-121, which resulted in higher area under the curve results on ChestX-ray14 for multilabel classification.

Segmentation

Image segmentation (Fig. 2) is the process of detection of boundaries of the targeted object (eg, tumor, anatomy) within 2D or 3D images based on certain image characteristics (eg, pixel value, relations with neighbors in volumetric space). Segmentation can be performed by solely computer vision techniques without any learning. For example, region growing, watershed, and level-set algorithms were successfully applied for the medical image segmentation, but with limited generalizability, and they are highly prone to noise.[17–19]

Segmentation can also be considered as a special case of supervised classification wherein the target is to assign a classification label to each pixel. CNN—a deep learning approach was successfully translated to a fully convolutional network that generates a pixel-level classification map as output by upsampling, tackling the image segmentation problem. In medical image processing, however, a slightly modified version of fully convolutional network —UNet (multiple upsampling layers)—is the most popular solution because it works with fewer training images and yields more precise segmentation (Fig. 3). UNet has been applied in segmentation from various modalities—ultrasound and MR and CT images.[20,21]

Anomaly Detection

In medical imaging applications, the unbalance between negative and positive data sets and the possible variation in degree of disease in positive data sets poses a difficulty for learning algorithms, as the outcomes will be biased toward the largest group. For anomaly detection tasks, algorithms are trained to flag data samples as being unusual or atypical. Anomaly detection can also be seen as a special case of classification with only one output class. Using anomaly detection algorithms addresses the issue of imbalanced data sets by training on negative cases only. In contrast to classification tasks that require large amounts of training data for each specific output label, anomaly detection algorithms can be trained on relatively small data sets. However, these algorithms are only able to identify whether data sets belong to this label or not without any further specification. An example of anomaly detection is given by this study of Wei et al[22] on the MURA database, containing 40,561 images of musculoskeletal radiographs, wherein each

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

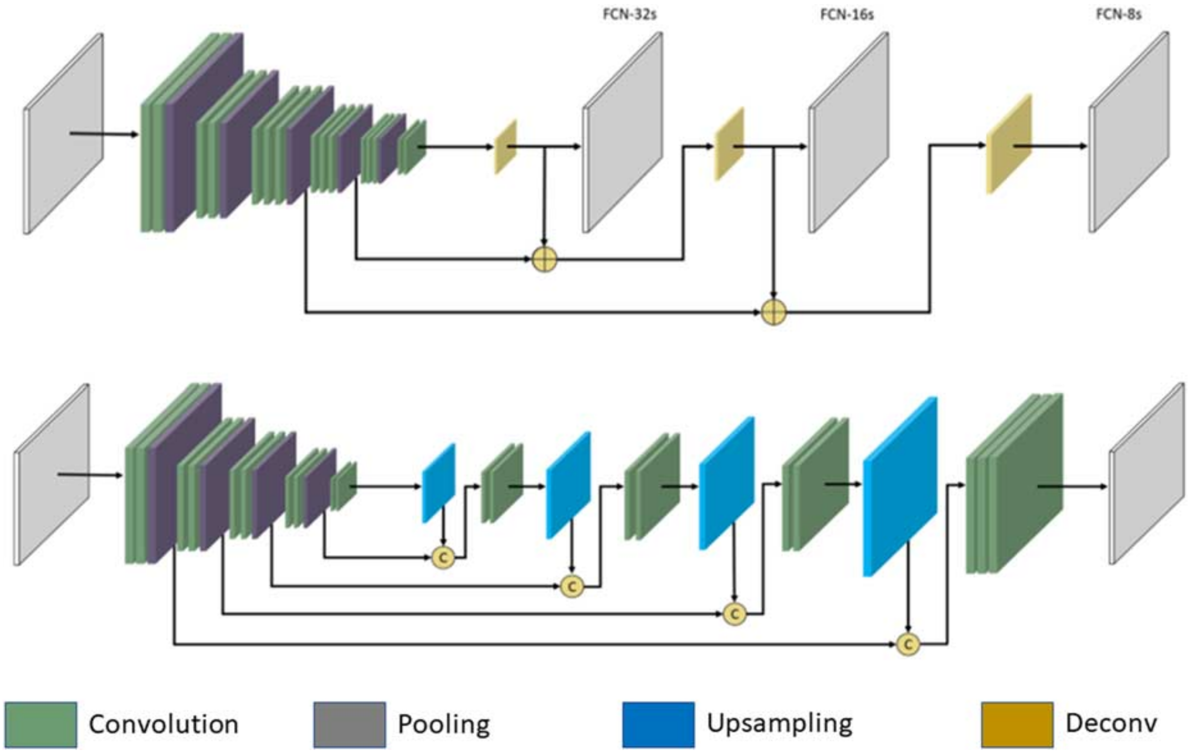

FIGURE 3. CNN-based segmentation architecture: FCN on top and UNet on bottom. full color online

study is manually labeled as either normal or abnormal. A 169-layer DenseNet baseline model, extension of the previously mentioned ResNet, was used to detect and localize abnormalities.

De-noising

De-noising algorithms are used to create a clean data sample, given a corrupted data sample as input. In the training phase, the clean sample is given as desired output, while the noise version serves at input. De-noising algorithms can play a role in optimizing image quality in images taken at lower image quality settings to reduce radiation dose. In medical imaging, CNNs have been used for noise reduction in low-dose CT scans, increasing the image quality to be equivalent to standard dose CTs.[23,24] By using low-dose CTs as input, the use of CNNs enabled the creation of CT images equal to the standard dose output CT images.

Each algorithm has its respective strengths and weaknesses, and an optimal algorithm needs to be selected on the basis of the specific task at hand. The consideration leading to the optimal choice of a specific AI model is based on the cognitive and computational complexity, desire accuracy, scalability, and interpretability of the algorithm, depending on the task and the data.

Core Ingredient of Success: Data

Besides the actual algorithm, the data used to train and validate the algorithm is another important factor. One of the main issues arising in current day medicine is the vast amount of data generated and digitally stored for each patient. This accumulation of data is caused by the increased use of electronic medical records and increased storage capabilities allowing for the collection of all sorts of data that were not previously recorded or saved. Current systems now allow for the collection and storage of data such as imaging data, interventional reports, lab values, and pathology reports. With the increase of available data, AI is a main candidate to play an essential role in the evaluation of all these data and offers the possibility of enhancing the ability of relevant data for patient care and present it in a digestible format. AI has the advantage of reviewing vast amounts of data to become proficient at using them for a wide variety of purposes.

The quality and amount of data on which the algorithm is trained are 2 important determinants of the performance and generalizability of the algorithm. Especially CNNs, one of the main AI algorithms used for imaging-related tasks, inherently need large amount of data to perform optimally. However, currently, a wide variety of protocols are being used depending on geography, image system manufacturer, and personal preferences. Combined with the variation in populations between continents, countries, and even hospitals, this variability leads to a steep increase in the number of data sets needed for adequate performance of ML algorithms. With current privacy laws and storage/sharing capabilities, there is a lack of sufficiently large data sets that are needed to train and validate AI algorithms to optimally perform their tasks. Several initiatives are being started to fill this gap in the AI workflow by combining and creating databases from multiple institutions all over the world. A second, data set related, limiting factor is the quality of the data. In the field of medical imaging, many of the analysis are carried out on the basis of visual assessment, leading to high inter-rater and intrarater variability, introducing bias to the reference standard given

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

to the ML algorithms. In order to further optimize the workings of AI algorithms for medical imaging purposes, a consistent and reliable reference label is of the essence.

IMPACT OF AI IN CARDIOTHORACIC IMAGING

Chest imaging is one of the major fields of interest for AI applications due to the high number of examinations and the availability of images in medical centers.[25] Chest radiography (CXR) is one of the most frequently performed procedures, representing a large percentage of the radiology workload in all institutions. One of the main focuses of AI application on CXR images is the automated detection of tuberculosis.[26] Tuberculosis is an important contributor to worldwide mortality, especially in underdeveloped areas, where there is a shortage of radiologists. Particularly, in this setting, automated AI applications can be of great assistance. Besides tuberculosis detection, pneumothorax and emphysema detection are other examples of AI applications.[25,27] The large amount of CXR examination performed has led to the emergence of several publicly available large databases of annotated chest radiographies. For example, the chestX-ray(8-14) database from the National Institutes of Health consists out of images corresponding to 14 different labels including atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis, effusion, pneumonia, pleural thickening, cardiomegaly, and nodules. CXR is also a field wherein anomaly detection is especially useful, differentiating normal and abnormal CXRs to help optimize clinical work-flow and help reduce reading time by the radiologist.

Another large group of chest examinations performed in radiology are the chest CTs. These types of examinations are especially increasing with the increase in large-scale lung cancer screening programs. Large lung cancer screening trials have been demonstrated to reduce lung cancer-related mortality.[28,29] It is expected that, on the basis of these results, the United States and European countries will start large scale screening. These increased numbers of chest CT examination will require a lot of resources, and AI application can help deal with the increased work load and reduce false-negative readings. In 2017, the Kaggle Data Science Bowl (KDSB17), focused on the prediction of lung cancer risk, because of the probability that a patient will be diagnosed with lung cancer within a year of the scan, on the basis of lung cancer screening CT examinations reaching accuracies of 94% using CNNs.[30,31] Compared with radiologists, the AI performance was proven to be equivalent or higher. Although nodule detection is probably the best documented field of AI research in chest CT imaging, the use of AI is not limited to nodule evaluation only and can also be applied to diagnose and stage chronic obstructive pulmonary disease and tuberculosis, and it can be used in the prediction of acute respiratory distress syndrome and mortality in smokers.[32]

Besides chest imaging, cardiac imaging is also gaining interest of the AI field. Cardiovascular diseases (CVDs) are a large attributor to the global mortality rate. A total of 17.9 million people die from CVDs every year, which accounts for 31% of all global deaths.[33] Cardiac imaging is a field that has been undergoing rapid innovation due to technological developments in hardware systems such as the imaging system itself, as in imaging analysis methods, allowing more complex forms of evaluation. With the increasing role of noninvasive cardiac evaluation in the clinical work-up of patients with or suspected of a cardiac disease, the field of cardiac imaging grows in volume and in complexity. With the focus on pattern recognition, inherent to the imaging nature of the field, AI holds great promise to help further this specific field in medicine.

For instance, noncontrast coronary artery calcium (CAC) scoring is computed to determine the presence and extent of atherosclerotic CVD, as it has proven to be an accurate risk factor for future cardiovascular events.[34,35] The number of CAC scoring acquisitions is rapidly increasing. CAC scoring is a fairly simple but time intensive task but needs manual segmentation and classification of calcified plaques throughout the entire coronary tree. Therefore, CAC scoring has been one of the first tasks tackled with AI algorithms. Research has shown that AI allows for CAC scoring, not only on dedicated CAC acquisitions, but also on chest CTs used for lung cancer screening.[36,37] Recent advances in image quality and computational capabilities also allow for the functional analysis of stenosis on CCTA images using computational fluid dynamics to calculate the fractional flow reserve (FFR). However, this approach needs extensive computational power and is time intensive. To enable real-time on-site evaluation of CT derived FFR, ML is successfully used.[38]

With the rise of new parameters such as CAC scoring and CT-FFR showing the value of imaging biomarkers for the diagnosis and treatment of CVD, new risk stratification and prognostic models, including these imaging markers, need to be developed. Because of its capabilities of analyzing large amounts of features and data, enabling the detection of complex relationships, AI algorithms are an ideal way to create these predictive models. Some early studies on relatively large data sets show the promise of using the ML algorithm for this purpose, combining clinical risk factors and imaging parameters.[39,40]

Besides cardiothoracic radiographs and CTs, AI applications in MRI are mainly focused on segmentation tasks, specifically cardiac structure segmentation for cardiac function evaluation. Before the possibilities of successfully using DNNs, multi-atlas registration and deformable shape models were the most used techniques for segmentation; however, most current approaches are currently based on CNNs.[41,42]

In the future, AI packages will become available, providing comprehensive analysis methods for pulmonary, cardiac, and bone analysis on single acquisitions. An example of one of these vendor-created software packages, combining multiple AI applications, is given in Figure 4.

In addition to the many image-related applications discussed above, there is another group of AI applications that have the potential to assist in the field of medical imaging. These are the algorithms that intervene in the planning and administrative side of the imaging field. With the use of natural language processing algorithms, applications for automated and structured reporting can be developed. In addition AI algorithms can be used to prioritize reading lists for the radiologist and optimize the work flow. All of these applications have as their main purposes the reduction in time and the increase in efficiency.

PITFALLS

Recently, ML/DL algorithms have shown excellent performance in a wide range of health care applications, which not only reflects in scientific publications but also in clinical practice. Despite all the debates about validation

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

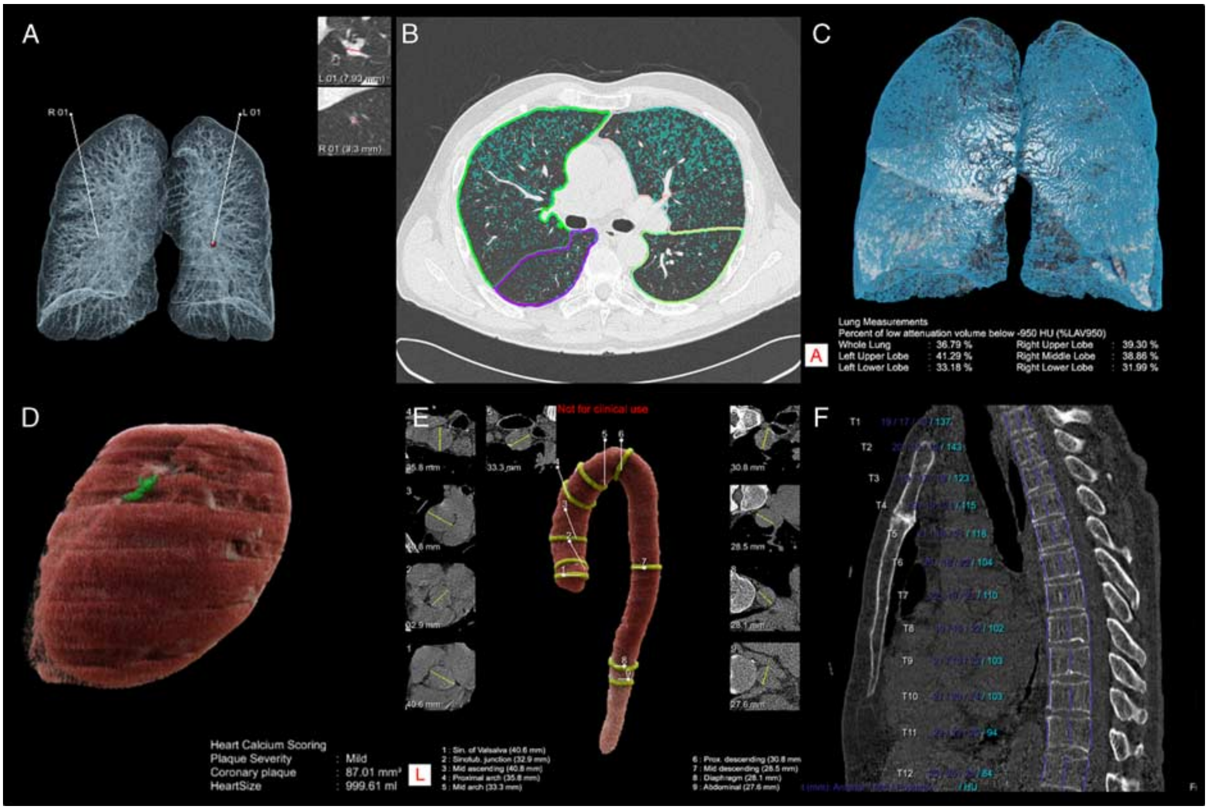

FIGURE 4. Example of a fully automated AI software program (AI-Rad Companion Chest CT, Siemens Healthineers) that allows advanced lung, cardiac, and bone assessment from chest CT images. With (A) lung nodule detection, (B, C) emphysema quantification, (D) coronary calcium and plaque analysis, (E) aorta measurements, and (F) bone density measurements.

and regulatory restriction, a wealth of new health care–focused software tools received clearance by the FDA. Interestingly, most of the tools are designed for medical image analysis with the applicability varying from analyzing CT images for stroke detection (Viz.ai's Contact, Viz.ai, San Francisco, CA) to autonomous detection of diabetic retinopathy from retinal camera images (IDx-DR, IDx, Coralville, IA). These AI systems are currently used in many clinical facilities (Iowa Health Care, IA, Johns Hopkins, Baltimore, MA, Erlanger Health System, Chattanooga, TN) across the United States. The IDx-DR solution is even currently being used in retail stores (Albertsons grocery), providing easy and convenient access to preliminary diagnosis of diabetic retinopathy. This wide span of AI systems' adoption clearly shows that AI in the health care industry is introduced as a means of smart automation to reduce burden and allow easy access to diagnostic tools.

However, AI algorithms in cardiac imaging, are struggling to make it into clinical practice. There are several factors contributing to the fact that AI is not living up to its full potential in this field. One of the main problems is the lack of large, well annotated, publicly available databases, as discussed earlier.

Another issue is the adaption by clinicians being hampered by the trust physicians put into these AI applications and the legal system protecting physicians and patients. For AI applications to be functional in clinical practice, the technical side (algorithm construction) and the clinical side (data annotation and clinical context) need to be perfectly attuned. Understanding and collaboration between the 2 sides are imperative, especially since the medical system is based on the physicians' ability to take well-informed decisions, even if they are largely based on AI algorithms often constructed by computer scientists. Therefore, it remains extremely important to provide a functional understanding of the algorithms used in a clinical context and make the AI process as transparent as possible, needing clinicians and computer scientists to maintain open and clear communication. Finally, workflow integration plays a fundamental role in the successful clinical implementation of the AI algorithm. In order to be implemented in the routine practice, implementation needs to be seamless and effortless in the busy radiologic workflow.

As with many new innovations, the regulatory system is fighting to catch up with the pace of innovation. Currently, only few regulatory efforts are finalized to help deal with the use of ML algorithms in clinical practice. Until now, physicians and other imaging professionals are still fully liable for the decisions they make, even if they are based on an ML algorithms in which they have little to no insight. With the grave consequences that come with some of these decisions, it cannot be expected that the imaging professionals involved are willing to take any risks. With regulations trying to reach consensus on how to deal with this responsibility, physicians will still be held responsible for their decisions; therefore, it remains important to increase the understanding of these algorithms and reduce the risk of mistakes.

It deserves to be noted that the creation of well-defined, annotated databases and the clinical interpretation of features and AI prediction are highly dependent on the efforts of dedicated clinicians.

In radiology, 2 types of AI applications can be distinguished, supporting applications and clinical interpretational applications. Supporting AI applications, such as quantification of coronary calcium, are currently covered by the US Food and Drug Administration (FDA), only requiring a 510(k) approval.[43] AI applications developed with the main goal of clinical interpretation of images will require FDA premarket approval (PMA), requiring results from clinical trials. In practice, this means that, for supportive AI applications, manufacturers only have to prove that they are substantially equivalent to similar legally marketed applications and do not require human trials to prove efficacy and safety. This is in contrast to the PMA approval, which requires data showing the device's performance in humans in a clinical setting, similar to the approval process of drugs. The difference in these 2 processes can have grave consequences. Between 2005 and 2009, a total of 113 PMA devices were recalled, of which 70% was cleared through the 510(k) process.[44] As with all medical applications using patient-related information, patient's privacy is an important issue. The development of AI application involves large amounts of patient data extracted from electronic health records, medical images, and lab results.

With the increased trust physicians put into the prediction of AI applications, security and safety of these algorithms becomes more and more important. With the increased use of AI, adversarial attacks on ML models have gained a lot of interest in the past years. An example is described in the paper by Thys et al,[45] showing that a small adversarial patch can hamper the accuracy of object detection using neural networks. There are other examples of wrongly interpreting AI results, clearly discussed by Cabitza et al.[46]

FUTURE PERSPECTIVE OF AI IN CARDIOTHORACIC IMAGING

The field of cardiothoracic imaging is facing new opportunities and challenges such as increasing acquisition volumes due to increased cancer screening programs and increasing numbers of cardiac examinations, now being strongly recommended in recent guidelines.

AI-based applications hold the promise to improve radiology workflow in the short term, providing pre-readings for the detection of abnormalities, accurate quantifications, for example, oncologic volume lesion tracking and cardiac volume and image optimization. For AI to be truly of assistance to the cardiothoracic imaging community, several issues need to be solved. First of all, integration into the radiologic workflow, security and protection of patient-sensitive data, especially for applications using cloud computing (cybersecurity), and liability are issues that require addressing.

In the long term, AI offers the possibility to improve patient care, providing information that visually cannot be extracted from medical images. Computers, in general, and AI algorithms specifically, are able to detect small changes, subtle deviation, and complex relationships, undetectable by human vision. This has already been demonstrated by the use of textural analysis and radiomics, using large amounts of imaging features with direct relationship to visually detectable differences. In contrast to the belief that AI will make the role of radiologists obsolete, AI is able to strengthen the position of the radiologist by making them the connecting factor between patients, data, algorithms, computer scientists, and all other specialists. AI offers the field of medical imagers the unique opportunity to be once again on the forefront of the fields by embracing technology in their clinical workflow not only easing workload and increasing efficiency but also improving patient care and patients' health, the ultimate goal of every physician.

REFERENCES

1. Turing A. Computing machinery and intelligence amplification. "Parsing the Turing Test." Comput Intell Expert Speak Springer, 2009:25–44.
2. Pannu A. Artificial intelligence and its application in different areas. *Int J Eng Innov Technol*. 2015;4:79–84.
3. Jackson P. *Introduction to Expert Systems*, 3rd ed. Boston, MA: Addison-Wesley Longman Publishing Co. Inc.; 1998.
4. James J. Data Never Sleeps 6.0. In: Domo Blog. 2012. Available at: www.domo.com/blog/wp-content/uploads/2012/06/DatainOneMinute.jpg. Accessed April 12, 2019.
5. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444.
6. Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18:500–510.
7. 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.
8. Focke NK, Yogarajah M, Symms MR, et al. Automated MR image classification in temporal lobe epilepsy. *Neuroimage*. 2012;59:356–362.
9. Lee CC, Chen SH. Gabor wavelets and SVM classifier for liver diseases classification from CT images. *Conf Proc—IEEE Int Conf Syst Man Cybern*. 2007;1:548–552.
10. Moayedi F, Azimifar Z, Boostani R, et al. Contourlet-based mammography mass classification using the SVM family. *Comput Biol Med*. 2010;40:373–383.
11. Lee SLA, Kouzani AZ, Hu EJ. Random forest based lung nodule classification aided by clustering. *Comput Med Imaging Graph*. 2010;34:535–542.
12. Yu W, Na Z, Fengxia Y, et al. Magnetic resonance imaging study of gray matter in schizophrenia based on XGBoost. *J Integr Neurosci*. 2018;17:331–336.
13. Gray KR, Aljabar P, Heckemann RA, et al. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *Neuroimage*. 2013;65:167–175.
14. Nanni L, Ghidoni S, Brahnam S. Handcrafted vs. non-handcrafted features for computer vision classification. *Pattern Recognit*. 2017;71:158–172.
15. Yao L, Poblenz E, Dagunts D, et al. Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv*. 2018:1–12.
16. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med*. 2018;15:1–17.
17. Li BN, Chui CK, Chang S, et al. Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. *Comput Biol Med*. 2011;41:1–10.
18. Ng HP, Ong SH, Foong KWC, et al. Medical image segmentation using k-means clustering and improved watershed algorithm. *Proc IEEE Southwest Symp Image Anal Interpret*. 2006;2006:61–65.
19. Pohle R, Toennies KD. Segmentation of medical images using adaptive region growing. *Med Imaging 2001 Image Process*. 2001;4322:1337–1346.
20. Manvel A, Vladimir K, Alexander T, et al. Radiologist-level stroke classification on non-contrast CT scans with deep U-Net BT—Medical Image Computing and Computer Assisted Intervention—MICCAI 2019. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap PT, Khan A, eds. Cham: Springer International Publishing; 2019:820–828.

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

21. Dubost F, Bortsova G, Adams H, et al. GP-Unet: lesion detection from weak labels with a 3D regression network BT—Medical Image Computing and Computer Assisted Intervention—MICCAI 2017. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S, eds. Cham: Springer International Publishing; 2017:214–221.
22. Wei Q, Ren Y, Hou R, et al. Anomaly detection for medical images based on a one-class classification. SPIE 10575, Med Imaging 2018 Comput Diagnosis 05751M, 2018.
23. Wolterink JM, Leiner T, Viergever MA, et al. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans Med Imaging*. 2017;36:2536–2545.
24. Wang Y, Yu M, Wang M, et al. Application of artificial intelligence-based image optimization for computed tomography angiography of the aorta with low tube voltage and reduced contrast medium volume. *J Thorac Imaging*. 2019;34:393–399.
25. Retson TA, Besser AH, Sall S, et al. Machine learning and deep neural networks in thoracic and cardiovascular imaging. *J Thorac Imaging*. 2019;34:192–201.
26. Lakhani P, Sundaram B. Deep learning at chest radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology*. 2017;284:574–582.
27. Eltorai AEM, Bratt AK, Guo HH. Thoracic Radiologists' Versus Computer Scientists' Perspectives on the Future of Artificial Intelligence in Radiology. *J Thorac Imaging*. 2019. Available at: https://doi.org/10.1097/RTI.0000000000000453.
28. Zhao YR, Xie X, De Koning HJ, et al. NELSON lung cancer screening study. *Cancer Imaging*. 2011;11:S79–S84.
29. Villa AD, Sammut E, Shome JS, et al. Combined high-resolution assessment of quantitative perfusion and late enhancement. Towards accurate estimation of the ischaemic burden in patients with coronary artery disease. *J Cardiovasc Magn Reson*. 2016;18(suppl 1):15.
30. Hammack D. Forecasting lung cancer diagnoses with deep learning. *arXiv*. 2017:1–6.
31. Kuan K, Ravaut M, Manek G , et al. Deep Learning for Lung Cancer Detection: Tackling the Kaggle Data Science Bowl 2017 Challenge, 2017.
32. González G, Ash SY, Vegas Sanchez-Ferrero G, et al. for the Copdg and E investigators. Disease staging and prognosis in smokers using deep learning in chest computed tomography. *Am J Respir Crit Care Med*. 2018;197:193–203.
33. World Health Organization. Key facts Cardiovascular Diseases (CVDs). 2017. Available at: www.who.int/cardiovascular_diseases/en/. Accessed March 18, 2019.
34. Ma S, Liu A, Carr J, et al. Statistical modeling of Agatston score in multi-ethnic study of atherosclerosis (MESA). *PLoS One*. 2010;5:e12036.
35. Agatston AS, Janowitz WR, Hildner FJ, et al. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15:827–832.
36. Kim H, Lee S-J, Davies-Venn C, et al. 64Cu-DOTA as a surrogate positron analog of Gd-DOTA for cardiac fibrosis detection with PET: Pharmacokinetic study in a rat model of chronic MI. *Nucl Med Commun*. 2016;37:188–196.
37. Takx RAP, De Jong PA, Leiner T, et al. Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability. *PLoS One*. 2014;9:e91239.
38. Tesche C, De Cecco CN, Albrecht MH, et al. Coronary CT angiography–derived fractional flow reserve. *Radiology*. 2017;285:17–33.
39. 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.
40. 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.
41. Dou Q, Yu L, Chen H, et al. 3D deeply supervised network for automated segmentation of volumetric medical images. *Med Image Anal*. 2017;41:40–54.
42. Bai W, Sinclair M, Tarroni G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks 08 Information and Computing Sciences 0801 Artificial Intelligence and Image Processing. *J Cardiovasc Magn Reson*. 2018;20:1–12.
43. Food and Drug Administration. Federal Register-Breakthrough Devices Program Draft Guidance for Industry and Food and Drug. 2017. Available at: www.federalregister.gov. Accessed January 15, 2019.
44. Zuckerman DM, Brown P, Nissen SE. Medical device recalls and the FDA approval process. *Arch Intern Med*. 2011;171:1006–1011.
45. Thys S, Van Ranst W, Goedemé T. Fooling automated surveillance cameras: adversarial patches to attack person detection. *Arxiv*. 2019;1904:08653.
46. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA*. 2017;318:517–518.

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.