Machine Learning in Medical Imaging

Maryellen L. Giger, PhD

Abstract

Advances in both imaging and computers have synergistically led to a rapid rise in the potential use of artificial intelligence in various radiological imaging tasks, such as risk assessment, detection, diagnosis, prognosis, and therapy response, as well as in multi-omics disease discovery. A brief overview of the field is given here, allowing the reader to recognize the terminology, the various subfields, and components of machine learning, as well as the clinical potential. Radiomics, an expansion of computer-aided diagnosis, has been defined as the conversion of images to minable data. The ultimate benefit of quantitative radiomics is to (1) yield predictive image-based phenotypes of disease for precision medicine or (2) yield quantitative image-based phenotypes for data mining with other -omics for discovery (ie, imaging genomics). For deep learning in radiology to succeed, note that well-annotated large data sets are needed since deep networks are complex, computer software and hardware are evolving constantly, and subtle differences in disease states are more difficult to perceive than differences in everyday objects. In the future, machine learning in radiology is expected to have a substantial clinical impact with imaging examinations being routinely obtained in clinical practice, providing an opportunity to improve decision support in medical image interpretation. The term of note is *decision support*, indicating that computers will augment human decision making, making it more effective and efficient. The clinical impact of having computers in the routine clinical practice may allow radiologists to further integrate their knowledge with their clinical colleagues in other medical specialties and allow for precision medicine.

Key Words: Machine learning, deep learning, radiomics, computer-aided diagnosis, computer-assisted decision support

J Am Coll Radiol 2018;15:512-520.

Advances in both imaging and computers have synergistically led to a rapid rise in the potential use of artificial intelligence in various radiological imaging tasks, such as risk assessment, detection, diagnosis, prognosis, and therapy response, as well as in multi-omics disease discovery. Although computer-aided detection (CADe) has been proposed, developed, and clinically used since 1966, especially in thoracic and breast imaging [1-5], the widespread progress in multiple clinical decision-making tasks and multiple disease sites has only advanced in the past decades with the corresponding access to large computational resources, including computer power, storage, and digital imaging, as well as increased electronic access to information at the time of interpretation (eg, clinical history, laboratory data, prior examinations).

A brief overview of the field is given here, allowing the reader to recognize the terminology, the various subfields, and components of machine learning, as well as the clinical potential. Figure 1 shows the number of publication counts in PubMed for searches on computer-aided diagnosis (CADx) in radiology, machine learning, and deep learning from 1972 to middle of 2017. Note that in each of these areas, there are numerous review publications; however, the aim of this article is to elucidate the concepts and generalities. The range in presentation of various subtle disease states, the need for large annotated clinical data sets, and the complex structure of many machine learning methods signify much need for continued research and development before full clinical incorporation and use.

CADe, CADx, AND DECISION SUPPORT

Medical image interpretation is the main undertaking of radiologists, with the tasks requiring both good image quality and good image interpretation. Image interpretation by humans is limited by the presence of structure noise (camouflaging normal anatomical background), incomplete

Department of Radiology, The University of Chicago, Chicago, Illinois.

Corresponding author and reprints: Maryellen L. Giger, PhD, University of Chicago, Department of Radiology, MC 2026, 5841 S Maryland Ave, Chicago, IL 60637; e-mail: m-giger@uchicago.edu.

Funded in parts by NIH U01CA195564, U01CA189240, and R01CA166945. M.L.G. is a stockholder in R2/Hologic, cofounder and equity holder in Quantitative Insights, and shareholder in QView and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba. It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest that would reasonably seem to be directly and significantly affected by the research activities.

1546-1440/17/$36.00 ■ https://doi.org/10.1016/j.jacr.2017.12.028

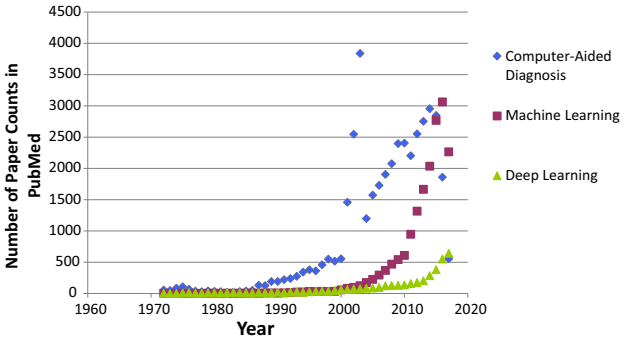

Fig 1. Number of paper counts in PubMed for searches on computer-aided diagnosis in radiology, machine learning, and deep learning from 1972 to middle of 2017.

visual search patterns, fatigue, distractions, the assessment of subtle or complex disease states, vast amounts of image data, and the physical quality of the image itself.

CADe and CADx have been under development for decades [1-5]. In fact, CADe systems have already been commercialized and have been in clinical use since the turn of the century [6]. In addition, over the past few decades, various investigators have been developing image analysis methods for CADx, such as the computer-assisted quantitative characterization of breast lesions on clinical images, as well as in the assessment of cancer risk [4].

There is no one-size-fits-all when it comes to computer algorithms and specific radiological interpretation tasks. Each computerized image analysis method requires customizations specific to the task as well as the imaging modality. For example, in breast cancer risk assessment, computer-extracted characteristics of breast density or breast parenchymal pattern are computed and related to breast cancer risk factors [7-12]. CADe methods involve a localization task and serve as a second opinion to radiologists in their task of finding suspicious regions within images, as in screening mammograms, leaving subsequent patient management decisions to the radiologist. CADx involves the characterization of a region or tumor, initially indicated by either a radiologist or a computer, after which the computer characterizes the suspicious region or lesion or estimates its probability of disease, again leaving the patient management to the physician [4].

RADIOMICS AND IMAGING GENOMICS (RADIOGENOMICS)

Effective diagnosis and treatment of disease rely on the integration of information from multiple patient tests involving clinical, molecular, imaging, and genomic data (ie, various "-omics"). Radiomics, an expansion of CADx, has been defined as the conversion of images to minable data [13-15]. Obtaining radiomic data may involve computer segmentation of a tumor from its background followed by computer extraction of various tumor features. The ultimate benefit of quantitative radiomics is to (1) yield predictive image-based phenotypes of disease for precision medicine or (2) yield quantitative image-based phenotypes for data mining with other -omics for discovery (ie, imaging genomics).

Radiomic features can be described as handcrafted or engineered, with intuitive features or deep-learned features. In this section, the focus is on handcrafted features for which computer algorithms are developed based on some analytical feature-extraction approach, such as the calculation of geometric shape of a tumor. For example, Figure 2 demonstrates a computer-aided design or radiomics pipeline for the computer extraction of various characteristics of breast tumors on dynamic

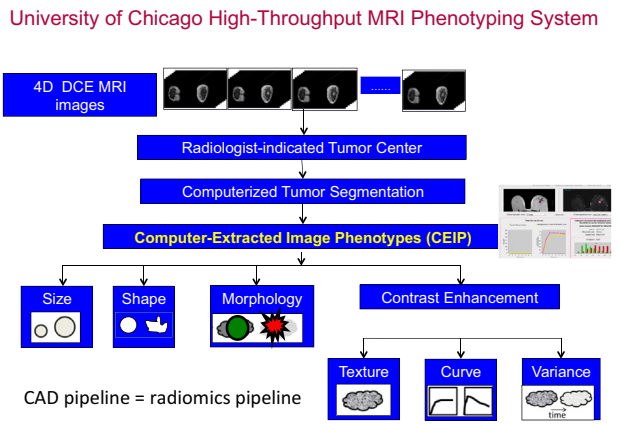

Fig 2. Schematic flowchart of a computerized tumor phenotyping system for breast cancers on DCE-MRI. The computer aided diagnosis (CAD) radiomics pipeline includes computer segmentation of the tumor from the local parenchyma and computer-extraction of "handcrafted" radiomic features covering six phenotypic categories: (1) size (measuring tumor dimensions), (2) shape (quantifying the 3-D geometry), (3) morphology (characterizing tumor margin), (4) enhancement texture (describing the heterogeneity within the texture of the contrast uptake in the tumor on the first postcontrast MRIs), (5) kinetic curve assessment (describing the shape of the kinetic curve and assessing the physiologic process of the uptake and washout of the contrast agent in the tumor during the dynamic imaging series, and (6) enhancement-variance kinetics (characterizing the time course of the spatial variance of the enhancement within the tumor) [16-21]. CAD = computer-aided diagnosis; DCE-MRI = dynamic contrast-enhanced MRI.

contrast-enhanced MRI [22,23]. After the tumor is delineated from the parenchymal background (ie, computer segmentation), the various radiomic features are calculated. The number of radiomics publications highlighting the role of quantitative imaging biomarkers is dramatically increasing with the focus going beyond CADx [4,13].

A major focus of radiomics is cancer. Cancers are spatially heterogeneous, and currently, many imaging biomarkers of cancerous tumors include only size and simple enhancement measures (if dynamic imaging is employed). Various genomic studies have demonstrated the heterogeneity of primary breast cancer tumors [24]. With radiomics, the goal is to obtain image-based phenotypes of the cancerous tumor including size, shape, margin morphology, enhancement texture, kinetics, and variance kinetic phenotypes. For example, enhancement texture phenotypes characterize the tumor texture pattern of contrast-enhanced tumors on the first postcontrast images and thus quantitatively characterize the heterogeneous nature of contrast uptake within the breast tumor [16,17,23]. For example, the larger the enhancement texture entropy, the more heterogeneous the pattern within the tumor, potentially reflecting the heterogeneous nature of angiogenesis and treatment susceptibility, serving as a location-specific "virtual digital biopsy."

A major gap in breast cancer research is the elucidation of the relationship between the macroscopic appearance of the tumor and its environment and biologic indicators of risk, prognosis, or treatment response. Imaging genomics (ie, "radiogenomics") aims to find relationships between imaging data and clinical data, molecular data, genomic data, and outcome data [25-28]. During this "discovery stage," the goal is identify associated radiomic features for later application in developing predictive models for use in risk assessment, screening, detection, diagnosis, prognosis, therapeutic response, risk of recurrence, and so on.

Basically, tumors are different, so can imaging capture the phenotypic differences and the heterogeneity within? Is it possible to decide targeted therapy based on imaging genomics association results? Can imaging features inform important genomics features? Can integration of imaging and genomics features lead to higher power in prediction? Can imaging serve as a virtual digital biopsy, because it is noninvasive, covers complete tumor, and is repeatable? It is important to note that the intention is not to use radiomics to replace conventional biopsies and genetic testing; however, from imaging genomics association studies, the aim is to ultimately understand the relationships between the image-based phenotypes and genetics, and bring imaging findings earlier into screening and treatment regimens to potentially avoid serial biopsies and provide virtual biopsies when actual biopsies are not practical.

These goals are to be achieved using image data from routine clinical imaging examinations. However, for radiomics to progress in biomedical discovery and clinical prediction, sufficient harmonization is necessary for clinical translation in terms of reproducibility and repeatability. Thus, studies have been conducted with focus on robustness of the imaging systems or robustness of the radiomic features. Various initiatives have focused on the aspects of quantitative and robustness, including those of the Quantitative Imaging Network of the National Cancer Institute [29] and the Quantitative Imaging Biomarker Alliance of the RSNA [30].

With such methods, investigators are phenotypically characterizing solid tumors to gain image-based information on the underlying genetic makeup. For example, in a multi-institutional National Cancer Institute collaboration, which used de-identified data sets of invasive breast carcinomas from The Cancer Genome Atlas and The Cancer Imaging Archive [31,32] the relationships between computer-extracted radiomic MRI tumor features and various clinical, molecular, and genomics markers of prognosis and risk of recurrence, including gene expression profiles, were investigated [23,27,28,33,34]. Statistically significant associations were seen between quantitative MRI radiomic features and various clinical, molecular, and genomic features in breast invasive carcinoma. Promising significant trends were observed between enhancement texture (entropy) and molecular subtypes (normal-like, luminal A, luminal B, HER2-enriched, basal-like), even after controlling for tumor size. Also discovered were some highly specific imaging-genomic associations, which may be potentially useful in (1) imaging-based diagnoses that can inform the genetic progress of tumor and (2) discovery of genetic mechanisms that regulate the development of tumor phenotypes. The authors noted that the computer-extracted MRI phenotypes show promise for high-throughput discrimination of breast cancer subtypes, which may yield a quantitative predictive signature for assessing prognosis.

In another example, a group characterized lung tumors on CT through a radiomic analysis of 440 features quantifying tumor image intensity, shape, and texture from 1,019 patients with lung or head-and-neck cancer

[35]. Using an independent data set, many of the radiomic features were shown to have prognostic power. The imaging-genomics association study noted that a prognostic radiomic signature, characterizing tumor heterogeneity, seemed associated with underlying gene-expression patterns. Figure 3 shows radiomics heat map from the unsupervised clustering of lung cancer patients and radiomic feature expression that revealed clusters of patients with similar radiomic expression patterns.

Radiomics also allows for the use of computer-extracted lesion features as image-based biomarkers (image-based phenotypes) in predicting a patient's response to a particular therapeutic treatment. For example, the functional tumor volume from breast MRI has been shown to a predictor of recurrence-free survival of patients undergoing neoadjuvant therapy in an evaluation using data from an ACRIN study [36].

MACHINE LEARNING

Computer-extracted (radiomic) features can serve as input to machine learning algorithms (ie, computer algorithms that "learn" a specific task given specific input data). With such machine learning methods, multiple radiomic features are merged into a single value, such as a tumor signature, which might be related to the likelihood of disease state (eg, see Clark et al [32]).

Various machine learning techniques have been applied across the decades, for example, linear discriminant analysis, support vector machines, decision trees and

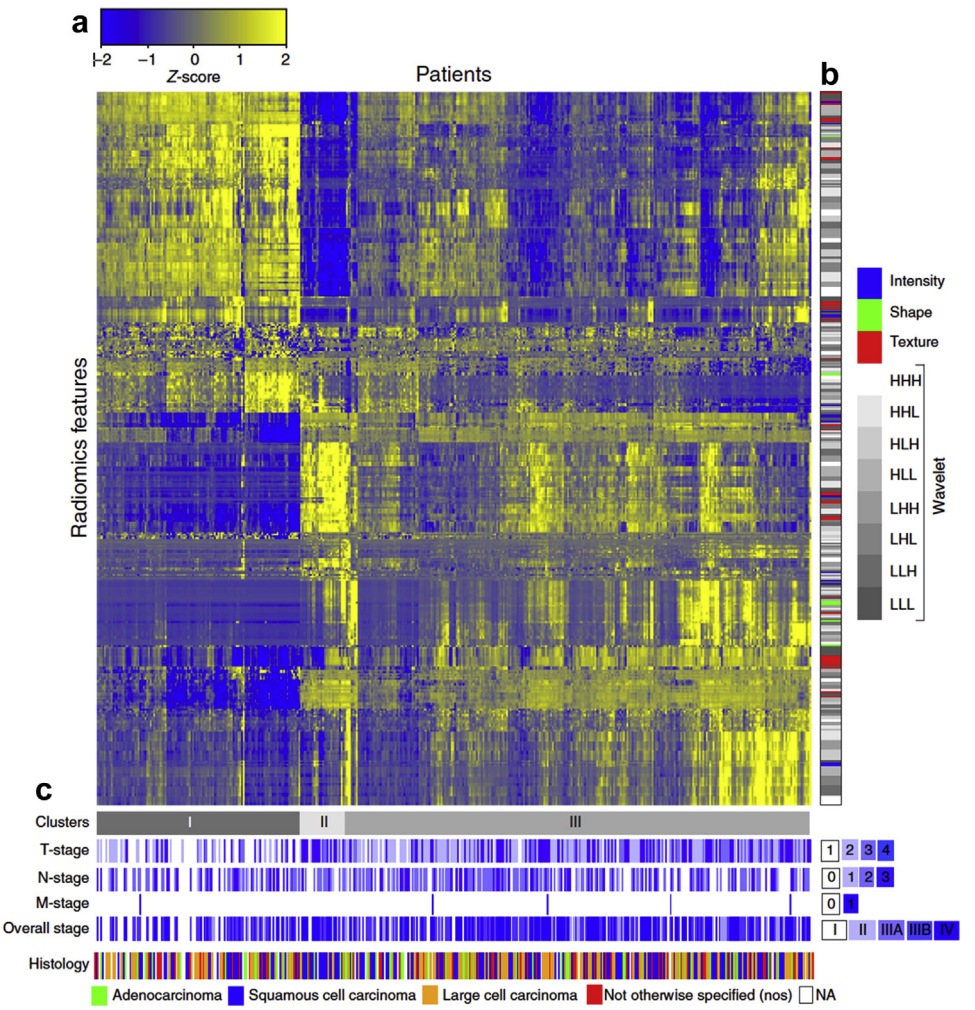

Fig 3. Radiomics heat map. (a) Unsupervised clustering of lung cancer patients (Lung1 set, n.422) on the y axis and radiomic feature expression (n.440) on the x axis revealed clusters of patients with similar radiomic expression patterns. (b) Clinical patient parameters for showing significant association of the radiomic expression patterns with primary tumor stage (T-stage; Po1_10_20, w2 test), overall stage (P.3.4_10_3, w2 test), and histology (P.0.019, w2 test). (c) Correspondence of radiomic feature groups with the clustered expression patterns. Reprinted with permission [35].

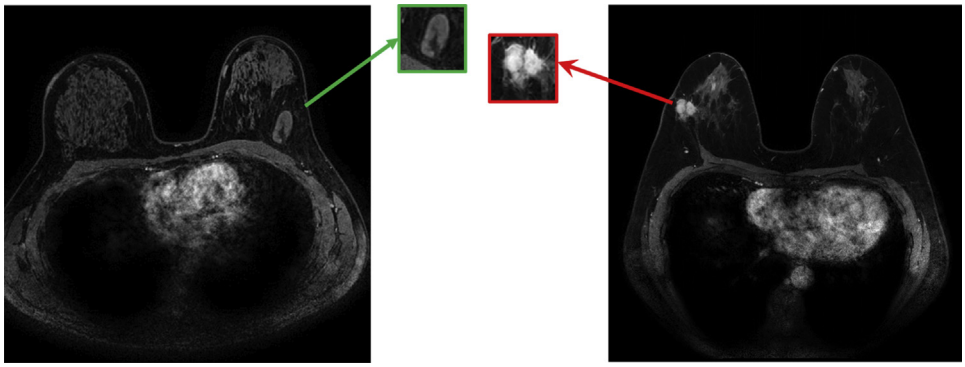

Fig 4. Examples of DCE-MRI transverse center slices with the corresponding regions of interest (ROIs) extracted. On the left is a benign case and on the right is a malignant case. These extracted ROIs are then input to a CNN for transfer learning. DCE = dynamic contrast-enhanced MRI. (Reprinted with permission [37])

random forests, and neural networks. Reviews of machine learning have been written over the past many years including those that serve as tutorials to new investigators into the field [15,38].

Given the ever-increasing variations of computer-extracted features, both handcrafted and deep-learned, appropriate feature selection techniques are important. Various studies have been conducted in which investigators, using moderately large data sets, have evaluated the combination of feature selection and classification methods [39-41]. Such analyses have taken into account both performance (such as the area under the receiver operating characteristic curve for a particular clinical task) and variability as a way to identify the optimal signature. That is, a computer-derived tumor signature needs to both perform well in its specific task and be generalizable across cases.

DEEP LEARNING

Deep learning is a subcategory of machine learning in which multiple-layered networks are used to assess complex patterns within the raw imaging input data. Most recently, deep learning has been conducted using deep convolutional neural networks (CNNs). Just as radiologists learn, during residency and beyond, by repeatedly correlating their visual interpretation of radiological images to actual clinical truth, so can machines. Although CNNs

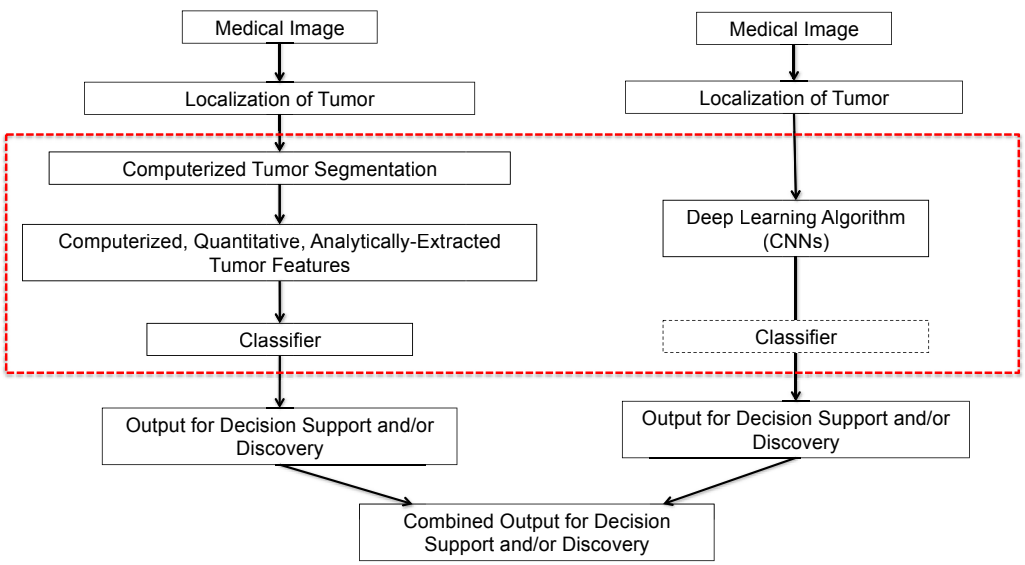

Fig 5. Schematic demonstrating the comparison of conventional hand-crafted computer-aided diagnosis and radiomic features, convolutional neural network (CNN)-extracted features, and an ensemble technique in the task of distinguishing between lesion type as used in Antropova et al [37] and Huynh et al [42].

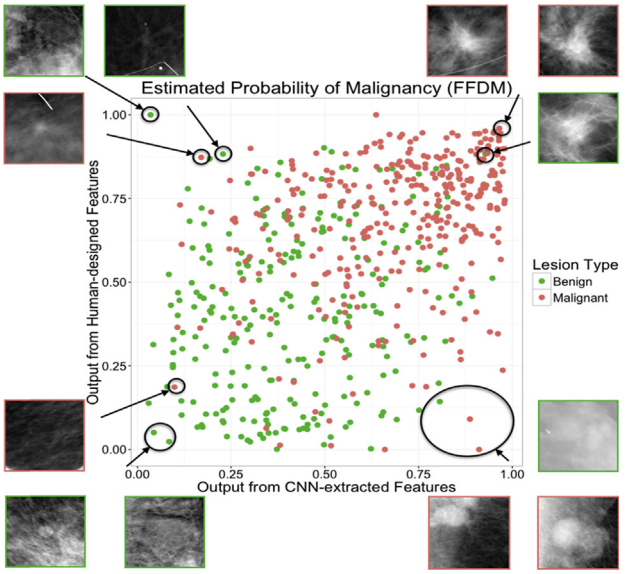

Fig 6. A diagonal classifier agreement plot between a convolutional neural network (CNN)-based classifier and a conventional computer-aided diagnosis (CADx) classifier for FFDM in the diagnostic task of estimating the probability of malignancy. The x axis denotes the output from the CNN-based classifier, and the y axis denotes the output from the conventional CADx classifier. Each point represents a region of interest (ROI) for which predictions were made. Points near or along the diagonal from bottom left to top right indicate high classifier agreement; points far from the diagonal indicate low agreement. ROI pictures of extreme examples of agreement and disagreement are included [37]. FFDM = full field digital mammography (Reprinted with permission [37])

have been used in CADe for decades, advances in computers have allowed for a dramatic increase in the number of layers within the CNN, thus resulting in the term *deep*. A comprehensive technical review of deep learning in medical image analysis is given by Shen et al [43].

Possibly the earliest journal publication of CNNs in medical imaging was in 1994 and was for the computerized detection of microcalcifications in mammography [44]. In that work, a CNN was trained to create filters within a shift-invariant artificial neural network, enabling the enhancement of microcalcifications for further analyses within a CADe system [44]. Other early uses of CNNs include a study of their use in the classification of biopsy-proven masses and normal tissue on mammograms [45].

Advances in recent years in deep learning have been quite noteworthy, with CNNs seeing great success in many benchmark image classification tasks [46-48]. However, to be trained, CNNs require very large and correctly labeled data sets, as well as substantial computational resources. Thus, implementation of deep learning in medical decision making is occurring through use of pretrained CNNs (ie, "transfer learning"—with and without "fine-tuning"). Many developments have been published demonstrating the role of transfer learning in radiology. Basically, training CNNs "from scratch" is often not possible for CAD and other medical image interpretation tasks. However, generic features can be transferred from an already-trained CNNs (ie, pretrained; eg, a CNN trained on natural scenes) to serve as features for input to classifiers focused on a medical imaging task. This process is known as transfer learning [49-52]. For example, the use of "off-the-shelf" CNNs pretrained on everyday objects, such as cats and dogs, can be used to characterize tumors on breast images [37,42] transferring knowledge from general object recognition tasks to medical imaging

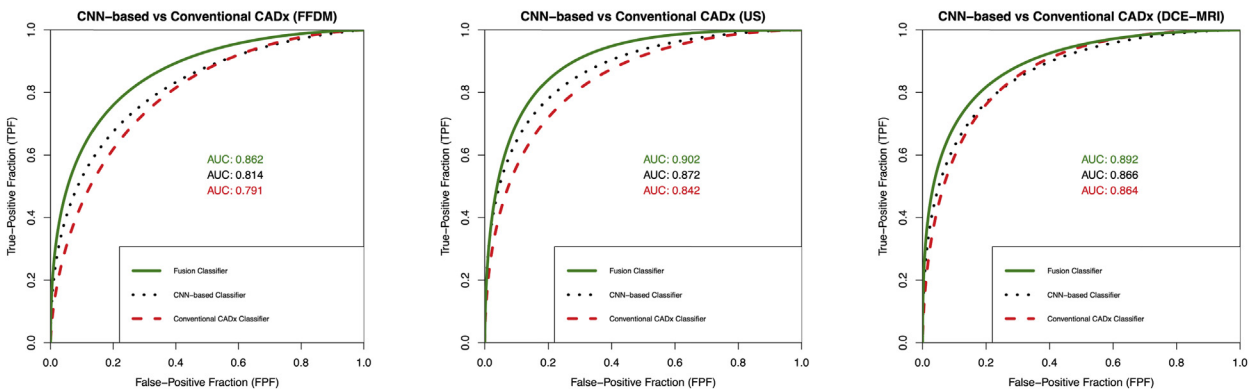

Fig 7. Receiver operating characteristic curves showing statistically significant improvement in diagnostic classification of breast lesions on FFDM, ultrasound, and breast MRI when output from conventional CADx and deep learning are combined [37]. AUC, area under the curve; CAD, computer-aided diagnosis; CNN, convolutional neural network; DCE = dynamic contrast-enhanced MRI; FFDM = full field digital mammography; US = ultrasound. Reprinted with permission [37].

classification tasks. In addition, fine-tuning of trained CNNs is occurring in which investigators use only a portion of a CNN trained for a different task and retrain the later layers of a CNN specifically for new task. These methods allow for harnessing the predictive power of deep neural networks without the need for extremely large data set or computational cost requirements.

For example, transfer learning has been successfully used in the diagnosis of breast tumors on mammography, ultrasound, and breast MRI [37,42]. In a breast imaging CADx system, deep neural networks with transfer learning were used, and specific layers of the CNN served as features for subsequent classifiers. Three scenarios were evaluated: a CADx system with computer-extracted handcrafted features, a CADx system with CNN-extracted features, and an ensemble classifier trained on both types of features. Across all three breast imaging modalities, the ensemble classifier performed best, indicating the potential for the complementary use of both handcrafted and deep-learned tumor features in medical decision making (Figs. 4, 5, 6, 7).

CNNs have been investigated for multiple tasks including the detection of colonic polyps on CT colonography as shown in Figure 1 of reference [53] and detecting patterns of interstitial lung disease on CT [54].

In another example, a CNN for detection was trained on digital mammography but then transfer learning was conducted to allow the image patterns learned from mammograms to be transferred to the analysis of breast tomosynthesis images, indicating the ability to transfer between radiological modalities [55].

Although deep learning allows for computers to learn directly from image data, for each clinical task, millions of images are needed for CNNs to be trained "from scratch." Such an example is that of detection of diabetic retinopathy in retinal fundus photographs [56].

DISCUSSION AND SUMMARY

Although many machine learning imaging publications are presented and published each year, there are still only a few methods that are able to handle the vast range of radiological presentations of subtle disease states. For example, the use of CNNs to distinguish trabecular bone structure or interstitial lung diseases involves subtle changes in texture-type patterns, which are quite different from everyday photos of cats and dogs.

The use of deep learning terminology has also caused concern in the use of a "black box" for medical tasks; however, there are methods to assess the learned parameters within a CNN to understand its decision-making focus and methodology.

In the future, to cover the entirety of radiology, there are challenges and potential pitfalls. For deep learning in radiology to succeed, recall that appropriately annotated large data sets are needed, deep networks are complex, computer software and hardware are evolving constantly, and subtle differences in disease states are more difficult to perceive than differences in everyday objects.

However, in the future, machine learning in radiology is expected to have a substantial clinical impact with imaging examinations being routinely obtained in clinical practice, providing an opportunity to improve decision support in medical image interpretation. The term of note is *decision support*, indicating that computers will augment human decision making, making it more effective and efficient. The clinical impact of having computers in the routine clinical practice may allow radiologists to further integrate their knowledge with their clinical colleagues in other medical specialties and allow for precision medicine.

TAKE-HOME POINTS

- Advances in both imaging and computers have synergistically led to a rapid rise in the potential use of artificial intelligence in various radiological imaging tasks.
- Radiomics, the -omics of images, is an expansion of CADx.
- Machine learning enables the use of radiomics in computer-learned tumor signatures.
- Deep learning, a subcategory of machine learning, allows computers to learn directly from image data; however, for each clinical task, millions of images are expected to be needed for CNNs to be trained "from scratch."
- Although many machine learning imaging publications are presented and published each year, there are still only a few methods that are able to handle the vast range of radiological presentations of subtle disease states. The range in presentation of various subtle disease states, the needs for large annotated clinical data sets, and the complex structure of many machine learning methods signify much need for continued research and development before full clinical incorporation and use.

REFERENCES

1. Lodwick GS. Computer-aided diagnosis in radiology. A research plan. Invest Radiol 1996;1:72-80.
2. Giger ML. Computerized image analysis in breast cancer detection and diagnosis. Sem Breast Dis 2002;5:199-210.
3. Giger ML, Chan H-P, Boone J. Anniversary paper: history and status of CAD and quantitative image analysis: the role of medical physics and AAPM. Med Phys 2008;35:5799-820.
4. Giger ML, Karssemeijer N, Schnabel J. Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer. Annu Rev Biomed Eng 2013;15:327-57.
5. Rao VM, Levin DC, Parker L, Cavanaugh B, Frangos AJ, Sunshine JH. How widely is computer-aided detection used in screening and diagnostic mammography? J Am Coll Radiol 2010;7:802-5.
6. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology 2001;220:781-6.
7. Alonzo-Proulx O, Packard N, Boone JM, et al. Validation of a method for measuring the volumetric breast density from digital mammograms. Phys Med Biol 2010;55:3027-44.
8. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006;25:273-82.
9. Huo Z, Giger M, Olopade O, Wolverton D, Weber B, et al. Computerized analysis of digitized mammograms of BRCA1 and BRCA2 gene mutation carriers. Radiology 2002;225:519-26.
10. Manduca A, Carston M, Heine J, et al. Texture features from mammographic images and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2009;18:837-45.
11. Nielsen M, Karemore G, Loog M, et al. A novel and automatic mammographic texture resemblance marker is an independent risk factor for breast cancer. Cancer Epidemiol 2011;35:381-7.
12. Li H, Giger M, Olopade O, Lan L. Fractal analysis of mammographic parenchymal patterns in breast cancer risk assessment. Acad Radiol 2007;14:513-21.
13. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol 2017;28:1191-206.
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563-77.
15. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: the promise of radiomics. Phys Med 2017;38:122-39.
16. Chen W, Giger ML, Bick U, Newstead G. Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI. Med Phys 2006;33:2878-87.
17. Chen W, Giger ML, Li H, Bick U, Newstead G. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. Magn Reson Med 2007;58:562-71.
18. Chen W, Giger ML, Bick U. A fuzzy c-means (FCM) based approach for computerized segmentation of breast lesions in contrast-enhanced MR images. Acad Radiol 2006;13:63-72.
19. Gilhuijs KGA, Giger ML, Bick U. Automated analysis of breast lesions in three dimensions using dynamic magnetic resonance imaging. Med Phys 1998;25:1647-54.
20. Chen W, Giger ML, Lan L, Bick U. Computerized interpretation of breast MRI: investigation of enhancement-variance dynamics. Med Phys 2004;31:1076-82.
21. Bhooshan N, Giger ML, Jansen S, Li H, Lan L, Newstead G. Cancerous breast lesions on dynamic contrast-enhanced MR images: computerized characterization for image-based prognostic markers. Radiology 2010;254:680-90.
22. Chen W, Giger ML, Newstead GM, et al. Computerized assessment of breast lesion malignancy using DCE-MRI: robustness study on two independent clinical datasets from two manufacturers. Acad Radiol 2010;17:822-9.
23. Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. NPJ Breast Cancer 2016:2. pii: 16012; Epub May 11, 2016.
24. Parker JS, Perou CM. Tumor heterogeneity: focus on the leaves, the trees, or the forest? Cancer Cell Previews 2015;28:149-50.
25. Li H, Giger ML, Sun C, et al. Pilot study demonstrating association between breast cancer image-based risk phenotypes and genomics biomarkers. Med Phys 2014;41:031917.
26. Gierach GL, Li H, Loud JT, et al. Relationships between computer-extracted mammographic texture pattern features and BRCA1/2 mutation status: a cross-sectional study. Breast Cancer Res 2014;23:424.
27. Zhu Y, Li H, Guo W. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. Sci Rep 2015;5:17787.
28. Guo W, Li H, Zhu Y, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. J Med Imaging (Bellingham) 2015;2:041007.
29. National Cancer Institute. Quantitative imaging network (QIN). Available at: https://imaging.cancer.gov/programs_resources/specialized_initiatives/qin.htm. Accessed January 1, 2015.
30. RSNA. Quantitative Imaging Biomarkers Alliance® (QIBA®). Available at: https://www.rsna.org/qiba/. Accessed January 25, 2018.
31. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61.
32. Clark K, Bruce V, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 2013;26:1045.
33. Burnside E, Drukker K, Li H, et al. Using computer-extracted image phenotypes from tumors on breast MRI to predict breast cancer pathologic stage. Cancer 2016;122:748-57.
34. Li H, Zhu Y, Burnside ES, et al. MRI radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of gene assays of MammaPrint, Oncotype DX, and PAM50. Radiology 2016;281:382-91.
35. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
36. Hylton NM, Gatsonis CA, Rosen MA, et al. Neoadjuvant chemotherapy for breast cancer: functional tumor volume by MR imaging predicts recurrence-free survival-results from the ACRIN 6657/CALGB 150007 I-SPY 1 trial. Radiology 2016;279:44-55.
37. Antropova N, Huynh BQ, Giger ML. A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys 2017;44:5162-71.
38. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics 2017;37:505-15.
39. Jamieson A, Giger ML, Drukker K, Li H, Yuan Y, Bhooshan N. Exploring non-linear feature space dimension reduction and data representation in breast CADx with Laplacian eigenmaps and t-SNE. Med Phys 2010;37:339-51.
40. Jamieson AR, Giger ML, Drukker K, Pesce L. Enhancement of breast CADx with unlabeled data. Med Phys 2010;37:4155-72.
41. Parma C, Grossmann P, Bussink J, Lambin P, Aerts H. Machine learning methods for quantitative radiomic biomarkers. Sci Rep 2015;5:13087.
42. Huynh B, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Med Imaging 2016;3:034501.
43. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221-48.
44. Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA. Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. Med Phys 1994;21:517-24.
45. Sahiner B, Chan HP, Petrick N, et al. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. IEEE Trans Med Imaging 1996;15:598-610.

46. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 2012;25:1097-105.
47. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014.
48. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. CoRR abs/1409.4842. 2014.
49. Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2010;22:1345-9.
50. Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks? CoRR abs/1411.1792 2014.
51. Razavian AS, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE; 2014:512-9.
52. Donahue J, Jia Y, Vinyals O, et al. Decaf: a deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531. 2013.
53. Roth H, Lu L, Liu J, et al. Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans Med Imaging 2016;35:1170-81.
54. Anthimopoulos M, Christodoulidis S, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 2016;35:1207-16.
55. Samala R, Chan H-P, Hadjiiski L, Helvie MA, Wei J, Cha K. Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. Med Phys 2016;43:6654-66.
56. Gulshan V, Peng L, Coram M, Stumpe MC, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402-10.